REGULAR ARTICLE

blood advances

Transcriptomic analysis reveals proinflammatory signatures associated with acute myeloid leukemia progression

Svea Stratmann,[1,*] Sara A. Yones,[2,*] Mateusz Garbulowski,[2] Jitong Sun,[1] Aron Skaftason,[3] Markus Mayrhofer,[4] Nina Norgren,[5] Morten Krogh Herlin,[6,7] Christer Sundström,[1] Anna Eriksson,[8] Martin Höglund,[8] Josefine Palle,[9] Jonas Abrahamsson,[10] Kirsi Jahnukainen,[11] Monica Cheng Munthe-Kaas,[12,13] Bernward Zeller,[13] Katja Pokrovskaja Tamm,[14] Lucia Cavelier,[1] Jan Komorowski,[2-17] and Linda Holmfeldt[1,18]

[1]Department of Immunology, Genetics and Pathology and [2]Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden; [3]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden; [4]National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Uppsala University, Uppsala, Sweden; [5]Department of Molecular Biology, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Umeå University, Umeå, Sweden; [6]Department of Clinical Medicine and [7]Department of Pediatrics and Adolescent Medicine, Aarhus University, Aarhus, Denmark; [8]Department of Medical Sciences and [9]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden; [10]Department of Pediatrics, Institute of Clinical Sciences, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden; [11]Children's Hospital, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland; [12]Norwegian Institute of Public Health, Oslo, Norway; [13]Division of Pediatric and Adolescent Medicine, Oslo University Hospital, Oslo, Norway; [14]Department of Oncology and Pathology, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden; [15]Swedish Collegium for Advanced Study, Uppsala, Sweden; [16]Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland; [17]Washington National Primate Research Center, Seattle, WA; and [18]The Beijer Laboratory, Uppsala, Sweden

Key Points

- Progression of AML is associated with proinflammatory mediators through altered expression levels of *CR1, DPEP1, IL1R1*, and *ST18.*
- Upregulated *CD6* and downregulated *INSR* are nodes in gene expression networks linked to AML relapse, according to machine learning analysis.

Numerous studies have been performed over the last decade to exploit the complexity of genomic and transcriptomic lesions driving the initiation of acute myeloid leukemia (AML). These studies have helped improve risk classification and treatment options. Detailed molecular characterization of longitudinal AML samples is sparse, however; meanwhile, relapse and therapy resistance represent the main challenges in AML care. To this end, we performed transcriptome-wide RNA sequencing of longitudinal diagnosis, relapse, and/or primary resistant samples from 47 adult and 23 pediatric AML patients with known mutational background. Gene expression analysis revealed the association of short event-free survival with overexpression of *GLI2* and *IL1R1,* as well as downregulation of *ST18*. Moreover, *CR1* downregulation and *DPEP1* upregulation were associated with AML relapse both in adults and children. Finally, machine learning–based and network-based analysis identified overexpressed *CD6* and downregulated *INSR* as highly copredictive genes depicting important relapse-associated characteristics among adult patients with AML. Our findings highlight the importance of a tumor-promoting inflammatory environment in leukemia progression, as indicated by several of the herein identified differentially expressed genes. Together, this knowledge provides the foundation for novel personalized drug targets and has the potential to maximize the benefit of current treatments to improve cure rates in AML.

Submitted 12 April 2021; accepted 23 August 2021; prepublished online on *Blood Advances* First Edition 7 October 2021; final version published online 7 January 2022. DOI 10.1182/bloodadvances.2021004962.

*S.S. and S.A.Y. contributed equally to this work.

Custom codes are available from the authors upon request. RNA-seq data are available under controlled access via doi.org/10.17044/scilifelab.13105229. Validation of differential gene expression results are based on data generated by The Cancer Genome Atlas Research Network (https://www.cancer.gov/tcga) and the Therapeutically Applicable Research to Generate Effective Treatments initiative (https://ocg.cancer.gov/programs/target), phs000218 (data available through: https://portal.gdc.cancer.gov/projects).

Requests for data sharing may be submitted to Linda Holmfeldt (linda.holmfeldt@igp.uu.se)

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

Introduction

Acute myeloid leukemia (AML) is believed to arise through a combination of genetic alterations and aberrant gene expression patterns caused by genetic and epigenetic changes, which in their composition also determine AML progression and therapy resistance. The AML inter- and intra-tumor heterogeneity at disease onset has been intensively investigated, resulting in improved disease classification[1] and novel treatment alternatives,[2-4] leading to complete remission in the majority of patients. Nevertheless, 40% to 60% of adults and 30% to 40% of children relapse within 3 years,[1,5-7] and these relapsed patients often fail to respond to conventional treatment. Together, this results in 5-year overall survival (OS) rates at only 28% and 70% for adults and children, respectively.[8,9]

Prognostic validation and treatment allocation in AML are currently based on morphologic, cytogenetic, and genetic features. However, risk stratification, including relapse prediction, remains challenging, especially for patients without causative genetic aberrations. To investigate novel treatment options that target each individual tumor, it is necessary to identify aberrant pathways that drive tumor progression and therapy resistance. RNA-sequencing (RNA-seq) provides a comprehensive picture of the cellular transcriptome, combining detection of various mutations and gene fusions with gene expression analysis. Furthermore, single-cell RNA-seq has emerged as a powerful tool for identifying cellular subgroups with independent characteristics.[10,11] Previous transcriptomic studies in AML have mainly been focused on gene expression signatures at diagnosis and their predictive potential.[12-15] Only a limited number of studies have investigated differential expression in relapsed and primary resistant (R/PR) AML, and most of these lacked patient-matched longitudinal samples and/or detailed knowledge about underlying genetic alterations in the form of whole-genome sequencing (WGS) or whole-exome sequencing (WES) data.[16,17] To the best of our knowledge, the largest published transcriptomic studies to date on patient-matched diagnosis/relapse samples comprised 24 adult[18] and 23 pediatric[19] AML cases; both of these studies were either partially or entirely based on microarrays, which do not allow for detection of various molecular aberrations.

Here we report an RNA-seq–based analysis, incorporating interpretable machine learning techniques, on longitudinal samples from 70 R/PR AML cases, all of which previously have been characterized by WGS or WES. We identified differentially expressed genes (DEGs) specific for relapse in AML (*CR1* and *DPEP1*). Furthermore, another set of DEGs was associated with short event-free survival (EFS; ie, *GLI2, IL1R1*, and *ST18*). Finally, rule networks derived from machine learning–based analysis were investigated, enabling the detection of additional putative drivers of leukemia progression and therapy resistance.

Patients and methods

Patient and control samples

Cryopreserved sequential AML samples from 47 adult patients and 23 pediatric patients with AML from the Nordic countries were included in this study. Criteria for selecting cases were as follows: available relapse or PR material of sufficient quality and yield via Uppsala Biobank and Karolinska Institute Biobank, collected between 1995 and 2016. Acute promyelocytic leukemia cases were excluded. Patients were clinically characterized according to the World Health Organization criteria,[20] and all samples were previously analyzed on the genomic level via WGS or WES[21] (supplemental Table 1). Sixty-three of the patients had de novo AML; 7 patients had either a prior diagnosis of myelodysplastic syndromes, therapy-related AML, or therapy-related myelodysplastic syndromes. The median length of EFS for relapse cases was 16.3 months (range, 1.1-126.0 months) for adults and 11.0 months (range, 2.3-33.6 months) for children (Table 1). Detailed biological characteristics and supporting clinical information are reported in supplemental Tables 2 and 3. CD34-expressing bone marrow (BM) cells from 5 healthy donors (AllCells Inc., Alameda, CA) were used as normal controls (supplemental Table 4).

The study was approved by the Uppsala Ethical Review Board (Sweden) and the Regional Ethical Committee South-East (Norway). Informed consent was obtained from all patients or their legal guardians according to the Declaration of Helsinki.

For validation purposes, RNA-seq data from another adult (The Cancer Genome Atlas [TCGA][13]; phs000178) and pediatric (Therapeutically Applicable Research to Generate Effective Treatments [TARGET][12]; phs000465) AML cohort were used.

Sample preparation

Mononuclear cells from BM aspirates or peripheral blood were isolated through Ficoll gradient centrifugation and cryopreserved until use. AML samples with leukemia cell content below 80% and sufficient amount of starting material were purified by immune-based depletion of nontumor cells (supplemental Tables 2 and 5). Total RNA was extracted via the AllPrep DNA/RNA/Protein Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions, incorporating DNase I treatment.

Transcriptome sequencing

Library preparation (Illumina TruSeq Stranded total RNA [ribosomal depletion] libraries) and RNA-seq (HiSeq2500 and/or NovaSeq6000, Illumina, San Diego, CA) were conducted by the SNP&SEQ Technology Platform, Science for Life Laboratory (SciLifeLab) (National Genomics Infrastructure, Uppsala, Sweden). Gene counts, gene fusions, single nucleotide variants, and small insertions and deletions (<50 base pairs) were retrieved and processed (detailed further in the supplemental Methods).

RNA-seq analysis

Differential gene expression analysis was conducted by using Qlucore Omics Explorer 3.6 (Qlucore AB, Lund, Sweden). In brief, read counts were filtered toward expressed, protein-coding genes and normalized by the trimmed mean of M values (TMM[22]) followed by normalization to the gene length. The Benjamini-Hochberg method was used to correct for multiple testing, and the fold change (FC) was calculated from the difference between the arithmetic averages over each group, based on log2-transformed normalized values. Furthermore, Gene Ontology (GO) enrichment analysis was performed by using Gene Ontology Enrichment Analysis and Visualization.[23,24]

Machine learning–based analysis

In preparation for machine learning–based analysis, zero variance genes across the leukemia samples were excluded. The expression levels for each of the remaining genes were discretized into 3 equal

Table 1. Patient cohort

Characteristic	Value
No. of patients, n (%)	70 (100)
Adult cases	47 (67.1)
Elderly (aged ≥60 y)	25 (35.7)
Adult (aged 40-59 y)	16 (22.9)
Young adult (aged 19-39 y)	6 (8.6)
Pediatric cases	
Adolescent (aged 15-18 y)	2 (2.9)
Child (aged 3-14 y)	14 (20.0)
Infant (aged <3 y)	7 (10.0)
Sex, female	37 (52.9)
Background	
De novo AML	63 (90.0)
Potential t-AML	3 (4.3)
MDS-AML	2 (2.9)
t-MDS-AML	2 (2.9)
No. of tumor samples, n (%)	122 (100)
Diagnosis	43 (35.2)
Relapse	73 (59.9)
R1 and R1-P	57 (46.7)
R2 and R2-P	13 (10.7)
R3	3 (2.5)
Primary resistant	6 (4.9)
Average age at onset, y	
Adult cases	59.5 (range, 20.5-83.1; median, 62.2)
Pediatric cases	7.7 (range, 0.4-17.5; median, 7.3)
Average length of EFS (D>R1), d	
Adult relapse cases	497 (range, 34-3844; median, 305.0)
Pediatric relapse cases	334 (range, 69-1026; median, 304.5)
Average length of OS, d	
Adult relapse cases	1109 (range, 45-8270; median, 509)
Pediatric relapse cases	1682 (range, 126-6557; median, 572)
Sample purity*	89% (>80% tumor cells; range, 41-100)
Cell viability	63% (≥75% viable cells; range, 10-94)
Average RIN	9.2 (range, 5.8-10.0; median, 9.3)
Sampling duration	1995-2016

Detailed biological and clinical data for each patient/sample are presented in supplemental Tables 2 and 3. D, diagnosis; EFS, EFS as time to first relapse; MDS, myelodysplastic syndromes; OS, OS as time to death or last follow-up; R1/2/3, sequential relapses; R1/2-P, persistent relapse specimen; RIN, RNA integrity number; t-AML, treatment-related AML.

*Single-nucleotide polymorphism–based calculation.

bins according to low, medium, and high gene expression. To create a ranking of the most informative features (ie, genes) that distinguish between diagnosis and relapse, the Monte Carlo Feature Selection (MCFS[25]; rmcsf v.1.2.6) algorithm was applied. Subsequently, interpretable machine learning models were built by using the R.ROSETTA R-package[26] version 2.2.9. That algorithm is based on the rough set theory. Following this supervised approach, classifiers are transparent, as the predictive model is constructed from a set of IF-THEN rules. Rules that represented copredictive mechanisms between genes were visualized as networks by using VisuNet R-package[27] version 1.3.5. Finally, for investigating similarities among cohorts, machine learning analysis was performed by using merged feature lists obtained from MCFS. These models were used for network comparison and discovery of copredictive genes visible in networks as highly connected nodes, also called hubs. Additional details are given in the supplemental Methods.

Statistics

Statistical tests were conducted by using GraphPad Prism version 7.02 and version 9.0.2 (GraphPad Software, La Jolla, CA) or R version 3.6 (R Foundation for Statistical Computing, Vienna, Austria). $P < .05$ was defined as statistically significant unless otherwise stated.

Results

We performed RNA-seq on 122 samples from 70 patients with AML, all of whom relapsed or had PR disease. These comprised samples collected at diagnosis (n = 43) and relapse (n = 73), as well as PR samples (n = 6) (Table 1; supplemental Table 1). CD34-expressing BM cells from 5 healthy donors were used as a source of normal control RNA ("BM-controls"). RNA-seq yielded an average of 41 million reads per sample (supplemental Tables 2 and 6). The composition of genomic alterations for all AML samples had previously been analyzed via WGS or WES.[21] Fifty-seven percent of the reported somatic protein-coding mutations could be validated via RNA-seq, as detailed in the supplemental Results, supplemental Tables 7 and 8, and supplemental Figure 1.

Recurrent gain of fusion transcripts during tumor progression

In addition to protein-coding mutations, we identified 26 and 23 gene fusions at diagnosis and/or R/PR in adult and pediatric cases, respectively (supplemental Tables 8 and 9). These comprised recurrent common AML-associated fusions, with translocations leading to the gene fusion *RUNX1-RUNX1T1* (adult, n = 4 cases; pediatric, n = 4 cases) and *NUP98* fusions (adult, n = 1; pediatric, n = 4) among the most frequent events (Figure 1). In addition, we detected previously unreported fusions involving known cancer-related genes, including, for instance, *FOS-PSAP*, *SRSF3-PLAG1*, *CEBPE-CEBPA*, and *REXO1-NF1*, with the latter leading to a frameshift. Transcribed gene fusions were mainly stable or gained over the course of the disease, with 37.5% of the fusions appearing during leukemia progression, including gain of *BCR-ABL1* (n = 2) (supplemental Figures 2A and 3), *RUNX1-RUNX1T1* (n = 1) (supplemental Figure 2B), and *SRSF3-PLAG1* (n = 1). In adults, 16 (61.5%) of 26 fusions were also determined by WGS, whereas 15 of 23 (65.2%) of the pediatric fusions were detected by both RNA-seq and WGS.[21] Of the remaining 18 gene fusion transcripts in adult and pediatric cases combined, 14 were undetected via WGS. The final 4 (eg, *NUP98-NSD1*) were identified in samples that at the DNA level were analyzed via WES, which is known to rarely be able to detect structural variants. On the contrary, out-of-frame *ETV6* and *NF1* transcripts predicted by WGS to be generated through genomic translocations could not be detected at the RNA level, potentially due to nonsense-mediated decay.

Gene expression profiling in R/PR AML

Although most types of genomic alterations can be detected via both WGS and RNA-seq, the latter holds the benefit of also containing gene expression information. As an initial step to investigate potential differences among the expression of protein-coding genes, we applied unsupervised principal component analysis on the entire cohort (supplemental Figure 4). Patient AML008 represents an adult hypodiploid AML case with a partial or total loss of 10 different chromosomes, resulting in a highly distinct transcriptome compared with the rest of the cohort (supplemental Figure 4A). Longitudinal samples from this patient were thus excluded from downstream analyses. As a likely consequence of the well-known heterogeneity in AML, unsupervised clustering of the remaining cases revealed no significant separation among the tumor samples. Sequential tumor samples from the same patient, however, were mainly grouped together by unsupervised clustering, whereas all BM-control samples formed a distinct group.

Pro-inflammatory signatures are associated with short EFS

We next investigated differences in the expression patterns between various subgroups within the cohort. To obtain initial information about poor outcome-associated transcriptomic alterations in AML, factors characterizing time of EFS were examined. Here, we focused solely on diagnosis samples, with the adult and pediatric cases combined to gain statistical power (short EFS, n = 18 [adults <0.5 year; children <1.0 year]; long EFS, n = 24) (supplemental Tables 8 and 10A). This comparison identified 996 DEGs ($P <$ .05) (supplemental Table 11; supplemental Figure 5). Short EFS was associated with 222 upregulated and 168 downregulated genes based on a |log2FC|>1 cutoff, with *GLI2* (|log2FC|=3.7), *IL1R1* (|log2FC|=2.2), and *ST18* (|log2FC|=2.9) found among the highest ranked genes (Figure 2). GO enrichment analysis of genes with higher expression in short EFS-associated samples revealed an overrepresentation of genes in pathways related to immune response (eg, *IL1R1*), regulation of cell differentiation (eg, *GLI2*), and exocytosis (Figure 2D; supplemental Table 12).

GLI2, a mediator of sonic hedgehog (Shh) signaling, is involved in the maintenance of normal and neoplastic hematopoietic stem cells,[28] and it plays an important role in the crosstalk between tumor cells and their microenvironment.[29] Overexpression of this gene was associated with short EFS (P_{EFS} = .0001) (Figure 2B). Furthermore, Kaplan-Meier analysis showed that higher *GLI2* levels correlated with poorer OS (P_{OS} = .045; 5-year OS $GLI2_{low}$, 38.5%; $GLI2_{high}$, 12.5%) (Figure 2C; supplemental Table 13). The correlation between higher *GLI2* expression and adverse outcome was via Kaplan-Meier analysis independently validated for both adult and pediatric AML by the TCGA[13] and TARGET[12] data sets, respectively (supplemental Figure 6). As previously reported,[30] *GLI2* overexpression was pronounced in AML positive for *FLT3*–internal tandem duplication (ITD) (supplemental Figure 7A). In our study, however, a link between elevated expression of this gene and poor outcome was seen also when excluding *FLT3*-ITD–positive samples from the differential expression and Kaplan-Meier analyses (supplemental Figure 7B-E).

High expression of *IL1R1* was associated with short EFS and OS (P_{EFS} = .0096; P_{OS} = .016; 5-year OS $IL1R1_{low}$, 47.4%; $IL1R1_{high}$, 13.0%) (Figure 2B-C; supplemental Table 13). Again, these observations could be independently validated by Kaplan-Meier analysis on the TCGA and TARGET cohorts (supplemental Figure 6). *IL1R1* encodes an interleukin-1 receptor involved in regulation of inflammatory and immune responses.[31]

ST18 encodes a zinc finger transcription factor associated with negative regulation of proliferation, apoptosis, and inflammation.[32,33] Expression of this gene was lower among samples associated with short EFS as well as with shorter OS in all 3 cohorts (local cohort, P_{EFS} = .027; P_{OS} = .017; 5-year OS $ST18_{low}$, 13.6%; $ST18_{high}$, 45.0%) (Figure 2B-C; supplemental Table 13; supplemental Figure 6). Elevated *ST18* expression, especially in cases harboring the good prognosis-associated inversion on chromosome 16 [inv(16)[34]], has previously been suggested as a marker for monitoring of pediatric AML minimal residual disease.[35] In the current study, AML samples carrying inv(16) showed exceptionally high *ST18* expression, albeit all of these cases eventually relapsed (supplemental Figure 6A).

CR1 and *DPEP1* are differentially expressed at relapse compared with patient-matched pretreatment samples

To gain insight into the gene expression programs preferentially active during AML progression, we explored gene expression differences between sequential tumor samples. The comparison of patient-matched diagnosis and relapse samples (adult, n = 22 pairs; pediatric, n = 17 pairs) (supplemental Tables 8 and 10B) resulted in 405 and 212 DEGs ($P <$ .05) identified in adults and children, respectively (supplemental Table 14; supplemental Figure 8A). Among these, 4 overlapping genes were found between the adult and pediatric cases with |log2FC|>1 (*CR1, DPEP1, MREG,* and *SHANK3*) (Figure 3A; supplemental Figure 8B). DEGs associated with relapse clustered in pathways predominantly related to immune response (eg, *CR1*) and cellular metabolic processes (eg, *DPEP1*) (supplemental Table 15; supplemental Figure 9A).

CR1 (also called C3b/C4b-receptor or *CD35*) encodes a complement receptor present on most circulating blood cells. Upon binding of its ligands (C3b/C4b opsonins) and ligand degradation, it acts as linkage between the innate and adaptive immune system.[36,37] Furthermore, CR1 negatively regulates complement activity by inhibiting C3 convertase.[38] We detected significantly lower expression of *CR1* at relapse compared with patient-matched diagnosis samples in adults (|log2FC|=1.8; P = .0001) and children (|log2FC|=1.6; P = .023) (Figure 3B; supplemental Figure 9B). Decreased expression of *CR1* over the course of the disease has previously been reported for chronic myeloid leukemia[39] but has not been described for AML.

Higher expression of *DPEP1* was seen at relapse both in adults (|log2FC|=1.3; P = .0008) and in children (|log2FC|=2.5; P = .0038) (Figure 3C; supplemental Figure 9C). *DPEP1* encodes dipeptidase one, a zinc-dependent metalloproteinase, involved in the regulation of apoptosis, inflammation, and cell migration.[40,41]

Interpretable machine learning for identification of copredictive biomarkers in AML

To find the most informative transcriptomic features characteristic for AML diagnosis vs relapse, we applied feature selection using the MCFS algorithm[25] (supplemental Figure 10). After selecting the

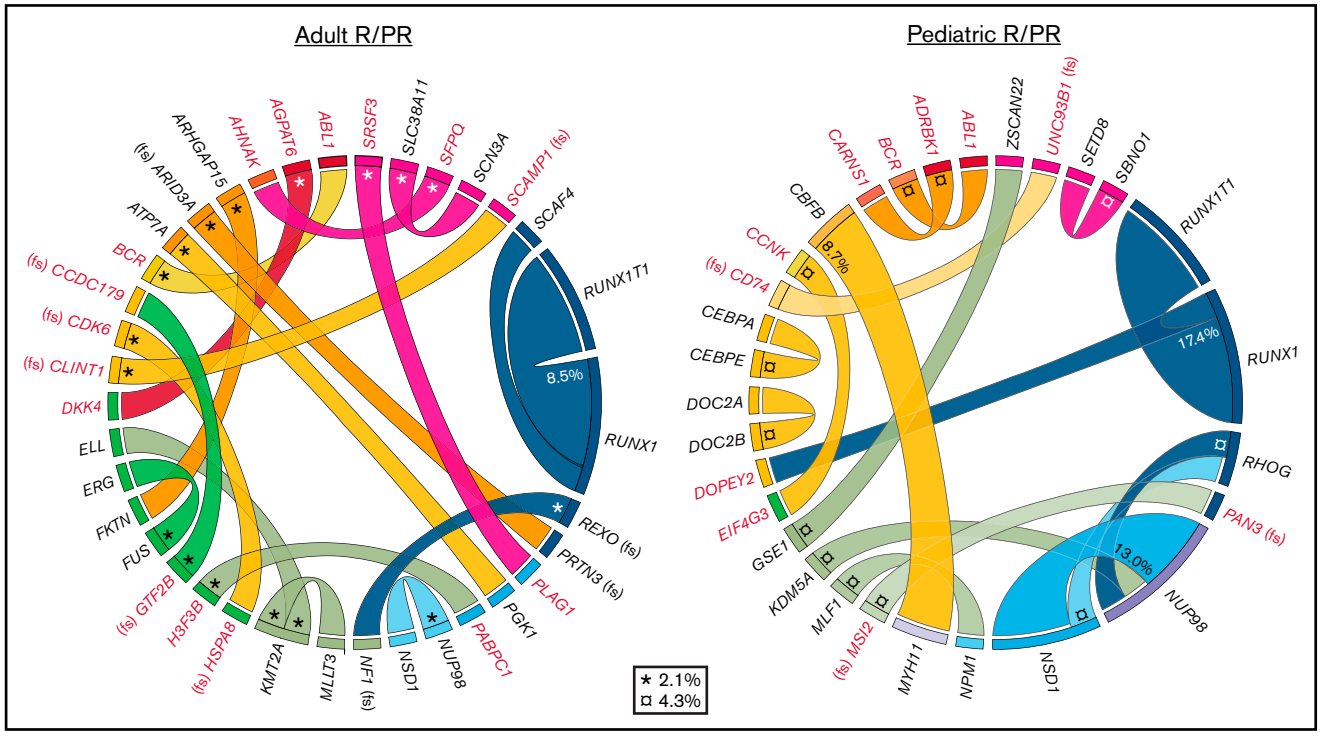

Figure 1. Fusion transcripts in adult and pediatric relapsed AML. Circos plots presenting gene fusion transcripts detected at relapse and in PR samples in adult (left) and pediatric (right) AML cases. Ribbon widths are proportional to the frequency of a fusion event among the respective patient cohort. Fusion transcripts that were gained at relapse or in PR samples are highlighted in red. In addition, one of the *RUNX1-RUNX1T1* fusions in pediatric AML was solely detected at relapse. *2.1%; ¤4.3%; fs, frameshift.

most important features from high-dimensional data, interpretable machine learning was performed by using the R.ROSETTA framework.[26] Subsequently, rule-based models were constructed, and copredictive features for each data set were estimated.

The trained models had a mean accuracy of 78% for fivefold cross-validation applied on the adult cohort and 90% for threefold cross-validation applied on the pediatric cohort ($P < .05$). The model based on the adult cohort identified *CD6* as the feature most commonly differentiating diagnosis from relapse, with overexpression of this gene at relapse (supported by 22 samples; 92% accuracy; $P = .00051$) (Figure 4; supplemental Table 16). This finding was verified by differential expression analysis (supplemental Figure 11A). *CD6* encodes a lymphoid-associated surface glycoprotein that is involved in cell adhesion and plays a role in immune synapse formation.[42]

Furthermore, *INSR* was found to be (via rule-based modeling) downregulated at relapse in adult AML, where it served as a node gene. This node was connected to, for instance, overexpression of *CD6* (supported by 18 samples; 95% accuracy; $P = .00094$) and upregulation of the transcription factor–encoding gene *ZNF773* (supported by 16 samples; 100% accuracy; $P = .00034$) compared with expression levels at diagnosis (Figure 4; supplemental Table 16; supplemental Figure 11). Of note, low *INSR* expression at diagnosis was associated with poor outcome in the TCGA cohort (supplemental Figure 12), in part supporting the findings generated through our relapse cohort that low *INSR* expression was associated with disease progression.

Likely due to the relatively small number of pediatric samples, the machine learning model was unable to detect strong node connections associated with pediatric AML relapse (supplemental Table 17; supplemental Figure 13). To overcome this obstacle, we merged the features identified via MCFS run independently on our pediatric cohort and on additional diagnosis (n = 20) and relapse (n = 38) samples available via TARGET, followed by generation of rule-based models (supplemental Tables 18 and 19; supplemental Figure 14). Relapses from our pediatric cohort and the TARGET cohort clustered together in a network comparison (Figure 5A). Distinct patterns of, for instance, lower expression of *NFATC4* and *KATNAL2* at diagnosis that were restored to normal at relapse compared with the BM-controls were found predictive in the merged setting (Figure 5; supplemental Figure 15). *NFATC4* encodes a transcription factor involved in cell quiescence,[43] whereas the *KATNAL2* gene product regulates microtubule stability.[44]

Discussion

Current diagnostic and prognostic tools in AML mainly rely on cytogenetic and genetic approaches, including only a few genes with targeted therapeutic potential. RNA-seq is beneficial because it can be used both for detection of various transcribed genetic alterations and for gene expression analysis. The identification of common dysregulated pathways among different patients with AML, independent of their mutational background, may provide novel options for therapeutic intervention. Furthermore, RNA-seq–based studies of R/PR AML cohorts with known genetic background are largely missing. Here, we show the importance of including RNA-seq for detailed

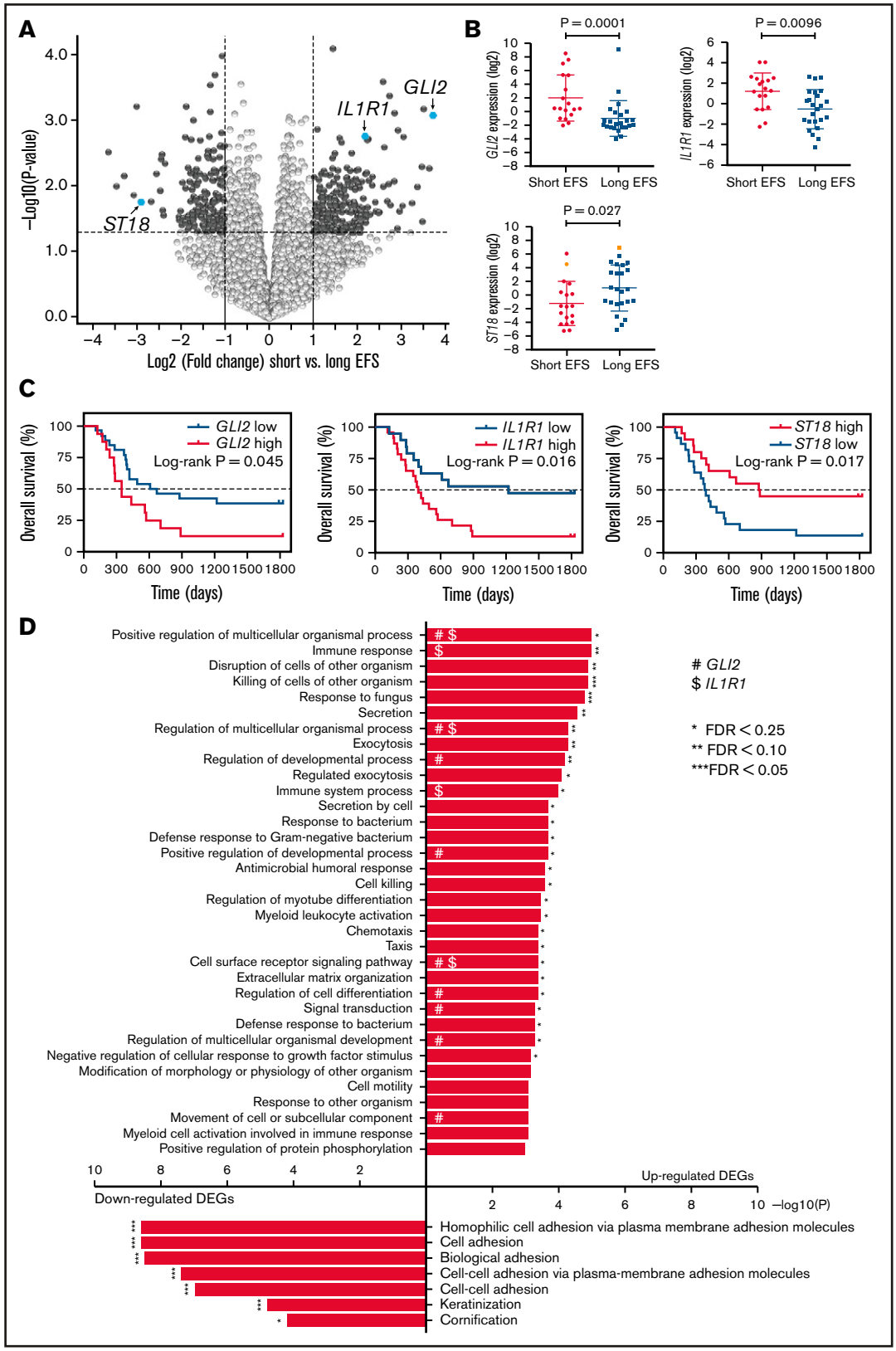

Figure 2. Differential gene expression between AML samples associated with short vs long EFS. (A) Volcano plot showing DEGs with genes downregulated (log2FC<0) and upregulated (log2FC>0), respectively, in diagnosis samples associated with short EFS compared with long EFS-associated diagnosis samples. Highly ranked genes are highlighted in dark gray ($P < .05$; |log2FC|>1), with *GLI2, IL1R1,* and *ST18* highlighted in blue. Visualization and underlying statistical calculations were

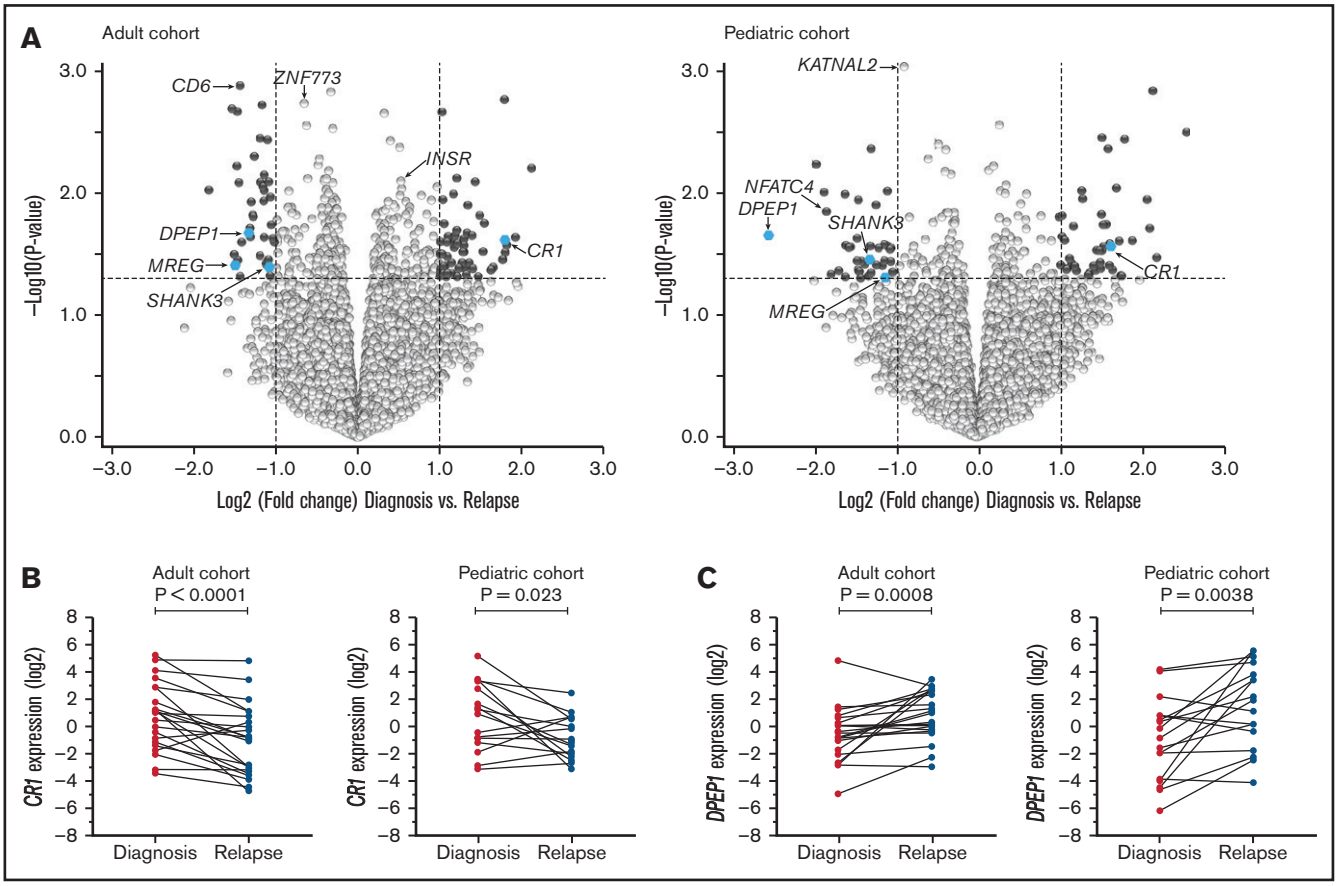

Figure 3. Differential gene expression between paired diagnosis and relapse samples. (A) Volcano plots presenting DEGs with genes downregulated (log2FC<0) and upregulated (log2FC>0), respectively, at diagnosis compared with relapse samples, for adult and pediatric cases separated. Highly ranked genes are highlighted in dark gray ($P < .05$; |log2FC|>1) or blue (highly ranked in both the adult and pediatric R/PR AML cohorts). Visualization and underlying statistical calculations were performed by using Qlucore Omics Explorer version 3.6. (B-C) Spaghetti plots displaying gene expression data comparing patient-matched diagnosis and relapse samples for *CR1* (B) and *DPEP1* (C). The y-axis represents log2-transformed, TMM-normalized expression values. Applied statistical test, Wilcoxon matched-pairs test. Supplemental Table 10B provides details regarding samples included in this figure, and supplemental Table 14 provides details for all DEGs.

analysis of R/PR AML to improve the understanding of molecular alterations promoting leukemia progression and therapy resistance.

RNA-seq allows for fusion detection, whereas fusions often are impossible to detect by WES. Using RNA-seq data, we verified all genomic aberrations leading to in-frame gene fusions previously detected by WGS in the respective AML samples.[21] Moreover, we detected 18 additional fusion transcripts that were not identified through WGS/WES. In 2 treatment-resistant cases, *BCR-ABL1* fusions were gained during leukemia progression, with one of these fusions only detected according to RNA-seq. This highlights the importance of screening for *BCR-ABL1* fusions at AML relapse to allow for utilization of tyrosine kinase inhibitors as optional treatment.

Shh signaling regulates normal hematopoiesis by controlling the frequency and maintenance of hematopoietic stem cells, whereas aberrant activation promotes myeloid leukemia progression and dormant leukemia stem cells.[28,45] *GLI1* and *GLI2* encode mediators of the Shh pathway and have previously been shown to be dysregulated in AML. Higher messenger RNA levels of *GLI1* have been linked to MYC and BCL2 activation, as well as to relapse and resistant disease in AML.[46] *GLI2* was previously found to be

Figure 2. (continued) performed by using Qlucore Omics Explorer version 3.6. (B) Scatter plots with mean and SD illustrating the log2-transformed, TMM-normalized expression values in samples associated with short vs long EFS for *GLI2*, *IL1R1*, and *ST18*. Applied statistical test, Mann-Whitney test. Samples highlighted in orange in the scatter plot illustrating *ST18* harbor an inversion on chromosome 16, leading to a *CBFB-MYH11* gene fusion. (C) Kaplan-Meier plots showing 5-year OS for cases with low expression (blue lines) and high expression (red lines) of *GLI2*, *IL1R1*, and *ST18* at diagnosis. The average expression value for the respective gene over all samples included in the analysis was used to discretize between low and high expression. *P* values were calculated by using the log-rank (Mantel-Cox) test. (D) GO analysis of DEGs between short vs long EFS-associated samples. GO terms presented above the x-axis are enriched among genes upregulated in samples associated with short EFS, whereas pathways below the x-axis are enriched among downregulated genes. Short EFS was considered as <0.5 year for adults and <1.0 year for pediatric patients. Supplemental Table 10A presents details regarding samples included in this figure, supplemental Table 11 presents details for all DEGs, and supplemental Table 13 presents details regarding statistical results associated with panel C. *False discovery rate (FDR) < 0.25, **FDR < 0.1, ***FDR < 0.05 (Benjamini-Hochberg correction). #*GLI2*, $*IL1R1*. SD, standard deviation.

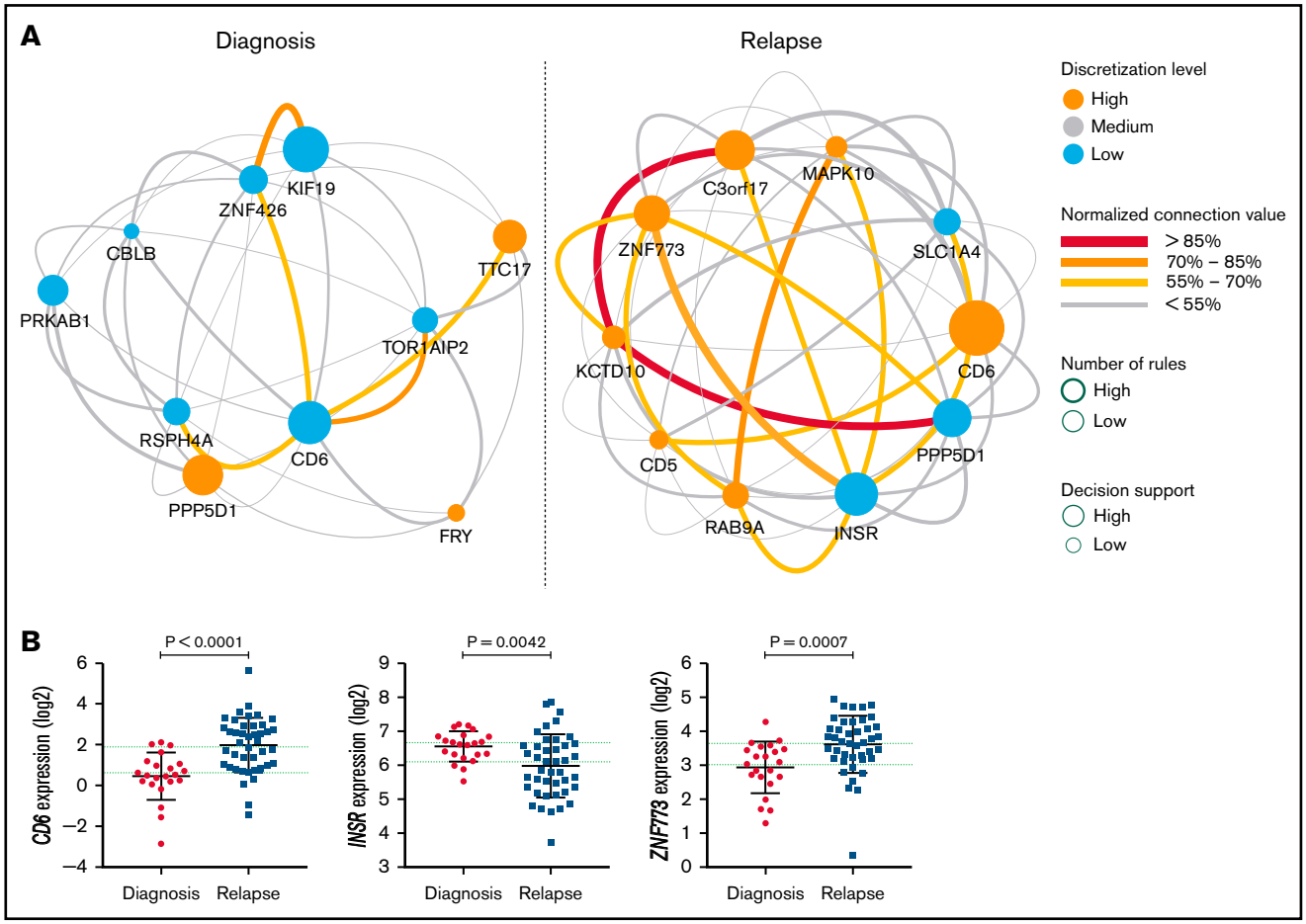

Figure 4. Predictive features for relapse in adult AML detected by machine learning analysis. (A) Relationships between copredictive features associated with diagnosis (left) and relapse (right) among adult AML cases are visualized by using VisuNet. The color of the nodes shows the expression level, with 3 bins for high (orange), medium (gray), and low (blue) expression. The rule support is indicated by the size of the respective node, whereas the support for each connection is visualized by the thickness and color of the connective line. Rules were filtered according to a false discovery rate <0.05. (B) Scatter plots with mean and standard deviation illustrating the log2-transformed, TMM-normalized expression values in diagnosis and unpaired relapse samples for *CD6*, *INSR*, and *ZNF773*. The borders of the 3 bins, corresponding to low, medium, and high expression of the respective genes, are indicated by green grid lines on the y-axis. Applied statistical test, Mann-Whitney test. Supplemental Table 10C provides details regarding samples included in this figure.

overexpressed predominantly in *FLT3*-ITD AML and associated with adverse outcomes.[30,47] We found that elevated *GLI2* expression, independent of *FLT3*-ITD status, correlates with shorter EFS and OS in AML (supplemental Figure 7). Altogether, this suggests that *GLI1/2* overexpression may give rise to an increased number of leukemia stem cells, known to be resistant to chemotherapeutic drugs. Glasdegib, an Shh inhibitor, has been approved for the treatment of patients with R/PR AML ineligible for standard treatment.[48] Our data (supplemental Figure 6) suggest that patients with elevated *GLI2* expression at diagnosis could benefit from Shh inhibitors in combination with standard treatment. Moreover, arsenic trioxide, a drug approved by the US Food and Drug Administration for the treatment of acute promyelocytic leukemia, reportedly inhibits GLI1 and GLI2.[49]

GO enrichment analysis revealed overrepresentation of pathways involved in immune response in the DEG analysis investigating associations with time of EFS as well as relapse. The dysregulation of the immune system via evasion of destruction and/or tumor-promoting inflammation are hallmarks of cancer.[50] *IL1R1* encodes a key regulator of inflammation and immune response and promotes tumor progression by activating numerous tumor-beneficial pathways such as NF-κB and MAPK, as well as antiapoptotic signaling.[31] Similar to these former findings, we report here an association between high *IL1R1* transcript levels and short EFS as well as OS in AML (Figure 2).

Downregulation of *ST18* has previously been reported to promote tumorigenesis through tumor-promoting regulation of inflammation.[32,33] Low *ST18* expression was associated with short EFS and OS in our cohort as well as in 2 independent validation cohorts (TCGA and TARGET) (Figure 2; supplemental Figure 6). The correlation between downregulated *ST18* and cancer was first reported for breast cancer, resulting from promoter hypermethylation detected in 80% of breast cancer samples.[32] Subsequent studies revealed ST18 as a negative regulator of tumor growth, inflammation, and apoptosis.[32,33] Altogether, our data suggest that decreased *ST18* expression can potentially be used as a biomarker in AML.

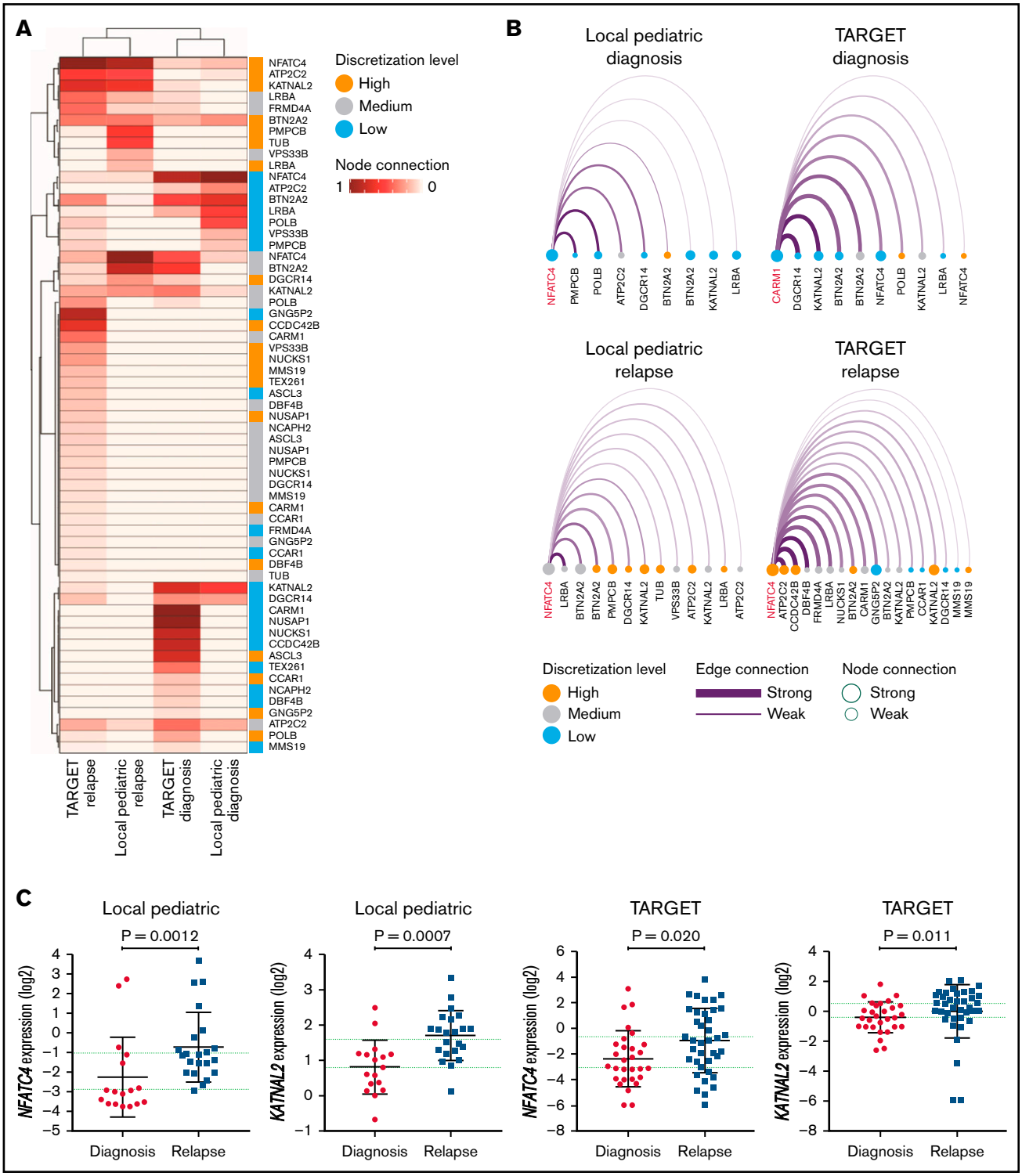

Figure 5. Predictive features for pediatric AML relapse detected through network comparison. (A) Heat map showing clustering of nodes and diagnosis and relapse sample groups from the local pediatric and TARGET cohorts based on their connection strength from networks, with the top 50 nodes pruned from the networks. (B) Arc diagrams representing the topmost connected node gene (highlighted in red; also referred to as a hub) and all of its connections for a given rule-based network. Each arc illustrates a connection between a hub gene and an associated gene, with the connection initially derived from a rule. Nodes are decreasingly sorted from the left by the edge connection value. The full networks are depicted in supplemental Figure 14. (C) Scatter plots with mean and standard deviation illustrating the log2-transformed, TMM-normalized expression values in diagnosis and unpaired relapse samples from the local pediatric and TARGET cohorts for *NFATC4* and *KATNAL2*. The borders of the 3 bins, corresponding to low, medium, and high expression for the respective genes, are indicated by green grid lines. Applied statistical test, Mann-Whitney test. Supplemental Table 10C-D provides details regarding samples included in this figure.

Former studies report the establishment of a pro-inflammatory microenvironment in hematopoietic organs through complement activation, after chemotherapy or radiation treatment.[51] Complement activation promotes tumor growth and oncogenesis.[50] Here, we found downregulation of *CR1/CD35*, encoding a negative regulator of the complement system, at AML relapse both in adults and children (Figure 3B). These results highlight the pronounced role of a tumor-promoting environment during disease progression in AML. The functional implication of CR1 in relapsed AML has yet to be elucidated, however.

Taken together, our data provide a rationale for targeting key regulators of pro-inflammatory pathways in AML. Inhibition of the IL1R1 signaling pathway has been shown to increase drug sensitivity and prolong EFS,[52,53] whereas numerous drugs targeting complement activation are under investigation. Although targeting immune regulators in AML holds great potential, the optimal point of intervention remains to be established.

Our analysis of DEGs during AML progression revealed upregulation of *DPEP1* at relapse both in adult and pediatric AML. *DPEP1* overexpression is best known as an adverse prognostic factor in colorectal cancer.[54] In addition, increased proliferation, survival, and resistance upon *DPEP1* overexpression in lymphocytic leukemia models have been observed,[55] and elevated levels of *DPEP1* messenger RNA have been detected in treatment-refractory samples in one AML study.[56] Little is known about the molecular mechanism by which *DPEP1* acts in different cancers, and its prognostic value is contradictory.[54,57] However, the overexpression of *DPEP1* found in R/PR AML merits investigation of its surface expression on AML cells for potential chimeric antigen receptor T cell targeting therapies, as previously indicated in pre–B-cell acute lymphoblastic leukemia.[58]

Machine learning models have gained popularity within biological big data analyses. Here, we relied on transparent classifiers, combined with preselection of features (ie, genes) to enhance predictive power as well as interpretability. Features were selected and ranked by using the MCFS algorithm (rmcfs[25]). Subsequently, R.ROSETTA was used to build predictive models to explore AML disease stages and to construct legible rule-based classifiers. The models allowed us to find patterns in the data based on rules that characterized copredictive mechanisms for genes with well-defined statistical properties. We found distinct *CD6* overexpression at AML relapse in adults, frequently combined with decreased *INSR* expression levels (Figure 4). This finding suggests that aberrant production of the lymphoid-associated surface glycoprotein CD6 may aid in therapy evasion, potentially mediated through interaction with the CD6-associated ligand ALCAM,[59] leading to adherence of the AML cells to a protected niche. Our results indicate that these AML cells simultaneously benefit from lower INSR levels, which hypothetically may lead to decreased cell proliferation and thus a more quiescent cell state associated with greater chemotherapy resistance. Furthermore, rule-based modeling identified downregulation of both *KATNAL2* and *NFATC4* at diagnosis in pediatric AML, with the expression levels at relapse being restored to what were seen for normal $CD34^+$ BM-controls (Figure 5; supplemental Figures 14 and 15). This suggests that low KATNAL2 and NFATC4 protein levels promote leukemia onset, whereas higher expression of the *KATNAL2* and *NFATC4* genes are selected for at relapse. *NFATC4* overexpression has been associated with cell quiescence and chemotherapy resistance in ovarian cancer.[43] Finally, the differential expression of *KATNAL2*, encoding a microtubule-severing enzyme,[44] raises the potential of introducing microtubule-targeting drugs such as colchicine or CYT997 (reviewed elsewhere[60]) as novel treatment alternatives at diagnosis to further distort microtubule function.

In conclusion, our results highlight the importance of complementary study approaches to fully elucidate the biological differences between leukemia blasts at diagnosis and their counterparts at a later stage during tumor progression. We identified novel or previously unappreciated DEGs associated with tumor progression in AML (eg, *IL1R1, CR1*, *GLI2*), many of which are expected to promote a pro-inflammatory tumor environment. Further studies are needed to investigate if targeting of their gene products or associated downstream pathways has therapeutic potential in AML or could improve treatment by sensitizing the cells to conventional drugs. We envision that knowledge gathered through this study gained by the combination of genomic and transcriptomic data, partially facilitated through machine learning approaches, will improve therapeutic innovations and help prolong patient survival.

Acknowledgments

Technical support was provided by Maria Lindström (Clinical Pathology, Uppsala University Hospital), Clinical Genomics Uppsala (SciLifeLab), and BioVis (Department of Immunology, Genetics and Pathology, Uppsala University). Patient samples were provided by U-CAN, Clinical Pathology and Clinical Genetics, Uppsala University Hospital (Sweden), and Nordic Society of Paediatric Haematology and Oncology. Transcriptomic sequencing was performed by the SNP&SEQ Technology Platform in Uppsala, part of the National Genomics Infrastructure Sweden and SciLifeLab. The SNP&SEQ Platform is supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. The computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) through the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX), partially funded by the Swedish Research Council through grant agreement no. 2018-05973, under Project SNIC sens2017604 and sens2018512. Computational assistance was provided by SciLifeLab-Wallenberg Advanced Bioinformatics Infrastructure (WABI) support at Uppsala University, and S.A.Y. was part of the Swedish Bioinformatics Advisory Program through the SciLifeLab bioinformatics platform. The authors thank Fredrik Barrenäs, who contributed to the improvement of rule clustering.

This work was supported by grants from the Knut and Alice Wallenberg Foundation (KAW 2013-0159), The Swedish Research Council (2013-03486), The Swedish Childhood Cancer Foundation (PR2013-0070 and TJ2013-0045), The Swedish Cancer Society (CAN2013/489), and The Kjell and Märta Beijer Foundation (L.H.), by grants from the Polish National Science Centre (DEC-2015/16/W/NZ2/00314), The University of Washington, Seattle, The National Institute of Allergy and Infectious Diseases, Division of AIDS, National Institutes of Health (ABL Contract No. HHSN272201700010I), and The eSSence program (J.K.).

Authorship

Contribution: S.S. performed experiments, analyzed data, and wrote the paper; S.A.Y. performed the machine learning–based analysis, merging of data sets for network comparisons, and wrote the corresponding part of the paper; M.G. performed network comparisons and visualization of machine learning–based results; J.S. performed RNA-seq gene fusion analysis; A.S., M.M., and N.N. analyzed RNA-seq data; M.K.H., C.S., A.E., M.H., J.P., J.A., K.J., M.C.M.-K., B.Z., K.P.T., and L.C. contributed clinical samples and/or data; J.K. contributed and supervised the machine learning–based analysis; L.H. designed the study, performed experiments, analyzed data, and wrote the paper; and all authors read and contributed to the final version of the manuscript.

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

ORCID profiles: S.S., 0000-0002-7438-9093; S.A.Y., 0000-0002-7201-2604; M.G., 0000-0002-2497-194X; N.N., 0000-0002-3823-1555; M.K.H., 0000-0001-7179-4643; C.S., 0000-0002-8160-5647; M.H., 0000-0003-2468-0226; J.K., 0000-0002-0766-8789; L.H., 0000-0003-4140-3423.

Correspondence: Linda Holmfeldt, Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Rudbeck Laboratory SE-75185 Uppsala, Sweden; e-mail: linda.holmfeldt@igp.uu.se.

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. Stone RM, Mandrekar SJ, Sanford BL, et al. Midostaurin plus chemotherapy for acute myeloid leukemia with a FLT3 mutation. *N Engl J Med.* 2017; 377(5):454-464.
3. Stein EM, DiNardo CD, Pollyea DA, et al. Enasidenib in mutant *IDH2* relapsed or refractory acute myeloid leukemia. *Blood.* 2017;130(6):722-731.
4. DiNardo CD, Stein EM, de Botton S, et al. Durable remissions with ivosidenib in IDH1-mutated relapsed or refractory AML. *N Engl J Med.* 2018; 378(25):2386-2398.
5. Verma D, Kantarjian H, Faderl S, et al. Late relapses in acute myeloid leukemia: analysis of characteristics and outcome. *Leuk Lymphoma.* 2010; 51(5):778-782.
6. Karlsson L, Forestier E, Hasle H, et al. Outcome after intensive reinduction therapy and allogeneic stem cell transplant in paediatric relapsed acute myeloid leukaemia. *Br J Haematol.* 2017;178(4):592-602.
7. Bejanyan N, Weisdorf DJ, Logan BR, et al. Survival of patients with acute myeloid leukemia relapsing after allogeneic hematopoietic cell transplantation: a Center for International Blood and Marrow Transplant Research study. *Biol Blood Marrow Transplant.* 2015;21(3):454-459.
8. Howlader NNA, Krapcho M, Miller D, et al. SEER Cancer Statistics Review (CSR) 1975-2016. SEER web site. Bethesda, MD: National Cancer Institute; 2018
9. Abrahamsson J, Forestier E, Heldrup J, et al. Response-guided induction therapy in pediatric acute myeloid leukemia with excellent remission rate. *J Clin Oncol.* 2011;29(3):310-315.
10. van Galen P, Hovestadt V, Wadsworth Ii MH, et al. Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell.* 2019;176(6):1265-1281.e24.
11. Wu J, Xiao Y, Sun J, et al. A single-cell survey of cellular hierarchy in acute myeloid leukemia. *J Hematol Oncol.* 2020;13(1):128.
12. Bolouri H, Farrar JE, Triche T Jr, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions [published corrections appear in *Nat Med.* 2018;24(4):526 and *Nat Med.* 2019;25(3):530]. *Nat Med.* 2018;24(1): 103-112.
13. Ley TJ, Miller C, Ding L, et al; 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.
14. Mou T, Pawitan Y, Stahl M, et al. The transcriptome-wide landscape of molecular subtype-specific mRNA expression profiles in acute myeloid leukemia. Am J Hematol. 2021;96(5):580-588.
15. Docking TR, Parker JDK, Jädersten M, et al. A clinical transcriptome approach to patient stratification and therapy selection in acute myeloid leukemia. *Nat Commun.* 2021;12(1):2474.
16. Wiggers CRM, Baak ML, Sonneveld E, Nieuwenhuis EES, Bartels M, Creyghton MP. AML subtype is a major determinant of the association between prognostic gene expression signatures and their clinical significance. *Cell Rep.* 2019;28(11):2866-2877.e5.
17. Hackl H, Steinleitner K, Lind K, et al. A gene expression profile associated with relapse of cytogenetically normal acute myeloid leukemia is enriched for leukemia stem cell genes. *Leuk Lymphoma.* 2015;56(4):1126-1128.
18. Toffalori C, Zito L, Gambacorta V, et al. Immune signature drives leukemia escape and relapse after hematopoietic cell transplantation. *Nat Med.* 2019;25(4):603-611.
19. Bachas C, Schuurhuis GJ, Zwaan CM, et al. Gene expression profiles associated with pediatric relapsed AML. *PLoS One.* 2015;10(4):e0121730.
20. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood.* 2016;127(20):2391-2405.

21. Stratmann S, Yones SA, Mayrhofer M, et al. Genomic characterization of relapsed acute myeloid leukemia reveals novel putative therapeutic targets. *Blood Adv.* 2021;5(3):900-912.
22. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11(3):R25.
23. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. *PLOS Comput Biol.* 2007;3(3):e39.
24. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics.* 2009;10(1):48.
25. Draminski M, Rada-Iglesias A, Enroth S, Wadelius C, Koronacki J, Komorowski J. Monte Carlo feature selection for supervised classification. *Bioinformatics.* 2008;24(1):110-117.
26. Garbulowski M, Diamanti K, Smolińska K, et al. R.ROSETTA: an interpretable machine learning framework. *BMC Bioinformatics.* 2021;22(1):110.
27. Smolinska K, Garbulowski M, Diamanti K, et al. VisuNet: an interactive tool for rule network visualization of rule-based learning models. GitHub repository: https://github.com/komorowskilab/VisuNet; 1 February 2019.
28. Zhao C, Chen A, Jamieson CH, et al. Hedgehog signalling is essential for maintenance of cancer stem cells in myeloid leukaemia. *Nature.* 2009; 458(7239):776-779.
29. Elsawa SF, Almada LL, Ziesmer SC, et al. GLI2 transcription factor mediates cytokine cross-talk in the tumor microenvironment. *J Biol Chem.* 2011;286(24):21524-21534.
30. Lim Y, Gondek L, Li L, et al. Integration of Hedgehog and mutant FLT3 signaling in myeloid leukemia [published correction appears in *Sci Transl Med.* 2015;7(295):295er6]. *Sci Transl Med.* 2015;7(291):291ra96.
31. Acuner Ozbabacan SE, Gursoy A, Nussinov R, Keskin O. The structural pathway of interleukin 1 (IL-1) initiated signaling reveals mechanisms of oncogenic mutations and SNPs in inflammation and cancer. *PLOS Comput Biol.* 2014;10(2):e1003470.
32. Jandrig B, Seitz S, Hinzmann B, et al. ST18 is a breast cancer tumor suppressor gene at human chromosome 8q11.2. *Oncogene.* 2004;23(57): 9295-9302.
33. Yang J, Siqueira MF, Behl Y, Alikhani M, Graves DT. The transcription factor ST18 regulates proapoptotic and proinflammatory gene expression in fibroblasts. *FASEB J.* 2008;22(11):3956-3967.
34. Hájková H, Fritz MH-Y, Haškovec C, et al. CBFB-MYH11 hypomethylation signature and PBX3 differential methylation revealed by targeted bisulfite sequencing in patients with acute myeloid leukemia. *J Hematol Oncol.* 2014;7(1):66.
35. Steinbach D, Bader P, Willasch A, et al. Prospective validation of a new method of monitoring minimal residual disease in childhood acute myelogenous leukemia. *Clin Cancer Res.* 2015;21(6):1353-1359.
36. Fearon DT. Identification of the membrane glycoprotein that is the C3b receptor of the human erythrocyte, polymorphonuclear leukocyte, B lymphocyte, and monocyte. *J Exp Med.* 1980;152(1):20-30.
37. Tedder TF, Fearon DT, Gartland GL, Cooper MD. Expression of C3b receptors on human be cells and myelomonocytic cells but not natural killer cells. *J Immunol.* 1983;130(4):1668-1673.
38. Iida K, Nussenzweig V. Functional properties of membrane-associated complement receptor CR1. *J Immunol.* 1983;130(4):1876-1880.
39. Lanza F, Castoldi G. Complement receptor 1 (CR1) expression in chronic myeloid leukemia. *Leuk Lymphoma.* 1992;8(1-2):35-41.
40. Toiyama Y, Inoue Y, Yasuda H, et al. DPEP1, expressed in the early stages of colon carcinogenesis, affects cancer cell invasiveness. *J Gastroenterol.* 2011;46(2):153-163.
41. Camano S, Lazaro A, Moreno-Gordaliza E, et al. Cilastatin attenuates cisplatin-induced proximal tubular cell damage. *J Pharmacol Exp Ther.* 2010; 334(2):419-429.
42. Castro MA, Oliveira MI, Nunes RJ, et al. Extracellular isoforms of CD6 generated by alternative splicing regulate targeting of CD6 to the immunological synapse. *J Immunol.* 2007;178(7):4351-4361.
43. Cole AJ, Iyengar M, Panesso-Gómez S, et al. NFATC4 promotes quiescence and chemotherapy resistance in ovarian cancer. *JCI Insight.* 2020: 5(7):e131486.
44. Ververis A, Christodoulou A, Christoforou M, Kamilari C, Lederer CW, Santama N. A novel family of katanin-like 2 protein isoforms (KATNAL2), interacting with nucleotide-binding proteins Nubp1 and Nubp2, are key regulators of different MT-based processes in mammalian cells. *Cell Mol Life Sci.* 2016;73(1):163-184.
45. Sadarangani A, Pineda G, Lennon KM, et al. GLI2 inhibition abrogates human leukemia stem cell dormancy. *J Transl Med.* 2015;13(1):98.
46. Terao T, Minami Y. Targeting hedgehog (Hh) pathway for the acute myeloid leukemia treatment. *Cells.* 2019;8(4):312.
47. Wellbrock J, Latuske E, Köhler J, et al. Expression of hedgehog pathway mediator GLI represents a negative prognostic marker in human acute myeloid leukemia and its inhibition exerts antileukemic effects. *Clin Cancer Res.* 2015;21(10):2388-2398.
48. Cortes JE, Heidel FH, Hellmann A, et al. Randomized comparison of low dose cytarabine with or without glasdegib in patients with newly diagnosed acute myeloid leukemia or high-risk myelodysplastic syndrome. *Leukemia.* 2019;33(2):379-389.
49. Beauchamp EM, Ringer L, Bulut G, et al. Arsenic trioxide inhibits human cancer cell growth and tumor development in mice by blocking Hedgehog/ GLI pathway. *J Clin Invest.* 2011;121(1):148-160.
50. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-674.

51. Lenkiewicz A, Bujko K, Brzezniakiewicz-Janus K, Xu B, Ratajczak MZ. The complement cascade as a mediator of human malignant hematopoietic cell trafficking. *Front Immunol.* 2019;10(1292):1292.

52. Carey A, Edwards DK V, Eide CA, et al. Identification of interleukin-1 by functional screening as a key mediator of cellular expansion and disease progression in acute myeloid leukemia. *Cell Rep.* 2017;18(13):3204-3218.

53. Wang Y, Sun X, Yuan S, et al. Interleukin-1β inhibits normal hematopoietic expansion and promotes acute myeloid leukemia progression via the bone marrow niche. *Cytotherapy.* 2020;22(3):127-134.

54. Eisenach PA, Soeth E, Röder C, et al. Dipeptidase 1 (DPEP1) is a marker for the transition from low-grade to high-grade intraepithelial neoplasia and an adverse prognostic factor in colorectal cancer. *Br J Cancer.* 2013;109(3):694-703.

55. Zhang J-M, Xu Y, Gale RP, et al. DPEP1 expression promotes proliferation and survival of leukaemia cells and correlates with relapse in adults with common B cell acute lymphoblastic leukaemia. *Br J Haematol.* 2020;190(1):67-78.

56. Horibata S, Gui G, Lack J, DeStefano CB, Gottesman MM, Hourigan CS. Heterogeneity in refractory acute myeloid leukemia. *Proc Natl Acad Sci U S A.* 2019;116(21):10494-10503.

57. Zhang G, Schetter A, He P, et al. DPEP1 inhibits tumor cell invasiveness, enhances chemosensitivity and predicts clinical outcome in pancreatic ductal adenocarcinoma. *PLoS One.* 2012;7(2):e31507.

58. Pölönen P, Mehtonen J, Lin J, et al. Hemap: an interactive online resource for characterizing molecular phenotypes across hematologic malignancies. *Cancer Res.* 2019;79(10):2466-2479.

59. Bowen MA, Patel DD, Li X, et al. Cloning, mapping, and characterization of activated leukocyte-cell adhesion molecule (ALCAM), a CD6 ligand. *J Exp Med.* 1995;181(6):2213-2220.

60. Chen X, Yang C, Xu Y, Zhou H, Liu H, Qian W. The microtubule depolymerizing agent CYT997 effectively kills acute myeloid leukemia cells via activation of caspases and inhibition of PI3K/Akt/mTOR pathway proteins. *Exp Ther Med.* 2013;6(2):299-304.