A META-ANALYSIS REVEALED DISTINCT MOLECULAR MARKERS IN PEDIATRIC AND ADULT ACUTE MYELOID LEUKEMIA

Iftikhar Aslam Tayubi
Faculty of Computing and Information Technology,
Rabigh, King Abdulaziz University, Jeddah - 21589, Kingdom of Saudi Arabia.
Email: iftikhartayubi@gmail.com

ABSTRACT

Acute Myeloid Leukemia (AML) is a blood-specific cancer characterized by diverse molecular alterations. We aimed to identify age-specific differently expressed genes through comparative analysis of pediatric and adult AML cohorts. Genes expression pattern were analysed across different age groups and potential hub genes were srudied for their molecular functions. The common differentially expressed genes in both adult and pediatric AML patients were filtered. These genes were utilized to develop protein-protein interaction network and top 16 genes were selected. Among the 16 differently expressed genes sharing common expression patterns in pediatric and adult AML, only 4 genes exhibited distinct expression patterns, suggesting potential differences in underlying mechanisms of AML between pediatric and adult patients. mRNA expression patterns of these genes were validated and age-dependent analysis was conducted. Notably, age-dependent analysis highlighted distinct expression patterns of CD44, FOXO3, PF4 and IL8 across different age groups. The survival analysis and functional enrichment analysis of hub genes further revealed the prognostic and biological relevance of identified genes in AML pathogenesis. This age-dependent meta-analyses contribute to a comprehensive understanding of gene expression patterns in distinct AML populations, paving the way for targeted therapeutic strategies and personalized treatment approaches in leukemia management.

Keywords: Leukemia, Acute Myeloid, AML, Adult-AML, Pediatric-AML, Age-Dependent Biomarkers

1. INTRODUCTION

Acute Myeloid Leukemia (AML) is a hematologic malignancy that is distinguished by the fast and uncontrolled growth of immature myeloid blast cells in the bone marrow (Behrmann et al., 2018). This disrupts normal blood cell production,

leading to life-threatening deficiencies related to red blood cells, white blood cells and platelets. These diseases predominantly affect adults age around 60-70 years and also occur in children and infants, often presenting with distinct clinical features and requiring different treatment approaches Pelcovits & Niroula, 2020; Boddu & Zeidan, 2019). In adults, a significant number of cases of AML are caused by a previous condition called myelodysplasia (MDS) or past exposure to radiotherapy or chemotherapy. Environmental factors, including radiation, tobacco smoke and benzene along with prior exposure to chemotherapeutic drugs, contribute to the multifaceted landscape of AML risk. However, in children, 95% of AML cases are considered to be new and not related to any previous condition (Roberts & Izraeli, 2014). Some pediatric AML patients exhibit a constitutional genetic susceptibility phenotypically with evident abnormalities like Down syndrome or a more modest familial syndrome (Niemeyer & Mecucci, 2017). About 70–80% of patients under 60 years old experience complete remission (CR) after few induction sessions (Sasaki et al., 2021). Nevertheless, in the context of pediatricAML, over 40% of patients experience a recurrence, leading to a comparatively modest overall survival rate of approximately 70% (Stubbins et al., 2022).

AML exhibits heterogeneity, with several subtypes classified based on specific genetic abnormalities and varying clinical behaviors. While some subtypes respond well to intensive chemotherapy, others are more resistant and carry dismal prognoses. The classification of AML involves various methods, considering factors such as etiology, genetics, immune phenotype and morphology (Taylor et al., 2017; Charrot et al., 2020). Myelodysplastic syndrome stands out as a prevalent risk factor, while myelofibrosis, aplastic anemia, down syndrome and bloom syndrome are associated with an elevated likelihood of developing AML (Burnett et al., 2011; Hartmann & Metzeler, 2019). The bone marrow microenvironment, comprising stromal cells, immune cells and cytokines, perform a vital role in the pathobiology of AML. Interactions between leukemic cells and the microenvironment influence disease progression, treatment response and the development of resistance (Marcucci et al., 2011). The primary method used to

establish treatment stratification in most Acute Myeloid Leukemia (AML) therapeutic protocols involves assessing the chromosomal and molecular anomalies found within the leukemic cells. In general, there are three distinct prognosis categories: favorable (good), intermediate and poor (adverse). Regarding clinical significance, it is not possible to apply adult classification systems, like the ones established by ELN (European LeukemiaNet), to classify childhood AML (Oran et al., 2015; Devillier et al., 2015). Because the genetic characteristics and risk associations of pediatric and adult AML differ based on age.

In pediatric acute myeloid leukemia, the prevailing cytogenetic abnormalities leading to chimeric fusion gene formation are balanced chromosomal rearrangements. However, it's worth noting that adult AML cases typically manifest only a subset of these fusion genes. Among these, core binding factor (CBF) leukemia, such as *t(8;21) (q22;q22)/RUNX1-RUNX1T1* and inv(16) (p13q22)/ CBFB-MYH11, demonstrate a favorable prognosis. Conversely, depending on the specific *KMT2A* partner gene, rearrangements involving *11q23/KMT2A (MLL)* exhibit an intermediate to poor prognosis. Rare instances occur where the rearrangement creates a positional effect, as seen in inv(3)(q21q26), wherein enhancer sequences are relocated closer to a protooncogene, leading to its increased expression (Gröschel et al., 2014; Calvo et al., 2021). In adults, recurrent mutations in genes like *FLT3, NPM1, CEBPA, RUNX1* and *IDH1/2* affect cell growth, survival and differentiation. Specific mutations like *NPM1*, often found in adults, can be associated with favorable outcomes. Similar to adults, children with imbalanced abnormalities like chromosome 5 and 7 monosomies have a poor prognosis (Blau, 2015).

There is still a significant research gap despite progress in understanding the molecular underpinnings of AML concerning the complex gene expression patterns specific to adult and pediatric populations. Existing studies often focus on either adult or pediatric AML, overlooking potential disparities or shared molecular features that may contribute to age-specific variations in disease presentation and response to treatment. This research gap underscores the need for a comprehensive

investigation that systematically explores and compares gene expression profiles in both adult and pediatric AML cohorts. Such an analysis is crucial for identifying age-specific molecular signatures and uncovering commonalities, ultimately informing more targeted and tailored therapeutic strategies for these distinct patient populations.In the molecular genetic landscape of pediatric and adult AML, a pivotal aspect lies in the application of differential gene expression analysis, which serves as a dynamic tool for explaining disease subtypes and identifying potential therapeutic targets. Through the systematic comparison of gene expression profiles across diverse AML patient cohorts, this analytical method contributes to a more accurate classification of AML by enabling the identification of signature genes and pathways linked to particular disease subtypes. The fundamental purpose of this work is to discover specific gene expression patterns in acute myeloid leukaemia (AML), focusing on the adult and pediatric populations. Identify genes that express differently and similarly across these age groups to understand the molecular characteristics of adult and pediatric AML.

2. MATERIALS AND METHODS

2.1 Data Mining

Gene expression profiled pediatric AML dataset (GSE2191) and adult AML dataset (GSE9476) was taken from GEO (Gene Expression Omnibus Clough & Barrett, 2016). GSE2191 contains 58 samples including 54 pediatric AML and 4 healthy bone marrow control samples obtained from the platform GPL8300 [HG_U95Av2] Affymetrix Human Genome U95 Version 2 Array and GSE9476 is comprised of 26 adult AML and 38 healthy control samples, retrieved from the platform GPL96 [HG-U133A] Affymetrix Human Genome U133A Array.

2.2. Microarray Data Pre-processing

The initial data from the dataset were exposed to background correction and normalization. Using an Entrez Gene ID converter, the first step in the data processing procedure included changing some gene symbols from probe IDs. When

multiple samples showed the same gene contribution, their mean value was calculated and taken into account as the final level of gene expression. The online statistical tool GEO2R was used to analyze the raw gene expression data; it included the R/Bioconductor and Limma package. The FDR and p-values were determined by applying the Benjamini and Hochberg (false discovery rate) and t-test procedures respectively with the GEO2R tool, in order to determine the DEGs. LogFC ≥ 1 and logFC ≤ −1 were regarded as indicators of upregulated and downregulated DEGs, respectively and P-value ($p < 0.05$) and log2fold change (| log (fold change) | > 1) to be statistically significant for the DEGs (Sultan et al., 2019). Additionally, the overlap and differences in DEGs across the two microarray profiles were visualized using FunRich V3.1.3 software (Benito-Martin & Peinado, 2015). Common molecular targets or pathways between the datasets are clearly depicted in FunRich's Venn diagram.

2.3. PPI Network Interaction of DEGs in AML

To visualize the PPIs between the statistically significant DEG-encoded proteins in both adult and pediatric AML, STRING (http://www.string-db.org/) was utilized (Szklarczyk et al., 2010). Given the dependability of these protein connections, an interaction threshold of a combined score >0.08 was considered significant. The PPI network obtained from the STRING database can be seen using the Cytoscape software v3.7.1 (https://cytoscape.org/). The intensity of the interactions between the proteins was depicted by the PPI network by assigning widths to the edges that connect them, using the combined score. Cytohubba identified cohesive gene clusters modules within the PPI network (Sultan et al., 2022).

2.4. Gene Ontology and Pathway Analysis

The Database for Annotation, Visualisation and Integrated Discovery (DAVID) (http://david.ncifcrf.gov) was employed to comprehend the function of differentially expressed genes in enriched pathways. Genes are grouped into molecular functions, cellular components and biological processes (BP) using systematic and standard annotations provided by this tool (Madar et al., 2021).

Furthermore, to identify pathways that showed significantly higher enrichment with the identified DEGs, KEGG pathway study was performed.

2.5. mRNA expression and Survival Analysis

UALCAN (http://ualcan.path.uab.edu/) (Chandrashekar et al., 2017), GEPIA (http://gepia.cancer-pku.cn/) (Tang et al., 2017) and KM plotter (Jia et al., 2019), were collectively used to analyze the survival rate and interactions between significantly expressed genes. KaplanMeier analysis was used in conjunction with log-rank testing for survival analysis. Using a significance threshold of $P < 0.05$, it was concluded that there was a statistically significant relationship between the levels of gene expression and patient survival. Expression validation was carried out by using AML patient data from TCGA) (Tomczak et al., 2015). Two groups were created from the quantifiable data, which were expressed as transcripts per million (TPM) values and the GEPIA database was used to visualize this stratification. The low-expressed group included patients whose TPM readings fell below the top quartile, whereas the highexpression group included patients whose TPM values exceeded the upper quartile.

2.6. Prediction of Drug-Gene Interaction

DGIdb2.0 (http://www.dgidb.org/) was utilized to forecast drug-gene interactions, leveraging differentially expressed genes (DEGs) identified within the module (Cotto et al., 2018). The Drug-Gene Interaction database (DGIdb) serves as a valuable resource for investigating drug-gene interactions and assessing gene druggability across various human diseases. Notably, the parameter for this analysis was configured to consider only interactions involving drugs that are FDA approved.

3. RESULTS AND DISCUSSION

3.1. Identification of Differentially Expressed Genes in AML

Gene expression profiling of GSE2191 and GSE9476 revealed 576 and 277 DEGs, respectively, utilizing the limma package, with a threshold of p-value < 0.05 and

log fold change > 1 (Figure 1). To validate the initial findings, FunRich_V3 was further used to identify overlapping DEGs between pediatric and adult AML datasets. Figure 2 (A) illustrates the number of common DEGs identified from each dataset through an intersecting venn diagram. Notably, we found 16 genes that were consistently regulated in both the GSE2191 and GSE9476 datasets.

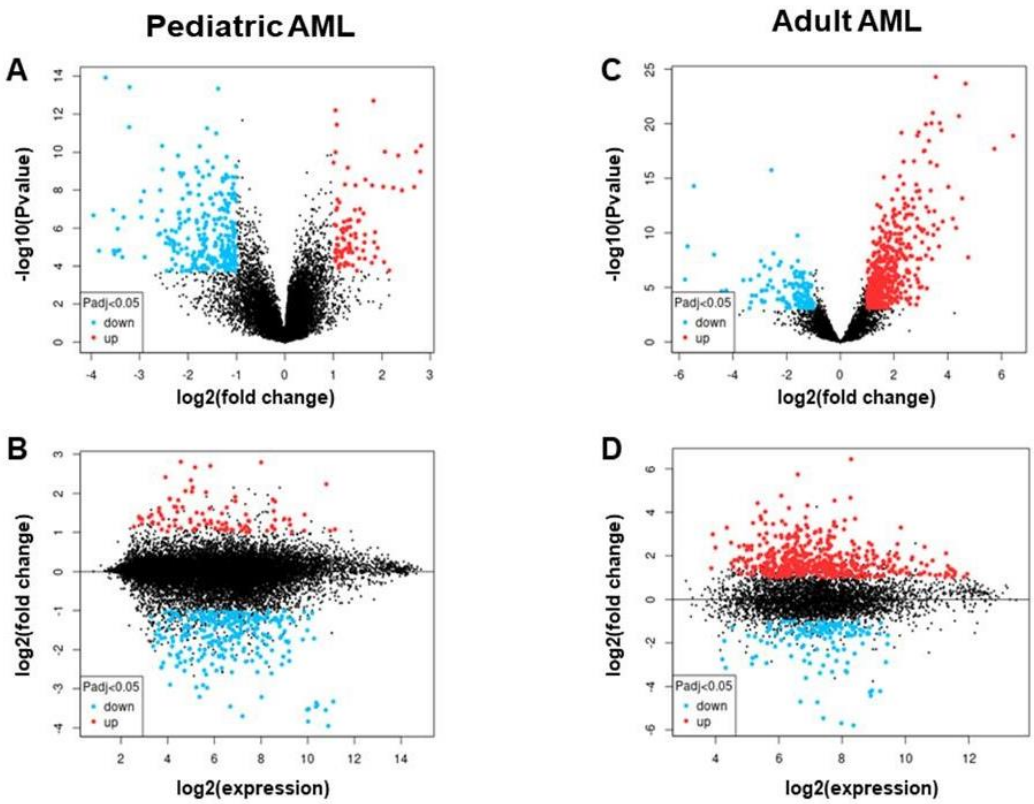


Figure 1: Differentially expressed genes and mean-difference plot in 2 different age groups (A, B: Pediatric AML), (C, D: Adult AML). Each red color circle represents up-regulated genes, blue color for down-regulated genes and black color indicates non-significant genes.

3.2. Common Genes Expression Analysis in AML Patients

Gene expression analysis in two categories of AML patients, i.e., pediatric AML and adult AML patients, revealed 16 genes with overlapping expression patterns. These genes were either upregulated (AHR, CD44, KIF2A, NEK7, SEC62, SLC25A36) or downregulated (ANK1, CBX7, FOXO3, FPR2, GCA, KCNH2, LUZP1, PF4, PRDX2, SLA) in both patient groups. Notably, 4 genes (KIF2A,

LUZP1, NEK7, SEC62) exhibited distinct expression patterns, suggesting potential differences in underlying mechanisms of AML between pediatric and adult patients (Figure 2B). These genes need further investigation as potential targets for developing age-specific therapeutic strategies.

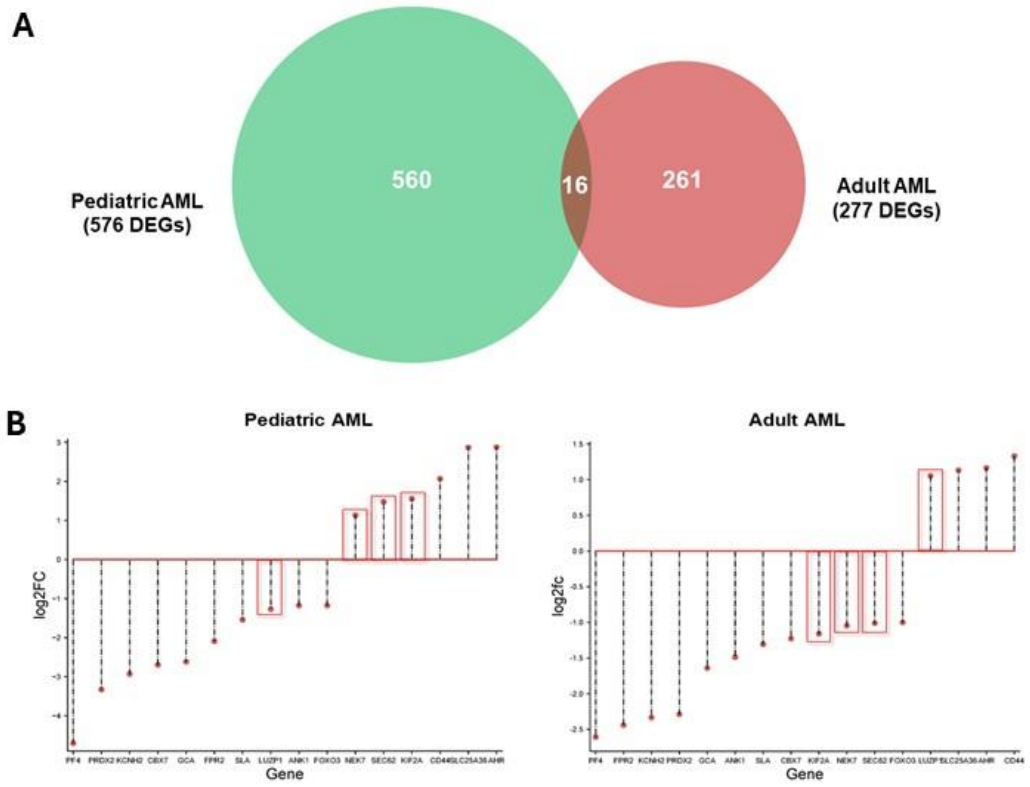
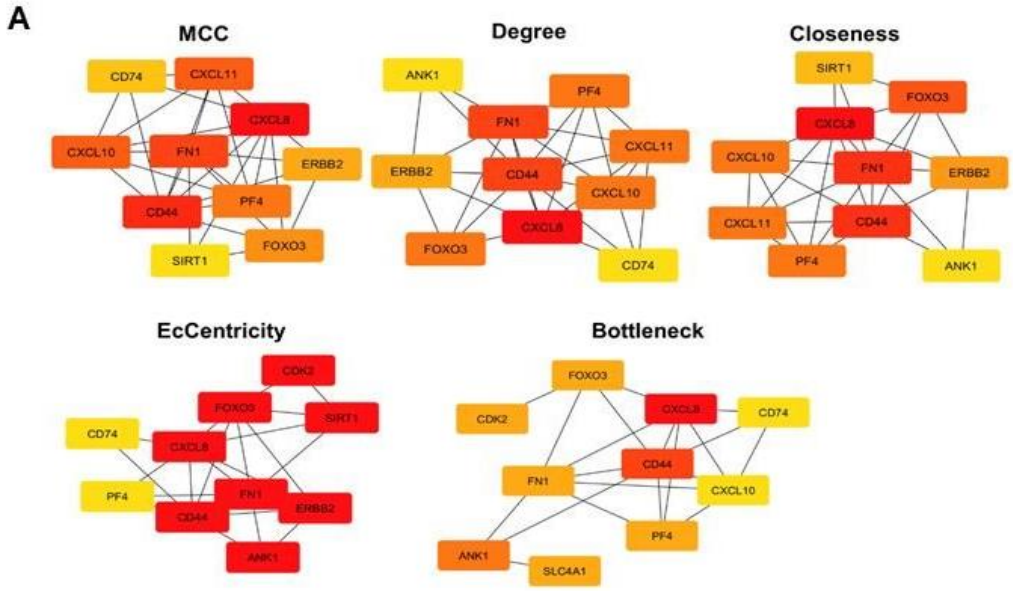


Figure 2: (A) Venn diagram denotes the overlapping DEGs from Pediatric AML and Adult AML datasets. (B) Expression pattern of common genes in both age groups.

3.3. Hub Gene Screening and PPI Network Construction

Network analysis was performed to investigate the connections among proteins encoded by 16 overlapping Differentially Expressed Genes (DEGs). To resulting PPI network comprised 26 nodes and 29 edges. To identify key players within this network, five hub genes; CD44, FN1, PF4, CXCL8, and FOXO3 were identified through various network analysis methods, including MCC, Bottleneck, EcCentricity, Closeness, and Degree (Figure 3). These hub genes demonstrated high centrality and connectivity, implying their potential as significant regulators in the

biological system under investigation. The topological properties of hub genes are outlined in Table 1. It provides a deeper understanding of the underlying biological processes at play, enriching our comprehension of the intricate mechanisms governing cellular function and disease pathology.



C

MCC	Degree	BottleNeck	EcCentricity	Closeness
CXCL8	**CXCL8**	**CXCL8**	ANK1	**CXCL8**
CD44	**FN1**	**CD44**	CDK2	**FN1**
FN1	**CD44**	ANK1	SIRT1	**CD44**
CXCL11	**PF4**	**PF4**	**FN1**	**FOXO3**
CXCL10	CXCL10	SLC4A1	**CXCL8**	**PF4**
PF4	CXCL11	CDK2	ERBB2	CXCL10
FOXO3	**FOXO3**	**FN1**	**CD44**	CXCL11
CD74	ERBB2	**FOXO3**	**FOXO3**	ERBB2
ERBB2	CD74	CD74	CD74	SIRT1
SIRT1	ANK1	CXCL10	**PF4**	ANK1

Figure 3: (A) Top 10 hub genes based on MCC, degree, Bottleneck, EcCentricity, and Closeness. (B) Venn diagram denotes the common genes in all the five methods. (C) Top 10 hub genes ranked by cytohubba.

Table 1: The topology properties of hub genes.

Gene	MCC	Degree	EcCentri-city	Closeness	Radiality	Between ness	Clustering Coefficient
CXCL8	176	10	0.294	11.667	3.466	38.181	0.467
CD44	174	9	0.294	11.333	3.466	26.081	0.556
FN1	156	9	0.294	11.333	3.466	27.205	0.528
FOXO3	32	6	0.294	9.833	3.277	22.705	0.6
CXCL11	144	6	0.221	9.417	3.088	1	0.867
CXCL10	144	6	0.221	9.417	3.088	1	0.867
PF4	122	6	0.221	9.417	3.088	5.600	0.733
ERBB2	30	5	0.294	9.333	3.214	3.700	0.8
SIRT1	8	4	0.294	8.833	3.151	11.348	0.667
CD74	24	4	0.221	8.417	2.962	0	1
ANK1	7	4	0.294	8.667	3.088	25.381	0.5
CDK2	3	3	0.294	7.500	2.773	18.619	0.333
FPR2	2	2	0.221	7.167	2.710	0	1
PRDX2	2	2	0.221	6.083	2.206	3.9	0
SLC4A1	2	2	0.221	6.583	2.521	7.281	0
AIP	1	1	0.118	1	0.353	0	0
AHR	1	1	0.118	1	0.353	0	0

3.4. mRNA expression patterns and survival analysis

The validation of hub gene expression patterns was performed using the GEPIA online software, confirming significant alterations in gene expression between Acute Myeloid Leukemia (LAML) tumor samples (173) and normal samples (70). Notably, CD44, FOXO3 and PF4 exhibited upregulation, among them CD44 showed significant upregulation suggesting an increased expression in the tumor context. This upregulation implies their potential involvement in the molecular mechanisms underlying leukemogenesis. In contrast, FN1 displayed significant downregulation, indicating a decreased expression level in LAML tumor samples compared to normal counterparts (Figure 4). This differential expression of hub genes underscores their potential role as molecular markers and implicates their involvement in the pathogenesis of AML. Figure 4 revealed that upregulated CD44 and downregulated FN1 were associated with poor overall survival for AML patients.

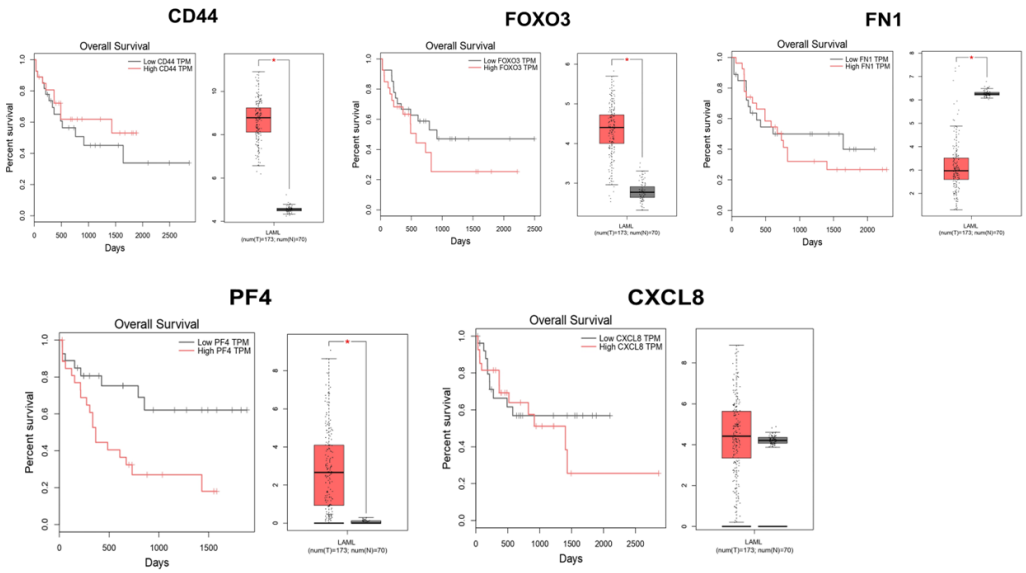


Figure 4: Overall survival curves of prognostic hub genes (CD44, FOXO3, PF4, CXCL8 and FN1) and corresponding mRNA expression levels in patients with acute myeloid leukemia (AML) from The Cancer Genome Atlas (TCGA).

3.5. Age-Dependent Biomarkers

In the investigation of hub gene expression across distinct age groups in AML, the UALCAN web tool was employed, utilizing the TCGA database. Transcript per million (TPM) served as the metric on the y-axis, while age groups spanning 21100 years were depicted on the x-axis. Remarkably, the CD44 gene and FOXO3 exhibited heightened expression in the 21-40 years age group, indicating potential roles in early adulthood leukemogenesis. IL8 emerged as a significant contributor, displaying increased expression in the 41-60 years age group, implicating its relevance in middle-aged AML patients. Notably, CD44 retained significance in the 61-80 years age group, suggesting a prolonged impact across different life stages. Intriguingly, in the 81-100 years age group, both FOXO3 and PF4 genes demonstrated substantial enrichment, possibly signifying a distinct molecular landscape in elderly AML patients. The age-stratified analysis enhances our understanding of the significant roles of CD44, FOXO3, PF4, and IL8 in AML progression (Figure 5), providing valuable insights for age-tailored therapeutic considerations in leukaemia management.

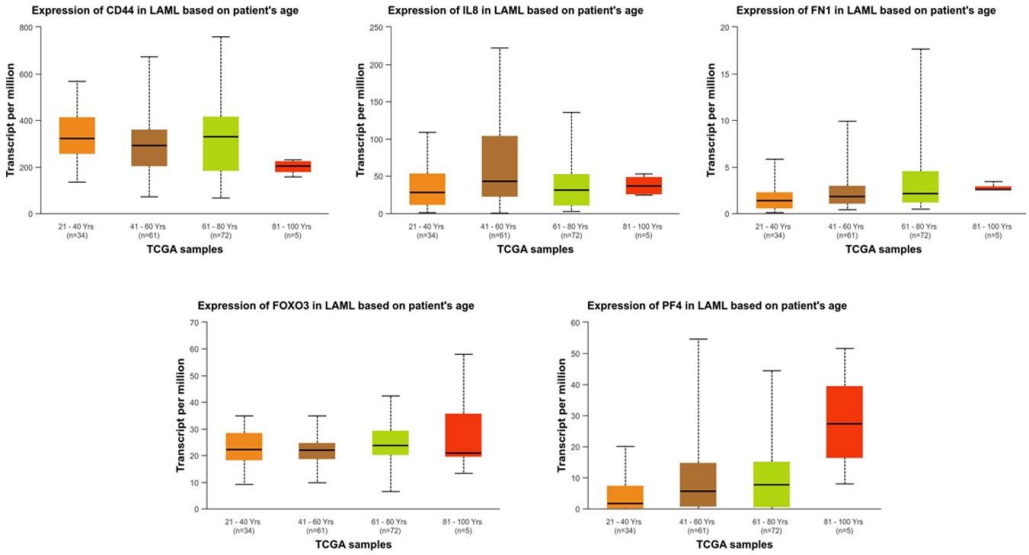


Figure 5: Co-expressed genes identified in AML patients between different age groups.

3.6. Functional Enrichment Analysis

The GO function annotations of overlapping DEGs revealed that the dysregulation of the identified hub genes is notably associated with crucial biological processes (Figure 6A) including RNA splicing, a fundamental step in gene expression, undergoes aberrations, impacting the generation of functional mRNA. The regulation of mRNA metabolic process experiences disruption, affecting the stability and turnover of mRNA molecules. Negative regulation of the cell cycle is compromised, leading to abnormal cellular growth, a defining trait of AML. Additionally, the dysregulated genes contribute to abnormalities in the regulation of chromosome organization, impacting genomic stability. Perturbations in the regulation of stem cell differentiation further contribute to the leukemic phenotype. Nuclear transport mechanisms are also affected, influencing the subcellular localization of key regulatory proteins. The differential expression of genes in Acute Myeloid Leukemia (AML) is intricately linked to alterations in various cellular components (Figure 6C). Nuclear specks, essential for pre-mRNA processing, undergo dysregulation, impacting gene expression. Perturbations at the nuclear periphery, including the nuclear membrane and envelope, suggest disruptions in nuclear organization. PML bodies, are known for their roles in genomic stability,

contributing to the genomic instability observed in AML. The nuclear matrix, is involved in chromatin organization, influencing overall genome structure. Inclusion bodies cause aberrant protein aggregation, while spliceosome bodies hint at disturbances in RNA splicing. Chromosomal regions change, reflecting genomic instability. Nuclear membrane and local adhesion alterations impact cellular structure and interactions. Whereas in molecular function, transcription co-regulator activity points to dysregulated gene expression control, emphasizing the intricate transcriptional landscape in AML (Figure 6B).

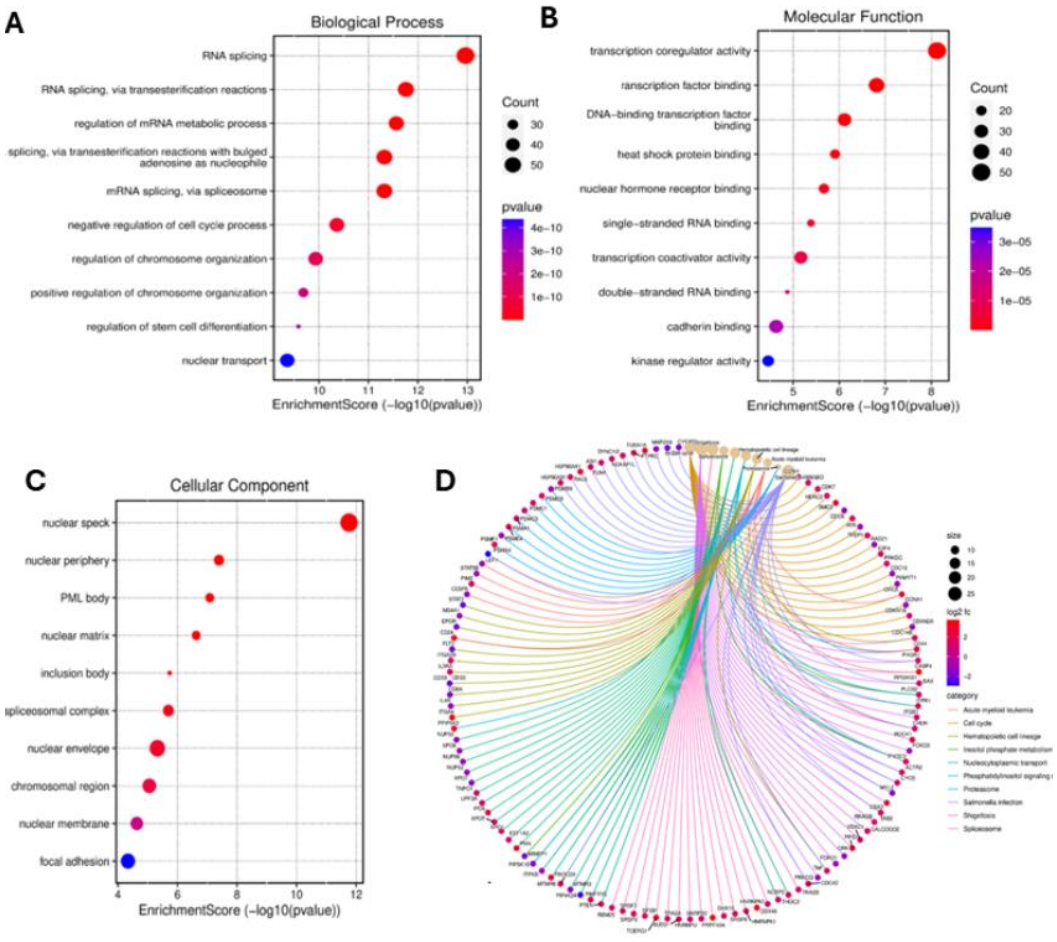


Figure 6. The enriched ontology terms and enrichment score of identified DEGs for Biological process (A), Molecular functions (B) and Cellular localization (C). The top 10 enriched KEGG pathways and respective scores (D).

DNA-binding transcription factor binding highlights alterations in transcriptional regulation, potentially contributing to abnormal cell proliferation. Heat shock protein binding suggests a response to cellular stress, a common feature in cancer. Nuclear hormone receptor binding signifies disruptions in hormone signaling pathways. Single-stranded RNA binding may indicate perturbations in RNA metabolism, influencing gene expression. Transcription coactivator activity points to enhanced transcriptional activation, contributing to leukemic phenotypes. Double-stranded RNA binding and cadherin binding imply diverse roles in RNA processing and cell adhesion, respectively. Kinase regulator activity suggests aberrations in signaling pathways, contributing to dysregulated cellular processes in AML (Figure 6 D). The aforementioned discoveries lay the groundwork for future investigations in both pediatric and adult AML, providing valuable insights that could guide further exploration and advancements in our understanding and management of this disease.

3.7. Identification of drug-gene interaction

DGidb results show significantly expressed gene with 24 drug-gene interactions, containing 8 FDA-approved drugs, as listed in Gene-drug interaction table, dacarbazine and docetaxel anhydrous are antineoplastic drug interact with the hub genes FN1 and CD44 with scores of 0.382987 and 0.053376 respectively. The drug-targeting hub genes are summarized in Table 2.

Table 2: Potential drug targets of the identified hub genes based on DGIdb database.

Gene	Drug	Regulatory approval	Indication	Interaction score
FN1	Dacarbazine	Approved	Anti-neoplastic agent	0.3830
	Ocriplasmin	Approved	-	0.2822
	L19tnfa	Not Approved	-	5.3618
	Daromun	Not Approved	-	5.3618
	As-1409	Not Approved	-	5.3618
	L19sip 131i	Not Approved	-	5.3618
	Mu-bc-1	Not Approved	-	5.3618
	Radretumab	Not Approved	-	5.3618
	Onfekafusp alfa	Not Approved	-	5.3618
	L19il2	Not Approved	-	10.7237
	As1409	Not Approved	Anti-neoplastic agent	1.7873
CD44	Docetaxel anhydrous	Approved	Anti-neoplastic agent	0.0534
	Mometasone	Approved	Anti-inflammatory	1.1342
	Gentamicin	Approved	-	0.3945
	Acetaminophen	Approved	Analgesic	0.1008
	Hyaluronic acid	Approved	Antileukopenic agent	1.5123
	Cisplatin	Approved	-	0.0304
	Recombinant tumor necrosis factor family protein	Not Approved	-	0.6049
	Bivatuzumab mertansine	Not Approved	-	0.2836
	Bivatuzumab	Not Approved	-	4.5369
	Gamma-interferon	Not Approved	-	1.0082
	Mpa	Not Approved	-	0.9074
	Pf-03475952	Not Approved	-	4.5369
	Gp-120 antigen	Not Approved	-	0.6980

In pediatric patients, the intricacies of the disease are heightened, necessitating a deeper understanding to tailor more effective diagnostic and therapeutic strategies (Kogan et al., 2019; Selim et al., 2019). Conventional treatments, such as chemotherapy and stem cell transplantation, are being used for managing AML. Nevertheless, the absence of accurate prognostic markers impedes the capacity to customize treatments for specific patients, resulting in unsatisfactory outcomes in

numerous instances. Recent years contributed to an increase in bioinformatics analysis for finding new therapy targets and cancer markers (Jiang & Liu, 2015; Tan & Wang, 2019). Yin, et. al. demonstrated the critical function of the G2/M checkpoint in the early stages of hepatocellular carcinoma, suggesting the potential prognostic and diagnostic significance of the checkpoints involving genes (Yin et al., 2017). In another study Tang, et. al. identified 100 hub genes hub genes in AML, revealing their enrichment in pathways related to immune regulation, apoptosis, and myeloid cell differentiation (Tan et al., 2020). However, when compared to this study, they performed gene expression profiling and utilized the module method to identify genes exhibiting high connectivity. Furthermore, the selected genes were solely validated using the Kaplan–Meier plotter database. Our study centered around a multi-step process that utilized several bioinformatics tools to conduct a thorough investigation of Acute Myeloid leukemia (AML). The process of data mining began by accessing the Gene Expression Omnibus (GEO) database. Specifically, datasets GSE9476 and GSE2191 were selected based on strict criteria to ensure the inclusion of samples from Acute Myeloid leukemia (AML) and corresponding healthy controls, along with sufficient clinical data. The mRNA expression pattern reveals that CD44 and FN1 show significant upregulation and downregulation respectively. CD44, a transmembrane glycoprotein known for its involvement in cell adhesion and migration, assumes significance in AML by promoting leukemic cell adhesion to the bone marrow microenvironment, thereby contributing to disease progression and chemoresistance (Chen et al., 2018). Upregulated CD44 expression has been linked to increased resistance to apoptosis (programmed cell death) in cancer cells. This anti-apoptotic effect allows leukemia cells to survive and proliferate, contributing to the uncontrolled growth characteristic of cancer (Ruffell & Johnson, 2008). Recent studies have reported a potential link between the CD44 protein and the characteristics of cancer stem cells. In acute myeloid leukemia, increased levels of CD44 expression may play a role in sustaining the leukemic stem cell population. Cancer stem cells represent a subgroup of cells within a tumor that possess the ability to self-renew and

differentiate into diverse cell types, thereby driving tumor growth and relapse (Wang et al., 2018). CD44 can activate various signaling pathways, including those involved in cell proliferation and survival. Upregulated CD44 may contribute to the dysregulation of signaling cascades, promoting the uncontrolled growth of AML cells. FN1, an extracellular matrix glycoprotein essential for cell adhesion and signaling, is involved in the leukemic stem cell microenvironment of AML. FN1 is also responsible for maintaining and expanding leukemic cells, thereby affecting the microenvironment to promote disease development (Liu et al., 2019).

Fibronectin 1 (FN1) is a crucial constituent of the extracellular matrix (ECM) and serves as a structural framework for cellular adhesion. Reducing the expression of Alterations in the FN1 gene may lead to a reduced capacity of leukemia cells to adhere to the extracellular matrix (ECM), hence disturbing normal cellular connections. Cell adhesion disruption can facilitate the liberation of leukemia cells from the bone marrow microenvironment, enabling their circulation in the bloodstream and potential infiltration of other organs (Spada et al., 2021). The study of gene-drug interaction reveals that CD44 has interaction with docetaxel anhydrous. Docetaxel is a member of taxanes, a group of anticancer drugs that work by causing tubulin to form polymers and halting the cell cycle. Furthermore, recent findings have demonstrated that taxanes, such as docetaxel, exert an influence on both cancer cells and neighbouring cells using many pathways, including the modulation of immune system responses, angiogenesis, and metastasis (Guppta et al., 2023). Dacarbazine, another antineoplastic medication, exhibited an interaction with FN1. Dacarbazine is a triazene compound that primarily causes methylation of O(6)-guanine by the action of methyl-diazonium ion, that exerts its cytotoxic (cell-killing) and mutagenic (DNA-damaging) effects through the formation of DNA O(6)-methylguanine adducts. These adducts represent DNA lesions where a methyl group is attached to the oxygen atom at the 6th position of the guanine nucleobase. The presence of these adducts in the DNA can interfere with proper DNA replication and transcription, leading to cell death or mutations (Hwa et al., 2020).

Further exploration is necessary to understand the pharmacological processes between the hub genes and drugs.

4. CONCLUSION

Acute myeloid leukemia is a complex blood-based cancer. The intricacies associated with both pediatric and adult AML necessitates a deeper understanding to tailor more effective diagnostic and therapeutic strategies. In this study, we aimed to perform a comprehensive in silico analysis in different age groups of AML patients, i.e., infants and adults. We identified sixteen overlapping DEGs that were consistently expressed in both groups, out of which five hub genes (CD44, FN1, CXCL8, FOXO3 and PF4) were found biologically linked with the onset and progression of AML. The gene correlation analysis revealed the potential association between the identified key regulatory hub genes and their clinical features, specifically age. Amongst the top 5 hub genes, FN1 and CD44 exhibit significantly varying expression at mRNA-level. Additionally, for CD44 and FN1, a total of 24 potentially druggable molecules were identified. Two of these drug molecules, i.e., dacarbazine and docetaxel anhydrous, were antineoplastic agents that may serve as potential drugs to target CD44 and FN1-derived AML. Alongside identifying the common biomarkers independent of patient's age, this study also offers additional insights into the therapeutic targets for the reported biomarkers that may serve equally well in both pediatric and adult AML patients.

REFERENCES

Behrmann, L., Wellbrock, J., & Fiedler, W. (2018). Acute myeloid leukemia and the bone marrow niche—take a closer look. *Frontiers in oncology*, *8*, 444.

Pelcovits, A., & Niroula, R. (2020). Acute myeloid leukemia: a review. *Rhode Island medical journal*, *103*(3), 38-40.

Boddu, P. C., & Zeidan, A. M. (2019). Myeloid disorders after autoimmune disease. *Best Practice & Research Clinical Haematology*, *32*(1), 74-88.

Roberts, I., & Izraeli, S. (2014). Haematopoietic development and leukaemia in D own syndrome. *British journal of haematology*, *167*(5), 587-599.

Niemeyer, C. M., & Mecucci, C. (2017, April). Practical considerations for diagnosis and management of patients and carriers. In *Seminars in Hematology* (Vol. 54, No. 2, pp. 69-74). WB Saunders.

Sasaki, K., Ravandi, F., Kadia, T. M., DiNardo, C. D., Short, N. J., Borthakur, G. & Kantarjian, H. M. (2021). De novo acute myeloid leukemia: a population-based study of outcome in the United States based on the Surveillance, Epidemiology, and End Results (SEER) database, 1980 to 2017. *Cancer*, *127*(12), 2049-2061.

Stubbins, R. J., Francis, A., Kuchenbauer, F., & Sanford, D. (2022). Management of acute myeloid leukemia: A review for general practitioners in oncology. *Current Oncology*, *29*(9), 6245-6259.

Taylor, J., Xiao, W., & Abdel-Wahab, O. (2017). Diagnosis and classification of hematologic malignancies on the basis of genetics. *Blood, The Journal of the American Society of Hematology*, *130*(4), 410-423.

Charrot, S., Armes, H., Rio-Machin, A., & Fitzgibbon, J. (2020). AML through the prism of molecular genetics. *British journal of haematology*, *188*(1), 49-62.

Burnett, A., Wetzler, M., & Lowenberg, B. (2011). Therapeutic advances in acute myeloid leukemia. *J Clin Oncol*, *29*(5), 487-494.

Hartmann, L., & Metzeler, K. H. (2019). Clonal hematopoiesis and preleukemia—Genetics, biology, and clinical implications. *Genes, Chromosomes and Cancer*, *58*(12), 828-838.

Marcucci, G., Haferlach, T., & Döhner, H. (2011). Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. *Journal of Clinical Oncology*, *29*(5), 475-486.

Oran, B., Jimenez, A. M., De Lima, M., Popat, U. R., Bassett, R., Andersson, B. S. & Champlin, R. E. (2015). Age and modified European LeukemiaNet classification to predict transplant outcomes: an integrated approach for

acute myelogenous leukemia patients undergoing allogeneic stem cell transplantation. *Biology of Blood and Marrow Transplantation*, *21*(8), 1405-1412.

Devillier, R., Gelsi-Boyer, V., Murati, A., Prebet, T., Rey, J., Etienne, A. & Vey, N. (2015). Prognostic significance of myelodysplasia-related changes according to the WHO classification among ELN-intermediate-risk AML patients. *Am J Hematol*, *90*(1), E22-E24.

Gröschel, S., Sanders, M. A., Hoogenboezem, R., de Wit, E., Bouwman, B. A., Erpelinck, C & Delwel, R. (2014). A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell*, *157*(2), 369-381.

Calvo, C., Fenneteau, O., Leverger, G., Petit, A., Baruchel, A., & Méchinaud, F. (2021). Infant acute myeloid leukemia: A unique clinical and biological entity. *Cancers*, *13*(4), 777.

Blau, O. (2015). Gene mutations in acute myeloid leukemia-incidence, prognostic influence, and association with other molecular markers. *LeukemiasUpdates and New Insights. Croatia: Intech Open*, 75-100.

Clough, E., & Barrett, T. (2016). The gene expression omnibus database. Statistical Genomics: Methods and Protocols, 93-110.

Sultan, G., Zubair, S., Tayubi, I. A., Dahms, H. U., & Madar, I. H. (2019). Towards the early detection of ductal carcinoma (a common type of breast cancer) using biomarkers linked to the PPAR (γ) signaling pathway. *Bioinformation*, *15*(11), 799.

Benito-Martin, A., & Peinado, H. (2015). FunRich proteomics software analysis, let the fun begin!. Proteomics, 15(15), 2555-2556.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P. & Mering, C. V. (2010). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic acids research, 39(suppl_1), D561-D568.

Sultan, G., Zubair, S., Madar, I. H., & Anandaram, H. (2022). Computational Approach to Identify Regulatory Biomarkers in the Pathogenesis of Breast Carcinoma. *International Journal of Advanced Computer Science and Applications*, *13*(6).

Madar, I. H., Sultan, G., Tayubi, I. A., Hasan, A. N., Pahi, B., Rai, A. & Rai, S. (2021). Identification of marker genes in Alzheimer's disease using a machine-learning model. Bioinformation, 17(2), 348.

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B. V., & Varambally, S. (2017). UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia, 19(8), 649-658.

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., & Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic acids research, 45(W1), W98-W102.

Jia, B., Zhao, X., Wang, Y., Wang, J., Wang, Y., & Yang, Y. (2019). Prognostic roles of MAGE family members in breast cancer based on KM-Plotter Data. Oncology letters , 18 (4), 3501-3516.

Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemporary Oncology/Współczesna Onkologia, 2015(1), 68-77.

Cotto, K. C., Wagner, A. H., Feng, Y. Y., Kiwala, S., Coffman, A. C., Spies, G. & Griffith, M. (2018). DGIdb 3.0: a redesign and expansion of the drug–gene interaction database. *Nucleic acids research*, *46*(D1), D1068-D1073.

Kogan, A. A., Lapidus, R. G., Baer, M. R., & Rassool, F. V. (2019). Exploiting epigenetically mediated changes: acute myeloid leukemia, leukemia stem cells and the bone marrow microenvironment. *Advances in Cancer Research*, *141*, 213-253.

Selim, A., Alvaro, F., Cole, C. H., Fraser, C. J., Mechinaud, F., O'Brien, T. A. & Moore, A. S. (2019). Hematopoietic stem cell transplantation for children with acute myeloid leukemia in second remission: a report from the

Australasian Bone Marrow Transplant Recipient Registry and the Australian and New Zealand Children's Haematology Oncology Group. *Pediatric Blood & Cancer*, *66*(8), e27812.

Jiang, P., & Liu, X. S. (2015). Big data mining yields novel insights on cancer. *Nature genetics*, *47*(2), 103-104.

Tan, J., Jin, X., & Wang, K. (2019). Integrated bioinformatics analysis of potential biomarkers for prostate cancer. *Pathology & Oncology Research*, *25*, 455460.

Yin, L., Chang, C., & Xu, C. (2017). G2/M checkpoint plays a vital role at the early stage of HCC by analysis of key pathways and genes. *Oncotarget*, *8*(44), 76305.

Tan, Y., Zheng, L., Du, Y., Zhong, Q., Zhu, Y., Liu, Z. & Zhang, Q. (2020). Identification of the hub genes and pathways involved in acute myeloid leukemia using bioinformatics analysis. *Medicine*, *99*(35).

Chen, C., Zhao, S., Karnad, A., & Freeman, J. W. (2018). The biology and role of CD44 in cancer progression: therapeutic implications. *Journal of hematology & oncology*, *11*, 1-23.

Ruffell, B., & Johnson, P. (2008). Hyaluronan induces cell death in activated T cells through CD44. *The Journal of Immunology*, *181*(10), 7044-7054.

Wang, L., Zuo, X., Xie, K., & Wei, D. (2018). The role of CD44 and cancer stem cells. *Cancer Stem Cells: Methods and Protocols*, 31-42.

Liu, T., Zhou, L., Li, D., Andl, T., & Zhang, Y. (2019). Cancer-associated fibroblasts build and secure the tumor microenvironment. *Frontiers in cell and developmental biology*, *7*, 60.

Spada, S., Tocci, A., Di Modugno, F., & Nisticò, P. (2021). Fibronectin as a multiregulatory molecule crucial in tumor matrisome: from structural and functional features to clinical practice in oncology. *Journal of Experimental & Clinical Cancer Research*, *40*(1), 1-14.

Gupta, R., Kadhim, M. M., Jalil, A. T., Alasheqi, M. Q., Alsaikhan, F., Mukhamedova, N. K. & Najafi, M. (2023). The interactions of docetaxel

with tumor microenvironment. *International Immunopharmacology*, *119*, 110214.

Hwa Yun, B., Guo, J., Bellamri, M., & Turesky, R. J. (2020). DNA adducts: Formation, biological effects, and new biospecimens for mass spectrometric measurements in humans. Mass spectrometry reviews, 39(1-2), 55–82.