British Journal of Cancer Research
2020; 3(3): 415 - 422 . doi: 10.31488/bjcr.154

Research article

Altered Transcriptome in Pediatric AML Compared with Normal Hematopoiesis

Yu Fan[1*], Ying Hu[1], Chunhua Yan[1], Qingrong Chen[1], Cu Nguyen[1], Barbara K. Dunn[2], Rhonda E. Ries[3], Hamid Bolouri[4], Jenny L. Smith[3], E. Anders Kolb[5], Soheil Meshinchi[3] and Daoud Meerzaman[1*]

1. Center for Biomedical Informatics and Information Technology, National Cancer Institute, Rockville, Maryland, United States of America
2. Division of Cancer Prevention, National Cancer Institute, Rockville, Maryland, United States of America
3. Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America
4. Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America
5. Nemours/Alfred I. duPont Hospital for Children, Wilmington, Delaware, United States of America

Corresponding author: Yu Fan, Center for Biomedical Informatics and Information Technology, National Cancer Institute, Rockville, Maryland, United States of America.

Received: May 01, 2020; Accepted: June 02, 2020; Published: June 04, 2020

Abstract

Gene expression profiling has been used to identify specific genes and pathways found within the different subsets of acute myeloid leukemia (AML). Prior studies have sought to characterize genetic differences within the subgroups themselves but no study to date has examined how gene expression patterns found in AML differ from normal hematopoiesis. In this study, we compared diagnostic bone marrow specimens obtained from 473 AML pediatric patients with normal bone marrow (NBM) from healthy individuals (N=20). An initial comparison of AML vs. NBM transcriptomes identified 2,625 differentially expressed genes (DEGs), with DUSP10 and MYBL1 as the most upregulated and downregulated genes. Comparing NBM with each AML cytogenetic subgroup resulted in 2,000-3,000 DEGs, with t(8;21) vs. NBM having the largest number of 3,367 DEGs. Three Cancer Testis Antigens (CTAs), MSLN, PRAME and CCNA1 were identified by all comparisons. Among them, CCNA1 was most significantly upregulated in AML vs. NBM. MSLN was highly expressed in most inv(16) samples whereas PRAME was highly expressed in the majority of t(8;21) samples. GSEA (Gene Set Enrichment Analysis) for DEGs showed tumorigenesis and immune related terms: inv(16) was linked to natural killer cell mediated cytotoxicity; t(8;21) and normal karyotype were related to T cell receptor signaling; MLL-rearrangement subtype was associated with PI3K-Akt signaling pathway. Distinctive linkage for increased or decreased interleukin levels were also indicated for each subtype. Together our findings suggest several genes and gene pathways within the AML subgroups that may serve as future biomarkers for disease or as targets for new immunotherapies.

Keywords: : Next-Generation Sequencing, NGS, RNA-seq, Pediatric Acute Myeloid Leukemia, AML, Gene Expression Profiling, Transcriptome, Differential Expression Analysis.

Introduction

Acute myeloid leukemia (AML) is composed of a group of hematopoietic malignancies. Those malignancies can be further classified into subgroups based on certain cytogenetic abnormalities found in bone marrow, including chromosomal alterations in t(8;21), inv(16), and MLL-rearrangement. The first two abnormalities, t(8;21) and inv(16), disrupt the function of core binding factors alpha (CBFA) and beta (CBFB), respectively, and are referred to collectively as CBF-AML. Abnormalities in the CBF or core binding factor gene are one of the most frequent aberrations found in AML. Acute leukemia bearing a mixed-lineage leukemia gene (MLL) rearrangement is most associated with an aggressive form of blood cancer found in pediatric patients. Next-generation sequencing (NGS), including RNAseq, has emerged as a powerful tool for discovering pathogenetic pathways and potential targets for clinical intervention in AML [1].Using whole-transcriptome sequencing, our previous work compared the transcriptome profiles of CBF-AML cases with cases that have unaltered (normal) karyotypes (NK). [2] That earlier analysis uncovered similarities and differences with respect to gene-expression signatures, splicing events, as well as RNA fusions, further defining the inv(16) versus the t(8;21) AML subtypes. A number of other studies have investigated the genomic and epigenomic features in adult AML blood samples[3-5]. Despite these prior omic analyses, no previous study

has specifically addressed omic differences between bone marrow from AML patients and normal bone marrow (NBM). A number of biomarkers, including several cancer testis antigens (CTAs), such as Mesothelin (MSLN), Preferentially Expressed Antigen in Melanoma (PRAME) , and cyclin-A1(CCNA1), as well as other genes involved in oncogenic pathways (e.g., cytokine receptor-like factor 2, CRLF2), are known to be overexpressed in AML [8-13].

In this study, we examine in greater depth the differences in gene expression profiles between AML and normal bone marrow. We compared N=473 pediatric AML bone marrow samples (inv(16) (n=60), t(8;21) (n=58), MLL (n=79), NK (n=104), other (n=139) and unknown (n=33)) with 20 NBM samples to investigate the differences between the two groups with respect to gene expression signatures and potential gene/pathway alterations. We expanded the list of differentially expressed genes in AML as well as possible gene networks for established biomarkers that were also dysregulated in AML. The subtle genomic discrepancies that emerged from our investigation may reflect alternate routes of AML pathogenesis and could hold potential for developing new biomarkers and treatment for disease.

Materials and Methods

Patient characteristics

The affected patient cohort included children and young adults diagnosed with de novo AML; 473 bone marrow specimens were obtained from these individuals. An additional 20 bone marrow specimens were derived from normal healthy individuals. All samples were obtained with written consent from the parents/guardians of the minors and were from the Children's Oncology Group clinical trial AAML1031. The Institutional Review Board at Fred Hutchinson Cancer Research Center reviewed and approved this study. It is filed under Institutional Review File #9950 (Biology of the Alterations of the Signal Transduction Pathway in Pediatric Cancer). All AML samples were enriched for mononuclear cells using the Ficoll enrichment process.

Rna sequencing in pediatric aml samples

RNA sequencing was performed on all 493 samples using the Illumina platform. Reads were mapped to 58,450 Ensembl Gene IDs, which belong to 31 gene biotypes, including protein-coding, non-coding, pseudogene, etc. Among Ensembl genes, approximately 21,500 were covered by at least one read and about 12,990 Ensembl genes had RPKM (Reads Per Kilobase per Million) mapped reads with values ≥ 1.

Identification of significant genes

Transcriptome sequencing data from the bone marrow specimens of 473 AML cases as a whole dataset or as individual AML cytogenetic subgroups were compared with 20 normal bone marrow (NBM) specimens to look for differences between AML and normal hematopoietic cells. The edgeR (version 3.30.3) [6,7], pipeline was applied to raw counts data for comparisons. After filtering genes with low expression, TMM (trimmed mean of M-values) normalization was performed to account for compositional difference between libraries. We then estimated dispersions and fit negative binomial (NB) generalized linear models (GLM) to each gene. Differential expression was determined using quasi-likelihood (QL) F-test. The log2 fold change (log2FC), p-value, and FDR (false discovery rate, Benjamini-Hochberg method) were calculated accordingly. We selected differentially expressed genes (DEGs) that occurred in significant levels using cutoffs $|log2FC| \geq 2$ and FDR<0.05. We generated violin plots using the in-house OmicPlot (v 0.1) R package to show gene expression distribution patterns.

Gene set enrichment analysis

Significant DEGs were subjected to Gene Set Enrichment Analysis (GSEA). We used an in-house OmicPath (v 0.1) R package to investigate potential molecular interactions and pathways associated with each set of DEGs identified. The OmicPath R package contains 24 databases as a resource for identifying existing pathways/networks. The hypergeometric distribution algorithm (or hypergeometric testing) was implemented within the GSEA application. Results with FDR<0.05 were considered significant. We mainly focused on the pathways from the KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene Ontology Consortium) databases as well as terms containing the keyword "leuk", which may be potentially related to AML. The UniProtKB [8], (http://www.uniprot.org) and Ingenuity® Pathway Analysis (IPA®) canonical pathways were also used to help inform pathway identification.

Results

Complete lists of DEGs can be found in Supplemental Table 1. Corresponding GSEA results showing significantly altered pathways are shown in Supplemental Table 2.

Overall AML compared with NBM

AML vs. NBM

To get a global sense of the differences between gene expression in the AML and NBM samples, we performed unsupervised clustering analysis. Multidimensional scaling (MDS) plots (Figure 1a) demonstrated variable expression patterns within AML, and minimal overlap was evident between the two groups.

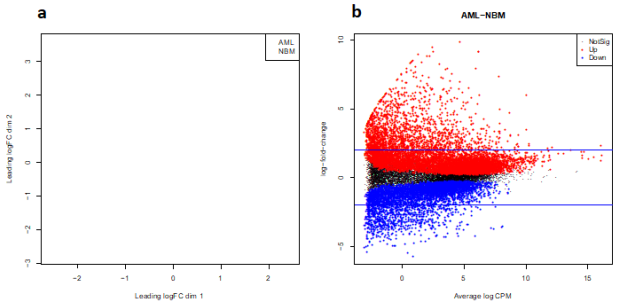

Figure 1. Multidimensional scaling (MDS) (a) and mean-difference (MD) (b) plots for NBM vs. AML. For plot (b), the log-fold-change for each gene is plotted against the average abundance; black dots stand for non-significant genes (FDR>0.05); red dots represent upregulated genes in AML and blue dots indicate downregulated genes. Horizontal blue lines show the fold-change cutoff as $|log2FC| \geq 2$.

As demonstrated in the MD plot (Figure 1b), statistical examination revealed that 11,966 genes were differentially expressed with FDR<0.05 in AML compared with NBM. By setting the fold-change cutoff as |log2FC|≥2, 2,625 genes were identified. Among these, 1,792 genes were upregulated whereas 833 genes were downregulated. Within the upregulated gene set, the DUSP10 (Dual specificity protein phosphatase 10) gene was the most significantly altered (log2FC=3.81; FDR=8.76E-29), followed by IRF2BP2 (Interferon regulatory factor 2-binding protein 2; log2FC=2.10; FDR=3.98E-22). The high expression of DUSP10 in AML was also reported in previous studies [9-11], among its mitogen-activated protein kinase (MAPK) targets, MAPK10 (JNK3) was the most upregulated one (log2FC=2.96; FDR= 0.000127). For the downregulated genes, MYBL1 (Myb-related protein A) was the most significantly altered (log-2FC=-4.60; FDR=6.74E-68), followed by FCRL6 (Fc receptor-like protein 6; log2FC=-5.69, FDR=1.05E-64). Although MYBL1 (A-Myb) and MYB (proto-oncogene c-Myb) are from the same gene family, they play distinct roles in different tissues [12,13], MYB, along with its co-regulators were not included in these 2,625 genes. High expression of MYBL1, but not MYB was observed in Burkitt's lymphoma, sIg+ B-acute lymphoblastic leukemia, and some chronic lymphocytic leukemias. [14], Unlike MYB, the role of MYBL1 in oncogenesis is not clearly established yet. The 20 most up- and down-regulated genes are shown in Figure 2.

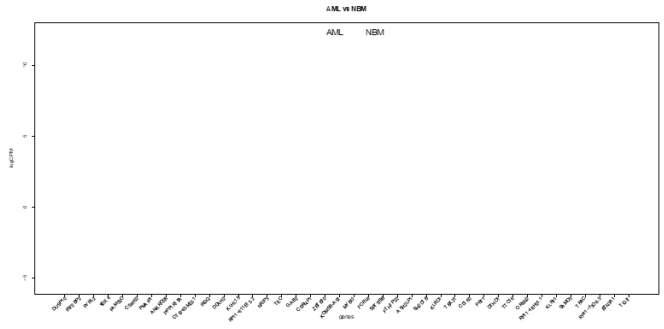

Figure 2. Top 20 up- and down-regulated genes from AML vs. NBM.

GSEA for DEGs

GSEA next was applied to the 2,625 differentially expressed genes with a focus on the KEGG and GO databases to elucidate the relevant pathways. The most significantly altered KEGG pathway was Hematopoietic cell lineage (hsa04640; FDR=0.0018), a known blood cancer–related network. Second was Primary immunodeficiency (hsa05340; FDR=0.0022), which is associated with pathogenesis in AML. Another KEGG pathway we identified was the Proteoglycans in cancer, which may suggest an essential role for protein glycosylation in AML. The MAPK signaling pathway and T cell receptor signaling pathway were also identified and are important in AML.

When GSEA was applied to the GO database, three terms were mapped in their entirety with extremely small p-values. These terms included leukocyte homeostasis (FLT3, NKX2-3) and neutrophil apoptotic process (IL6, HCAR2), and represent pathways that are upregulated. The other fully mapped term, which represented a downregulated pathway, was immunoglobulin production (VPREB1, IGKV1-5, IGKV4-1 and VPREB3). The GO pathway term that had the largest number of dysregulated genes (72 genes) was negative regulation of transcription from RNA polymerase II promoter (GO:0000122; FDR=0.022). The total number of genes in this pathway was 654.

Finally, a series of genes were found that related to increased interleukin-4 secretion, increased interleukin-17 secretion, decreased interleukin-5 secretion, increased circulating interleukin-18 level and decreased interleukin-2 secretion (MGI_GenePheno_hs; MGI_PhenoGenoMP_hs).

Stratification of GSEA by cytogenetic subtype

To better understand the differences that exist among cytogenetic subtypes, we compared four common AML karyotype classifications—inv(16), t(8;21), MLL, and NK—with NBM to establish a series of DEG lists and altered pathways. These genes and pathways enabled us to differentiate underlying features among the various subtypes. Likewise, unsupervised clustering analyses showed that each subtype separated well from NBM, and many of the subtype genes differed significantly in their expression compared with those from NBM. Here we summarize the significant GSEA results by cytogenetic subtypes.

Inv(16) vs. NBM

In addition to those identified from the initial comparison of overall AML to NBM, several new terms emerged. There were 27 genes belonging to the Natural killer cell mediated cytotoxicity pathway (hsa04650; FDR=0.0040), which had the smallest P-value. In GO, natural killer cell mediated immunity was also identified at the top with extremely small FDR. Other KEGG pathways of interest included NF-kappa B signaling pathway, Th17 cell differentiation, etc. These genes also were related to increased interleukin-4 secretion and increased interleukin-17 secretion.

t(8;21) vs. NBM

The top four KEGG pathways associated with this subtype were T cell receptor signaling pathway (hsa04660; FDR=0.0025), Primary immunodeficiency (hsa05340; FDR=0.0025), Viral protein interaction with cytokine and cytokine receptor (hsa04061; FDR=0.0032), and Rap1signaling pathway (hsa04015; FDR= 0.0032). B cell apoptotic process and natural killer cell mediated immunity were identified in GO with a very low FDR. Interleukin-related terms that emerged as significant were decreased interleukin-10 secretion and decreased interleukin-9 secretion.

MLL vs. NBM

Within the MLL subtype, the PI3K-Akt signaling pathway (hsa04151; FDR= 0.0033) and Chemokine signaling pathway (hsa04062; FDR=0.0043) were most significantly altered, along with Transcriptional misregulation in cancer (hsa05202; FDR=0.0063). In GO, cell morphogenesis involved in differentiation and RNA polymerase II transcription factor binding showed up first. Increased interleukin-4 secretion was also implicated.

NK vs. NBM

Within this subtype general terms also were identified, such as T cell receptor signaling pathway (FDR=0.0019) and MAPK

signaling pathway (FDR=0.0060). 24 genes were mapped to Gastric cancer pathway (hsa05226; FDR=0.0071). Decreased circulating interleukin-17 level had the highest significance.

Four previously established AML-related genes: MSLN, PRAME, CCNA1, and CRLF2

CCNA1, MSLN and PRAME, three CTAs, as well as CRLF2, all were included in the 2,625 DEGs from the initial comparison. These four genes had broader expression distributions in AML than in NBM. CCNA1 had the greatest median discrepancy between these two groups. (Figure 3a) By separating the AML samples based on their cytogenetic classification, we can see that MSLN and PRAME have the most cytogenetically distinct patterns: MSLN was highly expressed and similar to inv(16), MLL, and t(8;21) groups; the highest expression was observed in inv(16) samples. In contrast, PRAME was only highly expressed in the t(8;21) samples, which is consistent with previous studies. [15] (Figure 3b)

To investigate the underlying biological changes related to the high expression of MSLN and PRAME, we grouped all samples (N=493) based on customized logCPM (Counts Per Million) cutoffs to differentiate positive (+) and negative (-) expression for each. The level of positive expression for MSLN was 5 logCPM, which was same to PRAME. We then performed the same differential expression analysis using edgeR pipeline to identify DEGs (also setting the cutoff as |log2FC| ≥2 and FDR <0.05), followed by GSEA and pathway analysis. It was obvious that all NBMs were MSLN (-) and PRAME (-) (Figure 3b). With the 115 MSLN (+) AML samples, 290 DEGs were identified. Among the 92 PRAME (+) AML samples, 130 DEGs were identified. Additionally, when we applied the same analytic criteria for CCNA1 (logCPM cutoff=5) and CRLF2 (logCPM cutoff=1), 87 and 57 DEGs were identified among samples that

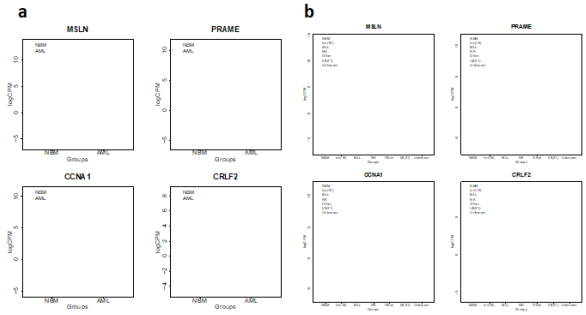

Figure 3. Expression patterns of MSLN, PRAME, CCNA1, and CRLF2 in AML vs. NBM (a) and among different cytogenetic subsets (b).

tested positive for the expression of those respective genes. The most significant GO/KEGG pathways suggested by the DEGs can be found in Table 1.

Discussion

The goal of this study was to compare gene expression from the bone marrow of AML patients with that of normal bone marrow to further characterize AML and its subgroups. Interrogating the transcriptome profile for all samples provided an overview of gene expression alterations in this disease. Interpreting individual subgroups using the same baseline fulfilled more knowledge for each subtype, which is helpful for designing useful biomarkers and therapies. Our analytic approach revealed genes and pathways unique to or shared by AML cytogenetic subtypes. By grouping samples based on expression cutoffs, we also expanded the potential gene networks for four established AML-related genes.

Our analyses revealed thousands of genes that were significantly and differentially expressed across a variety of AML sub-

Table 1. Top KEGG/GO pathways from DEGs mapped to four previously established AML-related genes.

Gene	Database	Gene mapped to Item	Item.ID	Item.name	FDR
MSLN	KEGG	CSF3 GP1BA GP5 IL9R ITGA2B ITGB3 MME	hsa04640	Hematopoietic cell lineage	0.0018
		CSF3 IL9R CCL3L3 PF4 PPBP CCL3 CCL3L1 CCL4 CCL20 CCL23 CCL24 CXCL12 GDF15	hsa04060	Cytokine-cytokine receptor interaction	0.0018
		ADCY2 ADCY9 ITGB3 CCL3L3 PLCB4 PTGER3 CCL3 CCL3L1 CCL4 CXCL12 CREB3L3	hsa05163	Human cytomegalovirus infection	0.0018
	GO	LRP4 MME SERPINF1	GO:0001822	kidney development	0.0046
PRAME	KEGG	LAMC3 COL4A5 COL6A6 CSF3 ITGB4 RELN BDNF SPP1	hsa04151	PI3K-Akt signaling pathway	0.00056
		CSF3 CXCL10 CCL20	hsa04657	IL-17 signaling pathway	0.0062
	GO	RELN LAMC3	GO:0000904	cell morphogenesis involved in differentiation	~0
CCNA1	KEGG	COL6A6 COL6A5 ITGB4 SPP1 CAV1	hsa04510	Focal adhesion	0.0073
	GO	KCNMA1 PLOD2	GO:0001666	response to hypoxia	0.014
		RAG1 SH2D1B	GO:0002250	adaptive immune response	0.014
CRLF2	KEGG	GP1BA GP5 IL9R	hsa04640	Hematopoietic cell lineage	0.0024
		COL6A6 FN1 COL6A5 SPP1	hsa04510	Focal adhesion	0.0024
		COL6A6 FN1 COL6A5 IGF2 SPP1	hsa04151	PI3K-Akt signaling pathway	0.0036
	GO	TMEM98 C6orf25	GO:0005783	endoplasmic reticulum	0.0032

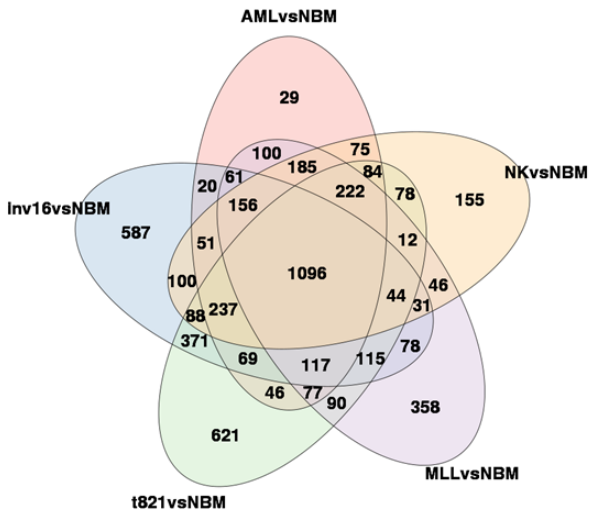

Figure 4. Five sets of all AML and subtype-specific comparator groups. The Venn diagrams show the differences and similarities among the five subtypes.

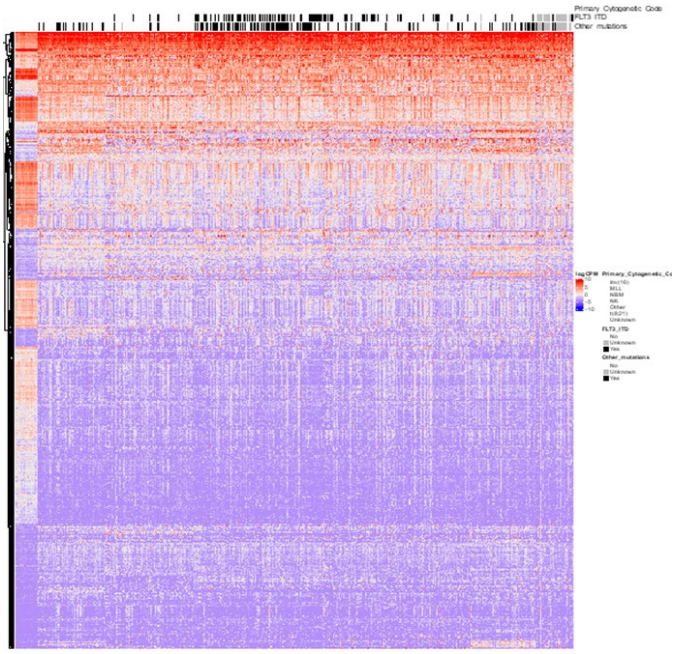

Figure 5. Heat map for NBM vs. AML with marked mutations. The different colors represent samples from different cytogenetic groups and are marked with the existence of FLT3/ITD (internal tandem duplication) and other mutations (NPM, CEBPA, WT1, or KIT mutations).

groups. The primary comparison showed a huge discrepancy between pathological AML bone marrow and normal bone marrow. In contrast to prior analyses using NK AML cases as the control which had confined the comparison between pathological samples, comparing the AML samples directly with NBM captured a much broader spectrum of transcriptome abnormalities. The various cytogenetic subtypes of AML were characterized according to different spectra of dysregulated genes and pathways, as well as mutations. The t(8;21) group had the largest number of DEGs and subtype specific DEGs compared with the other three subtypes, whereas NK had the smallest. The large overlap between the inv(16) vs. NBM and t(8;21) vs. NBM reflected in the 2,137 dysregulated genes suggests that extensive similarities exist between these two CBF-AML subgroups. Among these, 1,096 genes were identified in all five comparisons (all AML, inv(16), t(8;21), MLL, NK, each vs. NBM) (Figure 4). Figure 5 combines expression and mutation information for the 1,096 genes and shows distinct expression patterns for each cytogenetic subtype.

The two main data resources that we used for GSEA are KEGG and GO. Within the KEGG database, some smaller pathways were captured by more general KEGG terms used for other larger pathways. As a result, both large and small pathways may be identified in the same analysis. GSEA analysis using the GO database, which is based on different criteria for terminologies, yielded terms that contained only a few genes. These terms were more likely to be mapped with extremely small FDRs. In general, the most significant GSEA results (FDR<0.05) were immune response–related terms. This likely is because these terms are linked to other aspects of AML pathogenesis and because AML is often complicated by co-occurring diseases, such as infectious or autoimmune diseases that may or may not be related to the primary diagnosis of leukemia. Also, as expected, because AML is a highly proliferative and energy-consuming disease, many terms emerged that were related to basic metabolic mechanisms, cell cycle events, signaling pathways, and cell structure components.

Large portions of DEGs are shared. However, by using GSEA for individual cytogenetic subtypes and observing different significance levels of pathways we were able to gain further insight into these subtypes' unique mechanisms and clinical characteristics. Inv(16), t(8;21), and MLL were linked to NK cell mediated cytotoxicity, T cell receptor signaling, and PI3K-Akt signaling pathway, respectively. We also found distinct sets of interleukin levels, increased or decreased, within each subtype. The elevated number of FLT3/ITD mutations in NK bone marrow was another key finding, and supports previous observations that FLT3/ITD mutations are associated with normal cytogenetics in this disease [16], SRGN (Serglycin) was identified as a marker for immature myeloid cells and can be used to distinguish AML from several other hematologic malignancies [17]. In this study, it was most upregulated in the MLL subtype (log-2FC=2.06; FDR=1.37E-05).

CTAs are a family of well-characterized genes whose expression is mainly restricted to the testis, fetal tissues, and a variety of tumor types [18] Their strong in vivo immunogenicity make them potential prognostic/therapeutic biomarkers and ideal targets for tumor-specific immunotherapies [19]. MSLN, PRAME, and CCNA1 are known to be overexpressed in many AML patients. Not surprisingly, these three genes were found in significant proportions in our analyses. MSLN and PRAME often are used as part of a gene panel to sensitively and specifically monitor MRD (minimal residual disease), a disorder that is closely associated with remission and relapse in AML [20-23]. Given their AML-specific high expression and pro-oncogenic activities, these genes could be potential targets for developing new immunotherapy approaches for AML. Overexpression of CRLF2 was associated with greater activity in the JAK-STAT pathway. Increased expression of this gene is believed to be a factor in B-cell precursor acute lymphoblastic leukemia (ALL)

[24]. The gene may operate in a similar fashion in AML.

MSLN is involved in the cell cycle, cellular development, and cellular growth and proliferation. It has been studied as a potential biomarker for diagnosis and prognosis of ovarian cancer, mesothelioma, and pancreatic cancer (IPA Knowledgebase) [25-30]. Its overexpression in lung adenocarcinoma has been linked to tumor aggressiveness, higher risk of recurrence and poor outcomes [31].

PRAME expression is not found in normal bone marrow but is overexpressed in both acute and chronic hematological malignancies (AML, CML/chronic myelogenous leukemia, ALL, CLL/chronic lymphocytic leukemia, etc.), as well as in various solid tumors (skin, breast, lung, head and neck cancers; urological and neurological carcinomas; myeloma, etc.) [15,18,32-34]. PRAME is showing promise as an independent prognostic biomarker for breast cancer [18,35], and uveal melanoma, 36 and its expression may be useful in predicting increased metastatic risk or reduced survival. PRAME expression may have similar prognostic value in AML as it has been associated with increased overall survival and a favorable response to chemotherapy in patients, regardless of karyotype [15,37]. PRAME chiefly functions to repress retinoic acid receptor (RAR) signaling, which prevents retinoic acid (RA)-induced proliferation arrest, differentiation, and apoptosis [38], supporting the growth of cancer cells.

CCNA1 may help control cell cycle events at the G1/S (gap 1/DNA synthesis) and G2/M (gap 2/mitosis) transitions. It likely operates during meiosis in the germline and possibly in the mitotic cell cycle in some somatic cells. The Rb (retinoblastoma susceptibility gene) family of proteins and E2F-1 have been suggested as phosphorylation targets of CCNA1-associated kinase (UniProtKB) [39].The methylation status of CCNA1 is associated with HPV-induced cervical cancer and urothelial cell carcinoma [40-42]. Moreover, CCNA1 has been shown to be involved in the KEGG AML pathway (hsa05221) and linked with leukemogenesis and AML prognosis [43].

CRLF2 is the receptor for thymic stromal lymphopoietin (TSLP), which functions mainly on myeloid cells. CRLF2 forms a functional complex with TSLP and IL-7R and is capable of stimulating cell proliferation by activating phosphorylation in STAT3 and STAT5. Additional data suggest that CRLF2 may activate JAK2 and, thus, may contribute to the development of the hematopoietic system (UniProtKB) [44]. Patients with precursor B-ALL (pB-ALL) leukemia, an acute form of ALL, and who show the P2RY8-CRLF2 rearrangement and increased expression of CRLF2, have a significantly higher incidence of relapse (71% ± 19%) than patients without this mutation. Evaluating CRLF2 aberrations such as this may serve to further define this subpopulation, perhaps leading to more targeted treatment [45,46], Such assessments also may be applicable to AML patients.

Based on this evidence, we compared the transcription profiles of MSLN, PRAME, CCNA1, and CRLF2 in each AML subtype against those of NBM. Grouping samples by their expression levels of these four genes enabled in-depth analysis of their potential cooperators within certain networks. The DEGs and pathways we identified also intertwined with each other. For example, the PI3K-Akt signaling pathway was significant in both the PRAME-based and CRLF2-based results. Focal adhesion showed up in both the CCNA1-based and CRLF2-based analyses. These discoveries may aid in building new gene panels or networks that can be further refined to address AML.

Previous survival analyses of these four genes showed that MSLN's high expression levels were related not only to tumor progression but also to a better survival rate in gastric cancer [47].Increased expression of PRAME, along with an upregulation of the WT1 (Wilms' Tumor 1) gene, was associated with a favorable prognosis and is useful for monitoring minimal residual disease (MRD) in childhood ALL [48]. Finally, overexpression of CRLF2 indicated a poorer prognosis in pediatric T-cell ALL and in adult B-NEG ALL (B-cell acute lymphoblastic leukemia without recurrent fusion genes) patients [49,50].

Our study reveals both commonalities and differences in gene expression dysregulation among the various AML subgroups compared with NBM. These observations suggest potential diagnostic and therapeutic targets in AML that require further confirmation.

Abbreviations

AML: acute myeloid leukemia; NBM: normal bone marrow; DEG: differentially expressed gene; CTA: cancer testis antigen; GSEA: gene set enrichment analysis; CBF: core binding factor; NGS: next-generation sequencing; NK: normal karyotype; RPKM: reads per kilobase per million; TMM: trimmed mean of M-values; NB: negative binomial; GLM: generalized linear models; QL: quasi-likelihood; FC: fold change; FDR: false discovery rate; MDS: Multidimensional scaling; MD: mean-difference; CPM: counts per million.

Acknowledgements

The authors wish to acknowledge Xiaopeng Bian, Ph.D. for advice on data integration, Laura K. Fleming, Ph.D. and Barbara Vann for editorial critique.

Author Contributions

S.M., D.M., and Y.H. planned experiments; Y.F. analyzed data and wrote the manuscript. C.Y., Q.R., D.M., B.K.D., S.M., R.E.R., and J.L.S. provided advice for the manuscript. C.N., H.B., E.A.K. helped with the input data. All authors reviewed the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability

All data used or generated during this study have been deposited at The database of Genotypes and Phenotypes (dbGaP, http://www.ncbi.nlm.nih.gov/gap) under study accession phs000465.

References

1. Tarlock K, Meshinchi S. Pediatric acute myeloid leukemia: biology and therapeutic implications of genomic variants. Pediatr Clin North Am. 2015;62(1):75-93.
2. Hsu CH, Nguyen C, Yan C, et al. Transcriptome Profiling of Pediatric Core Binding Factor AML. PLoS One. 2015;10(9): e138782.
3. Cancer Genome Atlas Research N, Ley TJ, Miller C, et al. Ge-

nomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059-2074.
4. Lilljebjorn H, Agerstam H, Orsmark-Pietras C, et al. RNA-seq identifies clinically relevant fusion genes in leukemia including a novel MEF2D/CSF1R fusion responsive to imatinib. Leukemia. 2014;28(4):977-979.
5. Macrae T, Sargeant T, Lemieux S, Hebert J, Deneault E, Sauvageau G. RNA-Seq reveals spliceosome and proteasome genes as most consistent transcripts in human cancer cells. PLoS One. 2013;8(9):e72884.
6. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-140.
7. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40(10):4288-4297.
8. UniProt C. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47(D1):D506-D515.
9. Arora D, Kothe S, van den Eijnden M, et al. Expression of protein-tyrosine phosphatases in Acute Myeloid Leukemia cells: FLT3 ITD sustains high levels of DUSP6 expression. Cell Commun Signal. 2012;10(1):19.
10. Kabir NN, Ronnstrand L, Kazi JU. Deregulation of protein phosphatase expression in acute myeloid leukemia. Med Oncol. 2013;30(2):517.
11. Jimenez-Martinez M, Stamatakis K, Fresno M. The Dual-Specificity Phosphatase 10 (DUSP10): Its Role in Cancer, Inflammation, and Immunity. Int J Mol Sci. 2019;20(7).
12. Oh IH, Reddy EP. The myb gene family in cell growth, differentiation and apoptosis. Oncogene. 1999;18(19):3017-3033.
13. Pattabiraman DR, Gonda TJ. Role and potential for therapeutic targeting of MYB in leukemia. Leukemia. 2013;27(2):269-277.
14. Golay J, Luppi M, Songia S, et al. Expression of A-myb, but not c-myb and B-myb, is restricted to Burkitt's lymphoma, sIg+ B-acute lymphoblastic leukemia, and a subset of chronic lymphocytic leukemias. Blood. 1996;87(5):1900-1911.
15. Wadelin F, Fulton J, McEwan PA, Spriggs KA, Emsley J, Heery DM. Leucine-rich repeat protein PRAME: expression, potential functions and clinical implications for leukaemia. Mol Cancer. 2010;9:226.
16. Stirewalt DL, Kopecky KJ, Meshinchi S, et al. FLT3, RAS, and TP53 mutations in elderly patients with acute myeloid leukemia. Blood. 2001;97(11):3589-3595.
17. Niemann CU, Kjeldsen L, Ralfkiaer E, Jensen MK, Borregaard N. Serglycin proteoglycan in hematologic malignancies: a marker of acute myeloid leukemia. Leukemia. 2007;21(12):2406-2410.
18. Goodison S, Urquidi V. The cancer testis antigen PRAME as a biomarker for solid tumor cancer management. Biomark Med. 2012;6(5):629-632.
19. Fratta E, Coral S, Covre A, et al. The biology of cancer testis antigens: putative function, regulation and therapeutic potential. Mol Oncol. 2011;5(2):164-182.
20. Atanackovic D, Luetkens T, Kloth B, et al. Cancer-testis antigen expression and its epigenetic modulation in acute myeloid leukemia. Am J Hematol. 2011;86(11):918-922.
21. Ochsenreither S, Majeti R, Schmitt T, et al. Cyclin-A1 represents a new immunogenic targetable antigen expressed in acute myeloid leukemia stem cells with characteristics of a cancer-testis antigen. Blood. 2012;119(23):5492-5501.
22. Ortmann CA, Eisele L, Nuckel H, et al. Aberrant hypomethylation of the cancer-testis antigen PRAME correlates with PRAME expression in acute myeloid leukemia. Ann Hematol. 2008;87(10):809-818.
23. Steinbach D, Schramm A, Eggert A, et al. Identification of a set of seven genes for the monitoring of minimal residual disease in pediatric acute myeloid leukemia. Clin Cancer Res. 2006;12(8):2434-2441.
24. Russell LJ, Capasso M, Vater I, et al. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. Blood. 2009;114(13):2688-2698.
25. Hucl T, Brody JR, Gallmeier E, Iacobuzio-Donahue CA, Farrance IK, Kern SE. High cancer-specific expression of mesothelin (MSLN) is attributable to an upstream enhancer containing a transcription enhancer factor dependent MCAT motif. Cancer Res. 2007;67(19):9055-9065.
26. Hassan R, Remaley AT, Sampson ML, et al. Detection and quantitation of serum mesothelin, a tumor marker for patients with mesothelioma and ovarian cancer. Clin Cancer Res. 2006;12(2):447-453.
27. Frierson HF, Jr., Moskaluk CA, Powell SM, et al. Large-scale molecular and tissue microarray analysis of mesothelin expression in common human carcinomas. Hum Pathol. 2003;34(6):605-609.
28. McCarthy DM, Maitra A, Argani P, et al. Novel markers of pancreatic adenocarcinoma in fine-needle aspiration: mesothelin and prostate stem cell antigen labeling increases accuracy in cytologically borderline cases. Appl Immunohistochem Mol Morphol. 2003;11(3):238-243.
29. Ordonez NG. Application of mesothelin immunostaining in tumor diagnosis. Am J Surg Pathol. 2003;27(11):1418-1428.
30. Kendrick ZW, Firpo MA, Repko RC, et al. Serum IGFBP2 and MSLN as diagnostic and prognostic biomarkers for pancreatic cancer. HPB (Oxford). 2014;16(7):670-676.
31. Kachala SS, Bograd AJ, Villena-Vargas J, et al. Mesothelin overexpression is a marker of tumor aggressiveness and is associated with reduced recurrence-free and overall survival in early-stage lung adenocarcinoma. Clin Cancer Res. 2014;20(4):1020-1028.
32. Roman-Gomez J, Jimenez-Velasco A, Agirre X, et al. Epigenetic regulation of PRAME gene in chronic myeloid leukemia. Leuk Res. 2007;31(11):1521-1528.
33. Oehler VG, Guthrie KA, Cummings CL, et al. The preferentially expressed antigen in melanoma (PRAME) inhibits myeloid differentiation in normal hematopoietic and leukemic progenitor cells. Blood. 2009;114(15):3299-3308.
34. Pellat-Deceunynck C, Mellerin MP, Labarriere N, et al. The cancer germ-line genes MAGE-1, MAGE-3 and PRAME are commonly expressed by human myeloma cells. Eur J Immunol. 2000;30(3):803-809.
35. Doolan P, Clynes M, Kennedy S, Mehta JP, Crown J, O'Driscoll L. Prevalence and prognostic and predictive relevance of PRAME in breast cancer. Breast Cancer Res Treat. 2008;109(2):359-365.
36. Field MG, Decatur CL, Kurtenbach S, et al. PRAME as an Independent Biomarker for Metastasis in Uveal Melanoma. Clin Cancer Res. 2016;22(5):1234-1242.
37. Santamaria C, Chillon MC, Garcia-Sanz R, et al. The relevance of preferentially expressed antigen of melanoma (PRAME) as a

marker of disease activity and prognosis in acute promyelocytic leukemia. Haematologica. 2008;93(12):1797-1805.

38. Epping MT, Wang L, Edel MJ, Carlee L, Hernandez M, Bernards R. The human tumor antigen PRAME is a dominant repressor of retinoic acid receptor signaling. Cell. 2005;122(6):835-847.
39. Yang R, Muller C, Huynh V, Fung YK, Yee AS, Koeffler HP. Functions of cyclin A1 in the cell cycle and its interactions with transcription factor E2F-1 and the Rb family of proteins. Mol Cell Biol. 1999;19(3):2400-2407.
40. Zuo Q, Zheng W, Zhang J, et al. Methylation in the promoters of HS3ST2 and CCNA1 genes is associated with cervical cancer in Uygur women in Xinjiang. Int J Biol Markers. 2014;29(4):e354-362.
41. Yang N, Eijsink JJ, Lendvai A, et al. Methylation markers for CCNA1 and C13ORF18 are strongly associated with high-grade cervical intraepithelial neoplasia and cervical cancer in cervical scrapings. Cancer Epidemiol Biomarkers Prev. 2009;18(11):3000-3007.
42. Maldonado L, Brait M, Michailidi C, et al. An epigenetic marker panel for recurrence risk prediction of low grade papillary urothelial cell carcinoma (LGPUCC) and its potential use for surveillance after transurethral resection using urine. Oncotarget. 2014;5(14):5218-5233.
43. Stirewalt DL, Meshinchi S, Kopecky KJ, et al. Identification of genes with abnormal expression changes in acute myeloid leukemia. Genes Chromosomes Cancer. 2008;47(1):8-20.
44. Reche PA, Soumelis V, Gorman DM, et al. Human thymic stromal lymphopoietin preferentially stimulates myeloid cells. J Immunol. 2001;167(1):336-343.
45. Cario G, Zimmermann M, Romey R, et al. Presence of the P2RY8-CRLF2 rearrangement is associated with a poor prognosis in non-high-risk precursor B-cell acute lymphoblastic leukemia in children treated according to the ALL-BFM 2000 protocol. Blood. 2010;115(26):5393-5397.
46. Chen IM, Harvey RC, Mullighan CG, et al. Outcome modeling with CRLF2, IKZF1, JAK, and minimal residual disease in pediatric acute lymphoblastic leukemia: a Children's Oncology Group study. Blood. 2012;119(15):3512-3522.
47. Baba K, Ishigami S, Arigami T, et al. Mesothelin expression correlates with prolonged patient survival in gastric cancer. J Surg Oncol. 2012;105(2):195-199.
48. Khateeb EE, Morgan D. Preferentially Expressed Antigen of Melanoma (PRAME) and Wilms' Tumor 1 (WT 1) Genes Expression in Childhood Acute Lymphoblastic Leukemia, Prognostic Role and Correlation with Survival. Open Access Maced J Med Sci. 2015;3(1):57-62.
49. Palmi C, Savino AM, Silvestri D, et al. CRLF2 over-expression is a poor prognostic marker in children with high risk T-cell acute lymphoblastic leukemia. Oncotarget. 2016;7(37):59260-59272.
50. Chiaretti S, Brugnoletti F, Messina M, et al. CRLF2 overexpression identifies an unfavourable subgroup of adult B-cell precursor acute lymphoblastic leukemia lacking recurrent genetic abnormalities. Leuk Res. 2016;41:36-42.

To cite this article: Fan Yu, Hu Y, Yan C, et al. Altered Transcriptome in Pediatric AML Compared with Normal Hematopoiesis. British Journal of Cancer Research. 2020; 3:3.