Am J Clin Exp Immunol 2024;13(6):246-258
www.ajcei.us /ISSN:2164-7712/AJCEI0162141

Original Article

Predicting candidate biomarkers for COVID-19 associated with leukemia in children

Judy Bai[1], Qing Li[2]

[1]*Greenhills School, Ann Arbor, MI 48105, USA;* [2]*Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109, USA*

Received November 19, 2024; Accepted December 1, 2024; Epub December 25, 2024; Published December 30, 2024

Abstract: Since the COVID-19 pandemic, a significant number of pediatric leukemia patients have shown to have also contracted COVID-19 several weeks or months prior to the development of their cancer. Current research indicates the expression of MDA5, encoded by *IFIH1*, is associated with increased immunity to COVID-19 in children. Children are also known to have a much lower risk of developing leukemia. Our hypothesis is that *IFIH1* and its regulatory miRNAs are biomarkers associated with pediatric leukemia; the objective of our study is to identify genes, through miRNA targeting mechanisms, which may be biomarkers associated with COVID-19 infection and leukemia. The database TarBase was analyzed to identify miRNAs that target *IFIH1*, followed by the identification of other genes regulated by *IFIH1*'s targeting miRNAs, to construct a gene-miRNA targeting network. Protein-Protein Interaction (PPI) analysis and DAVID/KEGG pathway analysis were conducted to identify genes with meaningful biological interactions and pathways. We identified two significant miRNAs, *hsa-196a-5p* and *hsa-196b-5p*, and 51 of their targeted and highly expressed genes reported in the Acute Myeloid Leukemia (AML) samples from The Cancer Genome Atlas (TCGA) RNA sequencing database. When conducting additional analysis using the Gene Constellation module of the Immunological Genome Project for the top three candidate genes, several other genes were identified to be highly correlated with *STAT3* and *IFIH1* in our study. Based on our investigation into co-expression analysis, we found that *IFIH1* is a potential biomarker for AML. We are expanding our work to create a machine learning model to identify other biomarkers, examine the significance of various parameters (age, race, etc.), and perform comorbidity network analysis for other potential genes/miRNAs.

Keywords: Leukemia, COVID-19, RNA-Seq, differential expression

Introduction

Acute Myeloid Leukemia (AML) has a higher mortality rate compared to Acute Lymphoblastic Leukemia (ALL) in children, with most cases occurring in children who are either younger than 2 years old or teenagers. In AML, myeloid stem cells differentiate into unhealthy white blood cells called "blasts". These blasts, instead of fighting off infection and helping to protect the body, build up in bone marrow and leave less room for healthy blood cells. This type of leukemia also does not always stay in the blood or bone marrow and may travel to the skin, spine, or other organs.

Many studies have consistently reported that many children with AML have experienced a previous episode of COVID infection weeks to months prior to the onset of their cancer, suggesting an association between COVID-19 infection and AML development. In one study, the researchers discovered that the correlation between leukemia and COVID-19 infection is the strongest when compared to other cancers [1]. This is consistent with results from another study, which showed that COVID-19 infection is associated with acute leukemia and severe bone marrow involvement [2]. COVID-19 has also been associated with other types of blood cancers. For example, the diagnosis of Hair Cell Leukemia (HCL) in COVID-19 patients has been a reported relationship [3]. In addition, children who have blood malignancy are more likely to contract significant COVID-19 infection. Furthermore, patients with COVID-19 have an increased risk of developing a life-threatening illness if they are receiving one or more treat-

https://doi.org/10.62347/ULTA9461

ments intended to treat their blood cancer [4]. These studies highlight the possible association between COVID-19 infection and AML.

The goal of our project is to identify genome-wide candidate genes which could serve, in the context of miRNA targeting mechanisms, as biomarkers of COVID-19 and leukemia. Many studies show that children are less susceptible to contracting COVID-19 and are also less susceptible to developing AML compared to adults. Children's heightened immunity to COVID-19 has been proposed to stem from their high expression of the *IFIH1* gene, which encodes the MDA5 protein. It is known that MDA5 recognizes double-stranded RNA and prompts the innate immune response [5]. We used *IFIH1* as an entry point of the study, analyzed it using multiple databases, and identified potential upstream regulatory miRNAs of *IFIH1* and other potential target genes of *IFIH1* that are dysregulated in AML.

To examine the evident relationship between COVID-19 and AML, we used data from The Cancer Genome Atlas (TCGA), which has over 20,000 original tumor and matched normal samples from 33 different studied cancer types and is a part of a significant cancer genomics initiative project [6] (https://www.cancer.gov/ccg/research/genome-sequencing/tcga). The Level 3 (de-identified) RNA expression data generated from TCGA are publicly available to the community.

We also wanted to incorporate microRNAs (or miRNAs) into our study, since miRNAs are important regulators of cell function and gene expression and they have also not been widely examined in COVID-19 and leukemia. MicroRNAs are small (~17-22 bp) non-coding RNAs that can regulate gene expression through many mechanisms, including binding to the 3' Untranslated Region (UTR) of an mRNA and thus reducing protein translation efficiency. Additionally, miRNA dysregulation in cancer has been reported [7] for low levels of miRNA expression, which most likely causes overexpression of their target genes [8].

As more scientists recognize disease-causing mechanisms through miRNA post-transcriptional regulation channels, the process of identifying connections for genes and targeting miRNAs between diseases has become more streamlined. The correlation of miRNA-related diseases and potential influences have also been reported previously in other studies [9]. Our study has identified several candidate genes and miRNAs in the context of miRNA targeting mechanisms, based on TCGA Acute Myeloid Leukemia (LAML) RNA sequencing data, that could serve as comorbidity biomarkers of diseases (i.e., Coronavirus disease and leukemia cancer).

Material and methods

Identification of miRNAs that regulate the hub gene IFIH1 and their other target genes

Because previous studies have shown that high expression of *IFIH1* in children plays a role in their resistance to COVID-19, we first decided to search for miRNAs that target *IFIH1*. Using TarBase 7.0 [10], we inputted *IFIH1* and selected "*Homo sapiens*" as the species. Two upstream miRNAs (*hsa-miR-196a-5p* and *hsa-miR-196b-5p*) from the same family that had a prediction score greater than 0.7 were identified. Since miRNAs usually target multiple genes, we also decided to search for other genes targeted by the two miRNAs using TarBase. There were 77 targeted genes reported for *hsa-miR-196b-5p* and 101 targeted genes reported for *hsa-miR-196a-5p*.

Evaluation of expression of the target genes of miRNAs from TCGA RNA-Seq data

We employed putty Unix commands to search the above identified miRNA-targeted gene lists against TCGA LAML RNA-Seq data to retrieve expression information in terms of Transcripts Per Million (TPM). We wanted to focus on the genes that had the highest expression in leukemia, for those miRNA-targeted genes presumably have the highest effect (influence) on the development of leukemia. After obtaining the resulting lists of genes, we selected the ones with significantly high expressions (>500 TPM). From this, we acquired 51 candidate genes with a TPM score >500.

Protein-protein interaction of targeted genes

We used the STRING database [11] to conduct Protein-Protein Interaction (PPI) between the 51 genes with a significantly high expression (>500 TPM) to identify the genes that directly

interact with *IFIH1*. The default search parameters were employed to generate the interaction network.

Expression levels of IFIH1, STAT3, and MAP3K1 in acute myeloid leukemia

Genes that are highly expressed in certain diseases are usually an indicator that they play an important role in those diseases. Therefore, we wanted to verify that the three candidate genes, *IFIH1, STAT3,* and *MAP3K1*, are highly expressed in Acute Myeloid Leukemia by obtaining their expression data and determining if their expressions are higher in AML. To do this, we inputted all three candidate genes into the UCSC Xena database [12], selected the GDC TARGET-AML pediatric patient dataset [13], and obtained the graphs of each candidate genes' expression levels.

The expressions of IFIH1, STAT3, and MAP3K1 in normal blood cells

It is important to understand the expressions of these potential biomarkers in normal tissues; thus, we examined the expression levels of the three genes in published databases. We searched Immunological Genome Project (ImmGen) databases [14] for the three genes' expression data in mice and found that all three genes are ubiquitously expressed in all cell types (data not shown). We then focused on immune cells when retrieving expression information for all three candidate genes. We individually entered each candidate gene into the database (Results would be more detailed with individual genes compared to multi-gene queries) and observed the cell types that each gene was differentially expressed in.

Pathway analysis of IFIH1, STAT3, and MAP3K1

To identify the common pathways these three genes are involved, we ran DAVID [15]/KEGG [16] pathway databases to search for the functional importance for the three candidate genes (*IFIH1, STAT3, MAP3K1*). We entered the list of the three candidate genes into the database, selected "*Homo sapiens*" as the species, and selected "OFFICIAL_GENE_SYMBOL" as the ID. To ensure that we obtain the most significant and accurate results, we chose the "Disease" section (In hopes of seeing AML) and explored the first result that was shown under the "Disease" category.

Functional enrichment analysis of IFIH1 and STAT3

In the above pathway analysis, the Hepatitis B pathway was identified as the top hit (with the highest significance) out of the other pathways that the three candidate genes were involved in. We then wanted to confirm this through an independent analysis. The g:Profiler tool [17] was used to search for any roles/diseases the genes had. Because *MAP3K1* showed a few inconsistent results, we focused on the other two candidate genes *(IFIH1* and *STAT3)* and searched for functional enrichment or over-representation of the genes. Specifically, the two candidate genes were inputted into g:Profiler to detect statistically significant enriched terms that the genes were involved in. The enrichment analysis was performed using a cumulative hypergeometric test, and the non-standard multiple testing correction was adopted by g:Profiler.

Correlation of the expressions of IFIH1, STAT3, and MAP3K1

We next wanted to test if there was a correlation of the expressions of the three genes. Thus, we used the UCSC Xena database [12], an online exploration tool for multi-omic and clinical/phenotype data. We entered the three genes into the database and retrieved expression correlation graphs between all three genes, obtaining 3 sets of pair correlations. The genes' expressions were from the GDC TARGET AML dataset [13] to ensure that the three genes positively correlated with one another while also confirming that they are linked to the development of AML.

Identification of additional genes connected to the candidate genes that are potentially linked to AML

We also sought to identify additional genes that could possibly be related to *IFIH1* and *STAT3* to expand the number of drug target options for AML and COVID-19. Because *MAP3K1* analysis had given inconsistent results, we did not include this gene in the analysis. We employed ImmGen's gene constellations to model corre-

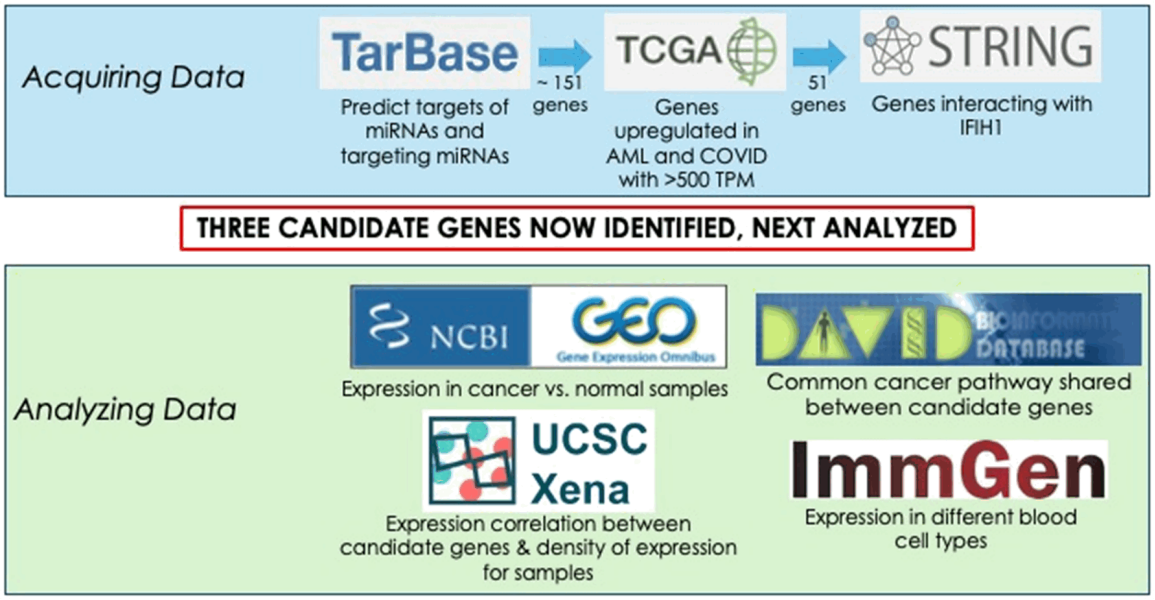

Figure 1. Workflow of bioinformatics tools and databases used in analysis.

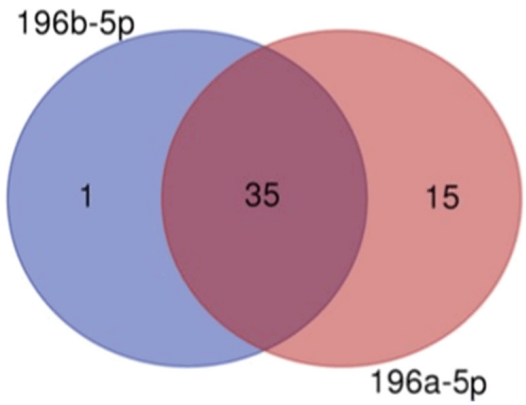

Figure 2. Target genes of miRNAs *hsa-miR-196a-5p* and *hsa-miR-196b-5p* with high expression in AML. The target genes of both miRNAs were filtered to only include the genes with a TPM score of >500 TPM in the TCGA LAML database; 51 genes were selected as a result. Of the 51 target genes, 35 are shared by both miRNAs, one is unique to *hsa-miR-196b-5p*, 15 are unique to *hsa-miR-196a-5p*.

lations between other different genes to find any additional genes that have a relationship with *IFIH1* and *STAT3*.

Verification for the interaction between the three candidate genes and the COVID-19 protein (SARS-CoV-2)

To test if the MDA5 and SARS-CoV-2 protein molecules interact for MDA5, we obtained the crystal structures of both from the Protein Data Bank. Specifically, we obtained 3GA3 (Crystal structure of the C-terminal domain of human MDA5), 6LU7 (SARS-CoV-2 protein), and 6NJS (*STAT3* Core in complex with compound SD36). We then used the ZDOCK server [18, 19] to visualize all the proteins' interactions. After obtaining the candidate genes' protein data from the Protein Data Bank and inputting them into the ZDOCK server, we developed two models that show the protein structure needed to ensure that the three proteins interact with each other.

The overall analysis workflow for this study is shown in **Figure 1**.

Results

MiRNAs targeting the hub genes

Because previous studies have shown that high expression of *IFIH1* in children contributes to their resistance against COVID-19, we decided to search for miRNAs that targeted *IFIH1*. Two upstream miRNAs (*hsa-miR-196a-5p* and *hsa-miR-196b-5p*) from the same family that had a prediction score greater than 0.7 were identified. Since miRNAs usually target multiple genes, we also decided to search for other genes targeted by the two miRNAs using TarBase [10]. There were 77 target genes reported for *hsa-miR-196b-5p* and 101 target genes reported for *hsa-miR-196a-5p*.

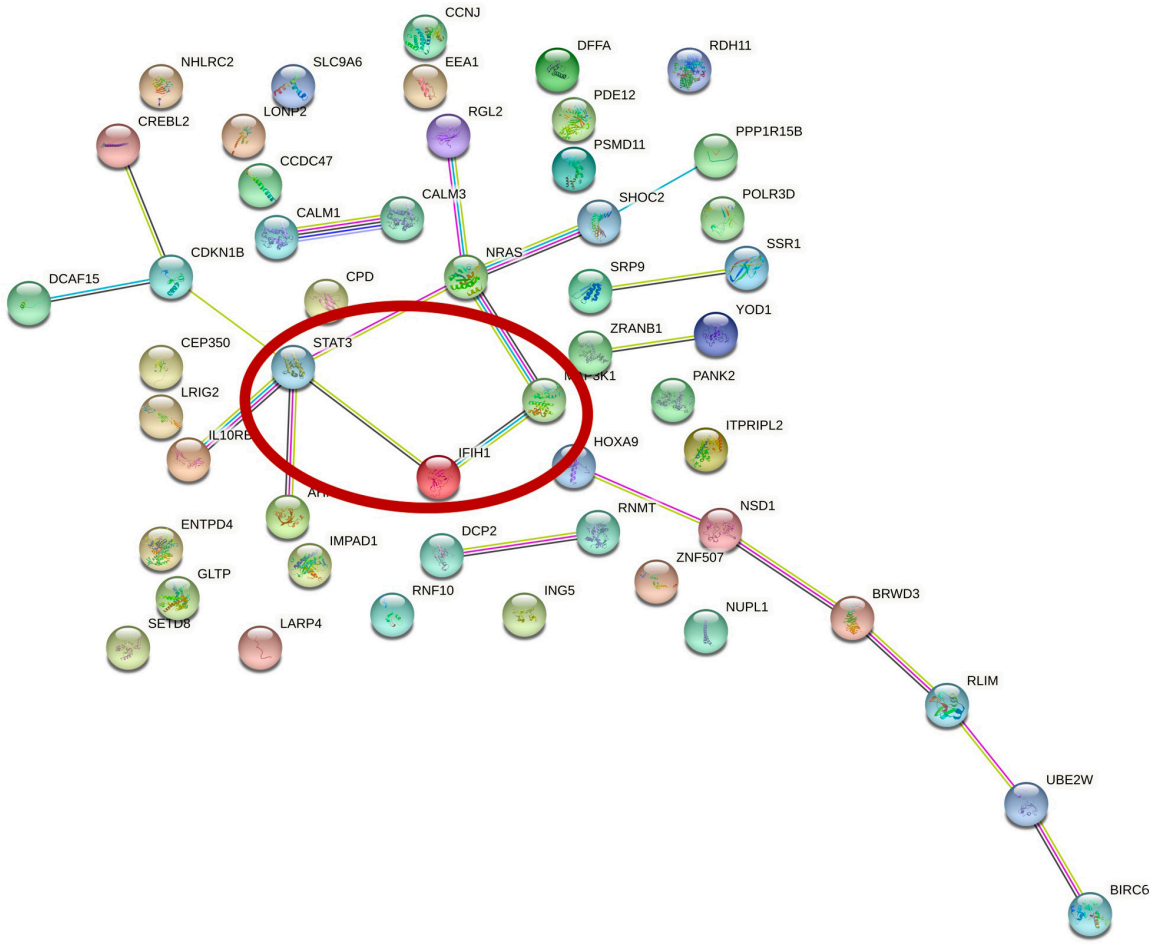

Figure 3. *IFIH1* is associated with *STAT3* and *MAP3K1*. After inputting the 51 potential target genes of miRNA hsa-miR-196b-5p and hsa-miR-196a-5p into the STRING database, we found that *IFIH1* had no direct connections other than the two genes, *STAT3* and *MAP3K1*, which were reported to closely interact with *IFIH1*.

Expression of the target genes of the two miRNAs

We employed putty Unix commands to search the above identified miRNA targeted gene lists against TCGA LAML RNA-Seq data to retrieve expression information (Transcripts Per Million or TPM).

Since both miRNAs were from the same family, 51 genes were obtained with a Transcript Per Million (TPM) score higher than 500 (**Figure 2**) from the combined list for downstream analyses (The target gene lists for the two miRNAs were combined because more genes could be found to be targets of the two miRNAs). Since a TPM score higher than 500 is a commonly used cut off, this threshold was used in LAML tumor data to narrow down the list of 77 genes to 36 genes for *hsa-miR-196b-5p* and the list of 101 genes to 50 genes for *hsa-miR-196a-5p*.

Protein-Protein Interaction gene set for IFIH1 gene

STRING database [11] search resulted in the identification of *STAT3* and *MAP3K1*, which were reported to closely interact with *IFIH1* (**Figure 3**).

Expression levels of IFIH1, STAT3, and MAP3K1

The UCSC Xena database [12] was used to produce the graphs of each candidate genes' expression levels. All three genes are highly expressed in AML samples (**Figure 4**). Taking into account that the minimum value considered high expression is 9.965, the values for

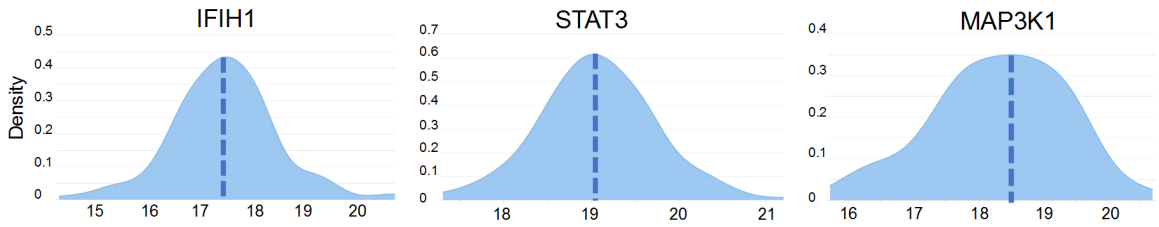

Figure 4. Majority of AML samples from TARGET-AML trial have high expression of *IFIH1, STAT3*, and *MAP3K1*. All three candidate genes have a significantly high expression in the GDC TARGET AML dataset [11], thus all three genes are involved in the development of AML.

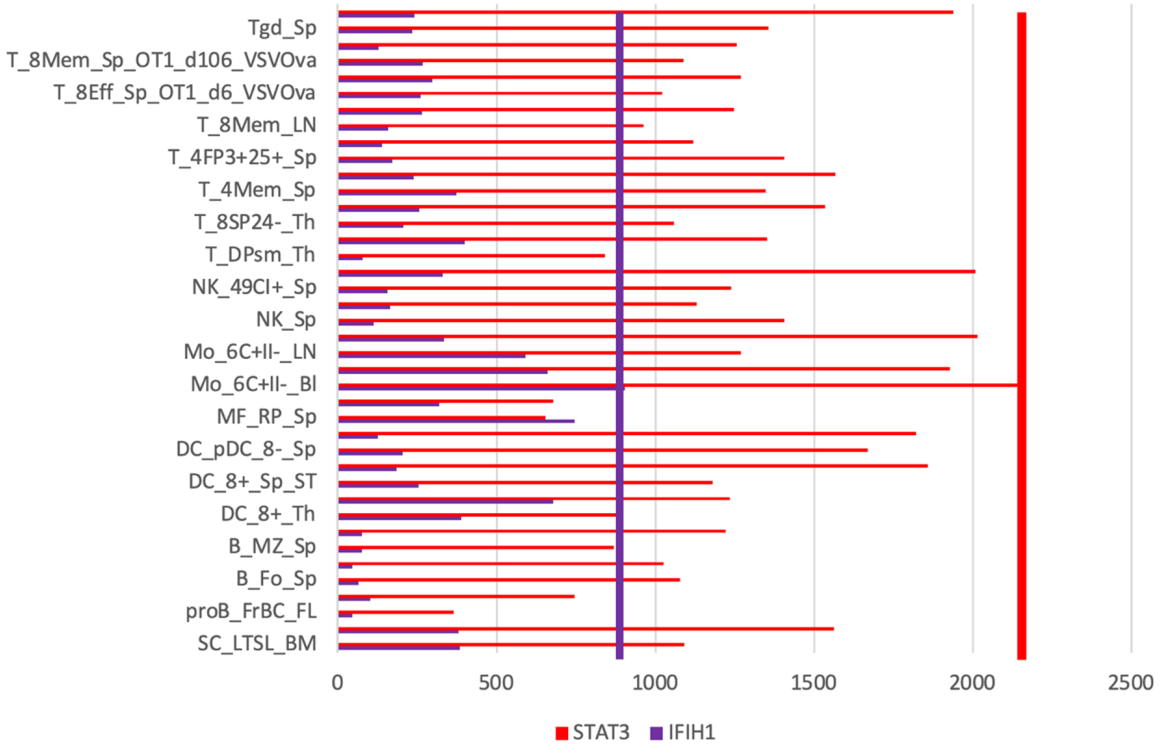

Figure 5. Expression of *STAT3* and *IFIH1* in immune cells. ImmGen [14] expression analysis of the mice model showed that the two genes, *STAT3* and *IFIH1*, are highly expressed in CD16+ monocytes. The purple line shows *IFIH1*'s highest expression, which is in monocytes, and the red line shows *STAT3*'s highest expression, which is in the same immune cell as *IFIH1*, monocytes. These results show that *IFIH1* and *STAT3* are linked. The *MAP3K1* gene is not shown in this figure because it did not present significant results for this test.

IFIH1, STAT3 and *MAP3K1* are 17.5, 19.1 and 18.5, respectively.

The ImmGen expressions of IFIH1, STAT3, and MAP3K1

To understand the expression of these potential biomarkers in normal tissues, we examined the levels of these three genes in published databases. In this case, we used the ImmGen databse. As shown in **Figure 5**, *IFIH1* and *STAT3* were highest expressed in monocytes, and *MAP3K1* was highest expressed in whole blood cells. Thus, *IFIH1, STAT3*, and *MAP3K1* could all be associated with leukemia, for leukemia affected the cells that all three genes were highest expressed in.

Pathway results of IFIH1, STAT3, and MAP3K1

It is important to identify the common pathways these three genes are involved in together, so we ran the DAVID [15]/KEGG [16] pathway databases to search for the functional impor-

Table 1. Analysis of pathways where *IFIH1, STAT3* and *MAP3K1* are involved

Term	Genes	*P*-value	Fold Enrichment
hsa05161: Hepatitis B	*IFIH1, MAP3K1, STAT3*	4.4130923260829284E -4	47.44137931034483
GO:0032755~positive regulation of interleukin-6 production	*IFIH1, STAT3*	1.0766391720932331E -2	123.5111111111111
GOTERM_BP_DIRECT GO:0032760~positive regulation of tumor necrosis factor production	*IFIH1, STAT3*	1.1073146675286262E -2	120.08024691358024
hsa04622: RIG-I-like receptor signaling pathway	*IFIH1, MAP3K1*	1.6361056013736306E -2	81.16431924882629
h_egfPathway: EGF Signaling Pathway	*MAP3K1, STAT3*	1.663585951940162E -2	60.111111111111114
h_pdgfPathway: PDGF Signaling Pathway	*MAP3K1, STAT3*	1.725200246456486E -2	57.96428571428572
KW-0219~Diabetes mellitus	*IFIH1, STAT3*	1.8166455428756008E -2	55.04651162790697
hsa04935: Growth hormone synthesis, secretion and action	*MAP3K1, STAT3*	2.7573784811262776E -2	48.02222222222222
hsa05162: Measles	*IFIH1, STAT3*	3.167660401229694E -2	41.7584541062802
hsa05171: Coronavirus disease - COVID-19	*IFIH1, STAT3*	5.296151957686342E -2	24.839080459770116

DAVID [15] functional annotation top 10 most significant results for candidate genes (*IFIH1, STAT3, MAP3K1*). The fold enrichment score and *p*-value for the top hit, Hepatitis B, is extremely significant, showing that all three genes are also involved in Hepatitis B pathway.

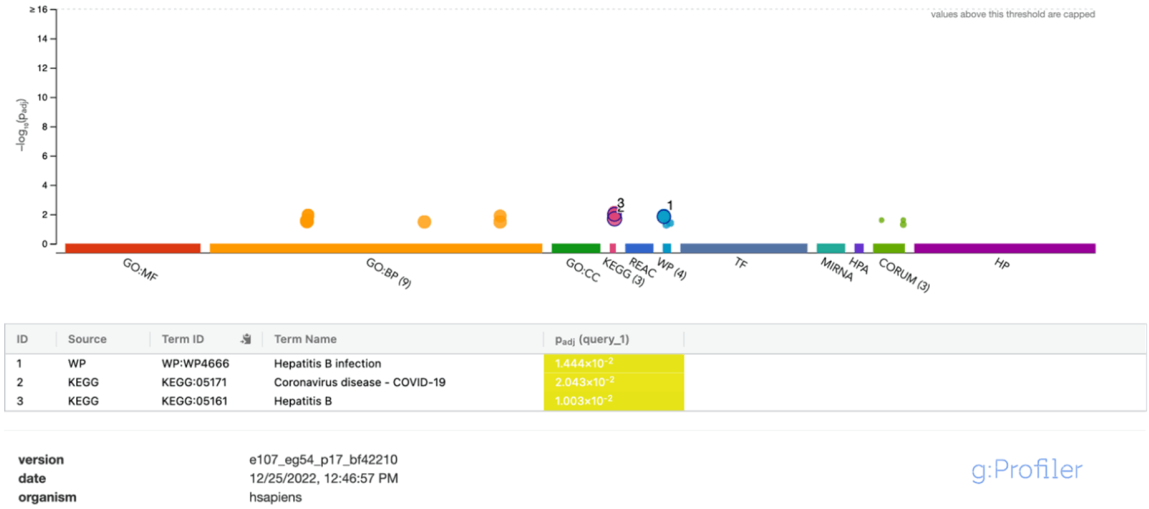

Figure 6. Functional annotation from g:Profiler for *IFIH1* and *STAT3*. The results from g:Profiler [17] show - and support the fact - that both genes are involved in Hepatitis B and the Coronavirus disease.

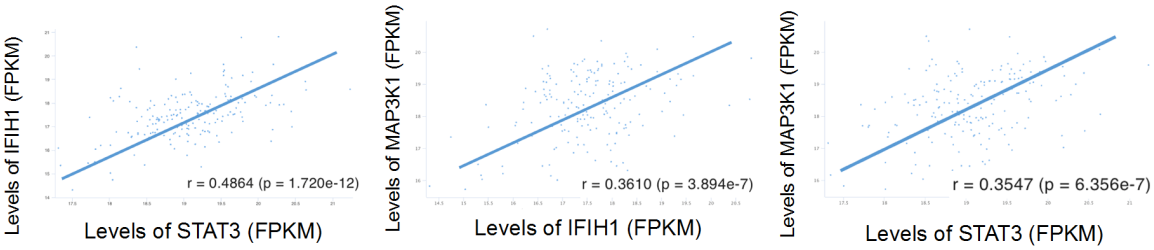

Figure 7. Correlation analysis results for *STAT3* and *IFIH1* in GDC TARGET-AML cohorts. The expression of *STAT3* versus *IFIH1* (Left), *MAP3K1* versus IFIH1 (middle), *STAT3* versus *MAP3K1* (right) in Acute Myeloid Leukemia are positively correlated, indicating all three genes are co-expressed in the development of AML.

tance of the three candidate genes (*IFIH1, STAT3, MAP3K1*). The analysis revealed infection and immune pathways as among the top hits (**Table 1**). Additionally, the Hepatitis B pathway had the highest hit among the others, thus it is a disease also potentially associated with *IFIH1, STAT3*, and *MAP3K1*.

Functional enrichment analysis confirms that IFIH1 and STAT3 are involved in the Hepatitis B and Coronavirus disease pathways

Identifying functionally important genes and associated terms are useful and informative. The Hepatitis B pathway was identified as the top hit in the above pathway analysis (with the highest significance) that the three candidate genes were involved in. It is clear that the two candidate genes (*IFIH1* and *STAT3*) are indeed involved in COVID-19 and Hepatitis B (**Figure 6**). These results support the DAVID [15]/KEGG [16] pathway analysis results (**Table 1**).

Correlation of the expressions of IFIH1, STAT3, and MAP3K1

After inputting all three genes into the UCSC Xena database and retrieving their expression information from the GDC TARGET-AML dataset, we found that they are all positively correlated with each other (**Figure 7**). This indicated that these genes are likely all co-expressed in leukemia samples.

Additional genes connected to the candidate genes that are potentially linked to AML

The genes closer to the center of the constellation graphs have a more significant correlation with the target gene. This revealed that the genes *ALPK1, OASL2, DYNC1LI1*, and *SAMD4B*

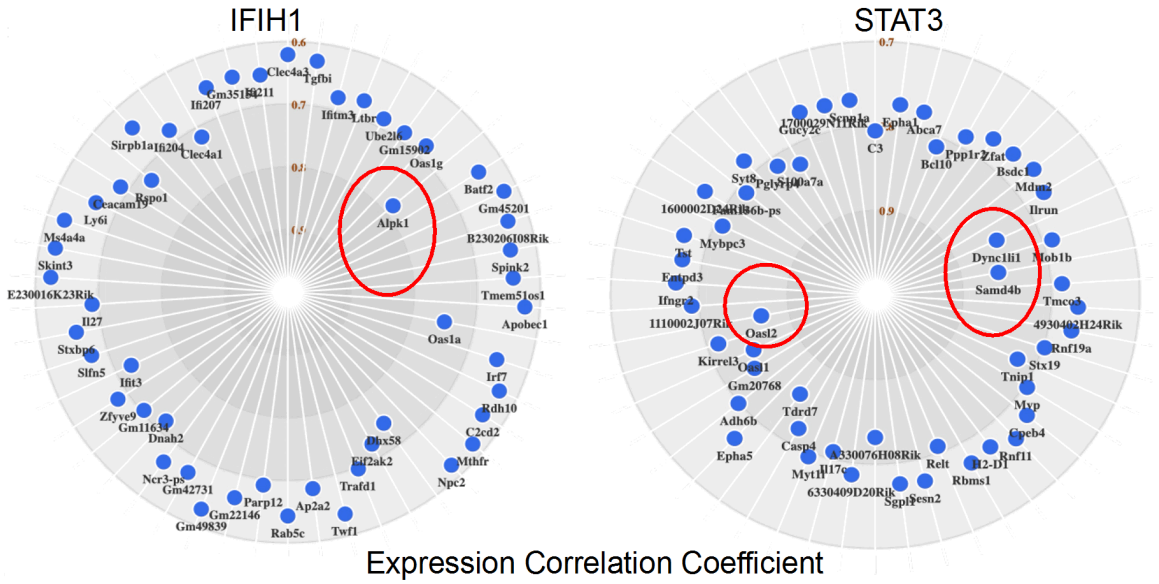

Figure 8. Mouse orthologous gene constellation graph for IFIH1 (left) and STAT3 (right). We found that *ALPK1* has a higher correlation alongside *IFIH1* and *DYNC1LI1, OASL2*, and *SAMD4B* have a higher correlation alongside *STAT3*. These additional genes could serve as possible drug targets for combating AML and COVID-19.

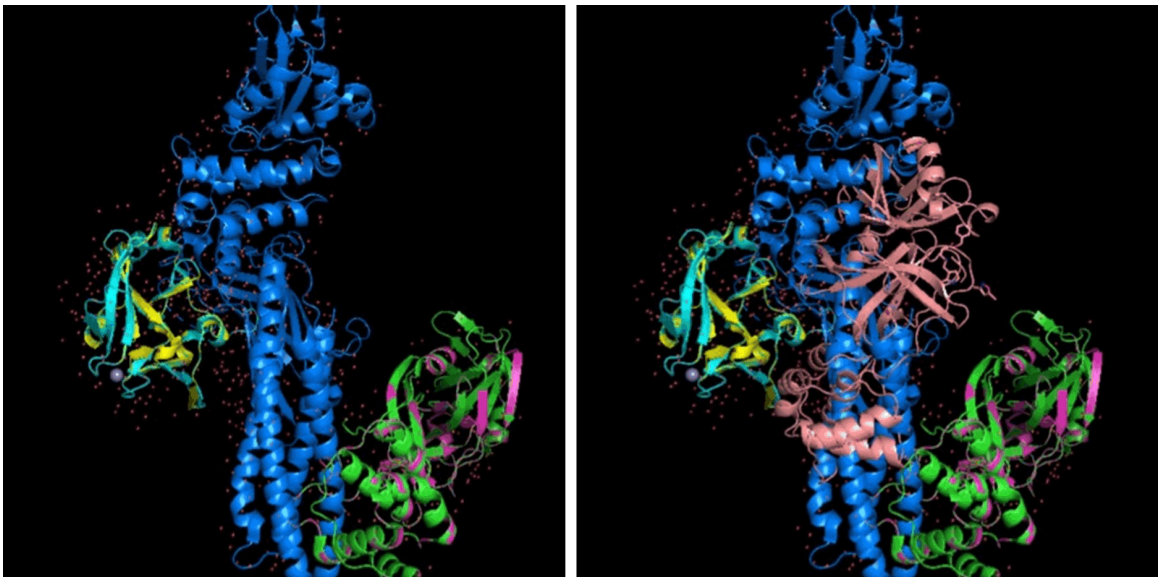

Figure 9. Molecular visualization of the structure of the 6LU7/3GA3/6NJS protein.

potentially interact with *IFIH1* and *STAT3* (**Figure 8**).

Interaction between the three candidate genes and the COVID-19 protein (SARS-CoV-2)

Revealing atomic-level accuracy of protein-protein interactions responsible for key biological processes is important. In order to test if the protein molecule MDA5 (encoded by *IFIH1*) interacts with SARS-CoV-2, we obtained both of their crystal structures from the Protein Data Bank to perform docking prediction of their protein complexes. As shown (Structure in pink) in **Figure 9**, SARS-CoV-2 did not directly interact with MDA5.

To understand this, SARS-CoV-2 does not directly interact with *IFIH1*; in reality, they interact through *STAT3*, which directly interacts with

Nsp1, which is a nonstructural protein component in COVID-19 [20]. Molecular docking predictions based on the ZDOCK server prediction tool [18, 19] of MDA5 and the main protease of COVID-19 also show the interaction complex between the two proteins (Visualization graphs were generated using the PyMOL software [21]).

Discussion

Our study identified three genes involved in COVID-19, AML, and HBV. *IFIH1* has been reported to play an important role in children's immunity against COVID-19 compared to adults [22, 23]. Both COVID-19 and leukemia diseases have been reported in literature to be associated with each other [24, 25]. Using *IFIH1* as the hub gene, we identified two miRNAs that regulate *IFIH1*. Then, through analysis of the targeted genes of these two miRNAs, we identified two genes (*STAT3* and *MAP3K1*) that have a direct interaction and are co-expressed with *IFIH1*. Moreover, our results indicated that all three genes could contribute to the comorbidity of COVID-19, AML, and HBV. Through extending the analysis for these three candidate genes, more genes could be further studied to evaluate their candidacy serving as biomarkers for comorbidity of complex diseases.

STAT3 has been demonstrated to be active in the upstream regulator analysis on Differential Expression of Genes (DEGs) linked with survival in End of Induction (EOI) residual blast cells obtained from AML bone marrow (BM) samples in a relatively recent scRNA-seq study discussing pediatric AML [26]. *STAT*s regulate the expressions of genes involved in immunological responses, inflammation, apoptosis, and oncogenesis as well as cell proliferation and differentiation [27], and can also regulate or be regulated by miRNAs through binding to its upstream enhancer regions [28]. Our results suggest that *STAT3* could be an important biomarker of COVID-19 and AML in children.

Interestingly, pathway analysis further revealed that there was also a connection between COVID-19 and Hepatitis B. HBV has been linked to chronic leukemia, particularly acute myeloid leukemia, according to evaluation of current bone marrow sample data in a published study [29]. Children with acute lymphoblastic leukemia (ALL) carry a high risk of hepatitis B virus (HBV) infection. Therefore, there could be shared pathways and gene regulation mechanisms between COVID-19 and HBV infections.

The miRNAs have been shown to play important roles in physiological regulation of blood cells and in leukemia development. Based on the Diana Tools - MiTED result [30], *hsa-miR-196b* belongs to one of the top 2% highly expressed miRNAs in Bone Marrow tissue and AML. *hsa-miR-196a-5p* is complementary to sites in the 3'UTR of *MAP3K1*, which exhibits upregulated expression at mRNA and protein levels [31]. By employing mirPath-v3 [32], we found *hsa-miR-196a* and its targeted genes, *STAT3*, *MAP3K1*, and *IFIH1*. Together, our data suggests that *hsa-miR-196b* and *hsa-miR-196a-5p* and their target genes *(STAT3*, *MAP3K1*, and *IFIH1)* can potentially be important biomarkers of COVID-19, AML and HBV.

MiRNA-196a has not only been reported to have elevated levels in AML: recent evidence related to estrogen receptor (ER)-positive breast tumors has shown up-regulation of the miRNA compared to ER-negative tumor tissues. Cervical cancer was also seen to be linked to the up-regulation of *miRNA-196a*, and cervical cancer patients with high levels of the miRNA are reported to have a decreased chance of survival. This miRNA has also been discovered to play a role in non-malignant disorders, such as chronic kidney disease, diabetic nephropathy, and more [33]. The trend of *miRNA-196a* up-regulation increasing disease severity can be applied to COVID-19, AML, and HBV. Additionally, SARS-CoV-2-induced dysregulation of miRNA targets may play a role in the dysregulation of miRNA targets in COVID-19 [34]. Thus, through dysregulating miRNAs such as the *miRNA-196a* family, SARS-CoV-2 likely hinders the immune system in defending against viruses such as HPV and increases the chance of AML development and progression.

One limitation of the miRNA database (TarBase) is that only experimentally targeted miRNAs were reported, so the study may not have fully captured all possible miRNAs. This may have limited our targeted gene set. Using the TCGA and ImmGen databases, we confirmed that the cell types that have a low expression of *hsa-miR-196a-5p*, such as monocyte cells, express high levels of their target genes, *IFIH1* and *STAT3*. Therefore, both genes are important in

the context of miRNA-mediated post-transcriptional regulation.

Future work will be focused on developing a Machine Learning (ML) model to investigate the importance of multiple factors (e.g. age, race, etc.) and conduct comorbidity network analysis for the candidate genes/miRNAs. Our next plan is to characterize the genes' interaction patterns contributing to comorbidity and elucidate cancer-causing mechanisms of candidate genes.

Acknowledgements

The TCGA LAML RNA-Seq gene expression results in this study here are in whole or part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga. TARGET's Study of Acute Myeloid Leukemia was originally published by the National Cancer Institute.

Disclosure of conflict of interest

None.

Address correspondence to: Judy Bai, Greenhills School, Ann Arbor, MI 48105, USA. E-mail: judycoolgymnast@gmail.com

References

[1] Wang Q, Berger NA and Xu R. Analyses of risk, racial disparity, and outcomes among US patients with cancer and COVID-19 infection. JAMA Oncol 2021; 7: 220-227.

[2] Nekooghadam SM, Moradi A, Karimi Toudeshki K and Pishgahi M. A case of acute leukemia following remission of COVID-19 infection; an urge to search for a probable association. Arch Acad Emerg Med 2021; 9: e51.

[3] Mandal S, Meenakshisundaram R and Karki A. Does COVID-19 increase your risk of developing hairy cell leukemia? Chest 2022; 162: A641.

[4] Blood cancer and COVID-19 - Leukaemia Foundation. Leukaemia Foundation. Published September 17, 2024. Accessed November 21, 2023. https://www.leukaemia.org.au/blood-cancer/covid-19/.

[5] IFIH1 protein expression summary - The Human Protein Atlas. Proteinatlas.org. Published 2014. Accessed November 22, 2023. https://www.proteinatlas.org/ENSG00000115267-IFIH1.

[6] McNeer NA, Philip J, Geiger H, Ries RE, Lavallée VP, Walsh M, Shah M, Arora K, Emde AK, Robine N, Alonzo TA, Kolb EA, Gamis AS, Smith M, Gerhard DS, Guidry-Auvil J, Meshinchi S and Kentsis A. Genetic mechanisms of primary chemotherapy resistance in pediatric acute myeloid leukemia. Leukemia 2019; 33: 1934-1943.

[7] Croce CM. Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet 2009; 10: 704-714.

[8] microRNA - What it is and How it Works|OSUCCC - James. The James - OSUCCC. https://cancer.osu.edu/microrna.

[9] Jin S, Zeng X, Fang J, Lin J, Chan SY, Erzurum SC and Cheng F. A network-based approach to uncover microRNA-mediated disease comorbidities and potential pathobiological implications. NPJ Syst Biol Appl 2019; 5: 41.

[10] Vlachos IS, Paraskevopoulou MD, Karagkouni D, Georgakilas G, Vergoulis T, Kanellos I, Anastasopoulos IL, Maniou S, Karathanou K, Kalfakakou D, Fevgas A, Dalamagas T and Hatzigeorgiou AG. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. Nucleic Acids Res 2015; 43: D153-9.

[11] Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, Bork P, Jensen LJ and von Mering C. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res 2023; 51: D638-D646.

[12] Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, Zhu J and Haussler D. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol 2020; 38: 675-678.

[13] TARGET's Study of Acute Myeloid Leukemia. Cancer.gov. Published 2024. https://www.cancer.gov/ccg/research/genome-sequencing/target/studied-cancers/acute-myeloid-leukemia.

[14] Heng TS and Painter MW; Immunological Genome Project Consortium. The Immunological Genome Project: networks of gene expression in immune cells. Nat Immunol 2008; 9: 1091-1094.

[15] Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, Stephens R, Baseler MW, Lane HC and Lempicki RA. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 2007; 8: R183.

[16] Kanehisa M and Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27-30.

[17] Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J and Peterson H. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res 2023; 51: W207-W212.

[18] Pierce BG, Hourai Y and Weng Z. Accelerating protein docking in ZDOCK using an advanced 3D convolution library. PLoS One 2011; 6: e24657.

[19] Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T and Weng Z. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. Bioinformatics 2014; 30: 1771-1773.

[20] Gordon DE, Hiatt J, Bouhaddou M, Rezelj VV, Ulferts S, Braberg H, Jureka AS, Obernier K, Guo JZ, Batra J, Kaake RM, Weckstein AR, Owens TW, Gupta M, Pourmal S, Titus EW, Cakir M, Soucheray M, McGregor M, Cakir Z, Jang G, O'Meara MJ, Tummino TA, Zhang Z, Foussard H, Rojc A, Zhou Y, Kuchenov D, Hüttenhain R, Xu J, Eckhardt M, Swaney DL, Fabius JM, Ummadi M, Tutuncuoglu B, Rathore U, Modak M, Haas P, Haas KM, Naing ZZC, Pulido EH, Shi Y, Barrio-Hernandez I, Memon D, Petsalaki E, Dunham A, Marrero MC, Burke D, Koh C, Vallet T, Silvas JA, Azumaya CM, Billesbølle C, Brilot AF, Campbell MG, Diallo A, Dickinson MS, Diwanji D, Herrera N, Hoppe N, Kratochvil HT, Liu Y, Merz GE, Moritz M, Nguyen HC, Nowotny C, Puchades C, Rizo AN, Schulze-Gahmen U, Smith AM, Sun M, Young ID, Zhao J, Asarnow D, Biel J, Bowen A, Braxton JR, Chen J, Chio CM, Chio US, Deshpande I, Doan L, Faust B, Flores S, Jin M, Kim K, Lam VL, Li F, Li J, Li YL, Li Y, Liu X, Lo M, Lopez KE, Melo AA, Moss FR 3rd, Nguyen P, Paulino J, Pawar KI, Peters JK, Pospiech TH Jr, Safari M, Sangwan S, Schaefer K, Thomas PV, Thwin AC, Trenker R, Tse E, Tsui TKM, Wang F, Whitis N, Yu Z, Zhang K, Zhang Y, Zhou F and Saltzberg D; QCRG Structural Biology Consortium; Hodder AJ, Shun-Shion AS, Williams DM, White KM, Rosales R, Kehrer T, Miorin L, Moreno E, Patel AH, Rihn S, Khalid MM, Vallejo-Gracia A, Fozouni P, Simoneau CR, Roth TL, Wu D, Karim MA, Ghoussaini M, Dunham I, Berardi F, Weigang S, Chazal M, Park J, Logue J, McGrath M, Weston S, Haupt R, Hastie CJ, Elliott M, Brown F, Burness KA, Reid E, Dorward M, Johnson C, Wilkinson SG, Geyer A, Giesel DM, Baillie C, Raggett S, Leech H, Toth R, Goodman N, Keough KC and Lind AL; Zoonomia Consortium; Klesh RJ, Hemphill KR, Carlson-Stevermer J, Oki J, Holden K, Maures T, Pollard KS, Sali A, Agard DA, Cheng Y, Fraser JS, Frost A, Jura N, Kortemme T, Manglik A, Southworth DR, Stroud RM, Alessi DR, Davies P, Frieman MB, Ideker T, Abate C, Jouvenet N, Kochs G, Shoichet B, Ott M, Palmarini M, Shokat KM, García-Sastre A, Rassen JA, Grosse R, Rosenberg OS, Verba KA, Basler CF, Vignuzzi M, Peden AA, Beltrao P and Krogan NJ. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. Science 2020; 370: eabe9403.

[21] The PyMOL Molecular Graphics System version 2.0 (Schrödinger, LLC, 2015).

[22] Mallapaty S. Kids and COVID: why young immune systems are still on top. Nature 2021; 597: 166-168.

[23] Asgari S, Schlapbach LJ, Anchisi S, Hammer C, Bartha I, Junier T, Mottet-Osman G, Posfay-Barbe KM, Longchamp D, Stocker M, Cordey S, Kaiser L, Riedel T, Kenna T, Long D, Schibler A, Telenti A, Tapparel C, McLaren PJ, Garcin D and Fellay J. Severe viral respiratory infections in children with IFIH1 loss-of-function mutations. Proc Natl Acad Sci U S A 2017; 114: 8342-8347.

[24] Li W, Wang D, Guo J, Yuan G, Yang Z, Gale RP, You Y, Chen Z, Chen S, Wan C, Zhu X, Chang W, Sheng L, Cheng H, Zhang Y, Li Q, Qin J; Hubei Anti-Cancer Association; Meng L and Jiang Q. COVID-19 in persons with chronic myeloid leukaemia. Leukemia 2020; 34: 1799-1804.

[25] Martín-Moro F, Núnez-Torrón C, Pérez-Lamas L, Jiménez-Chillón C, Marquet-Palomanes J, López-Jiménez FJ and Herrera-Puente P. The impact of lockdown during the COVID-19 pandemic on newly acute myeloid leukemia patients: Single-centre comparative study between 2019 and 2020 cohorts in Madrid. Leuk Res 2021; 101: 106518.

[26] Mumme H, Thomas BE, Bhasin SS, Krishnan U, Dwivedi B, Perumalla P, Sarkar D, Ulukaya GB, Sabnis HS, Park SI, DeRyckere D, Raikar SS, Pauly M, Summers RJ, Castellino SM, Wechsler DS, Porter CC, Graham DK and Bhasin M. Single-cell analysis reveals altered tumor microenvironments of relapse- and remission-associated pediatric acute myeloid leukemia. Nat Commun 2023; 14: 6209.

[27] Gharibi T, Babaloo Z, Hosseini A, Abdollahpour-Alitappeh M, Hashemi V, Marofi F, Nejati K and Baradaran B. Targeting STAT3 in cancer and autoimmune diseases. Eur J Pharmacol 2020; 878: 173107.

[28] Zou S, Tong Q, Liu B, Huang W, Tian Y and Fu X. Targeting STAT3 in cancer immunotherapy. Mol Cancer 2020; 19: 145.

[29] Mansouri N, Movafagh A, Sayad A, Ghafouri-Fard S, Darvish H, Zare-Abdollahi D, Emamalizadeh B, Shahvaisizadeh F, Ghaedi H, Bastami M, Kayyal M, Hashemi M, Heidari M, Nejatizadeh A and Zamani M. Hepatitis B virus infection in patients with blood disorders: a concise review in pediatric study. Iran J Ped Hematol Oncol 2014; 4: 178-87.

[30] Kavakiotis I, Alexiou A, Tastsoglou S, Vlachos IS and Hatzigeorgiou AG. DIANA-miTED: a microRNA tissue expression database. Nucleic Acids Res 2022; 50: D1055-D1061.
[31] Shiohama T, Fujii K, Miyashita T, Takatani T, Ikehara H, Uchikawa H, Motojima T, Uchida T and Shimojo N. MicroRNAs profiling in fibroblasts derived from patients with Gorlin syndrome. J Hum Genet 2019; 64: 757-765.
[32] Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, Dalamagas T and Hatzigeorgiou AG. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. Nucleic Acids Res 2015; 43: W460-6.
[33] Khalilian S, Hosseini Imani SZ, Abedinlou H, Omrani MA and Ghafouri-Fard S. miR-196a in the carcinogenesis and other disorders with an especial focus on its biomarker capacity. Pathol Res Pract 2024; 260: 155433.
[34] Roustai Geraylow K, Hemmati R, Kadkhoda S and Ghafouri-Fard S. miRNA expression in COVID-19. Gene Rep 2022; 28: 101641.