Analysis of Cellular and Biochemical Factors Involved in Acute Myeloid Leukemia Taking Into Account the Dysfunctional Module

JIAQI CHEN¹, JIANJUN TANG², CHAOKE BU³, JIANYUN LIAO³, JIEMING YU¹, TINGTING LUO¹, YIQI XV¹, RONGFANG DING¹*

¹ Department of Pediatrics, Nanfang Hospital, Southern Medical University, Guangzhou, 510000, China
² Department of Anesthesiology, Nanfang Hospital, Southern Medical University, Guangzhou, 510000, China
³ Nanfang-Chunfu Children’s Institute of Hematology & Oncology, Taixin Hospital, Dongguan, 523000, China

Abstract: Acute myeloid leukemia (AML) is a malignant hematological disease caused by the abnormal hematopoietic system. Its specific pathogenesis is still unclear. Therefore, we have carried out a series of related studies. Firstly, co-expression analysis of acute myeloid leukemia related genes in the Human Mendelian Genetic Database (OMIM) was carried out, and a co-expression module was constructed. Further enrichment analysis of module-related genes with Go and KEGG was carried out. Subsequently, the modules are analyzed by Crosstalk to show their interaction. Also, pivot prediction is used to identify critical factors in the regulatory module and potential target drugs. Results A total of 14 co-expression modules were obtained, which were found to be significantly involved in cell proliferation and apoptosis, Ras and Wnt signaling pathways. The modules with higher connectivity (UBB, RAC1, TP53, PIK3CA) were also counted. Also, it is found that there is significant crosstalk between these modules. On the other hand, we screened 2815 drivers of regulatory modules, including TP53, E2F1, MYC-based transcription factors, and non-coding RNA (ncRNA) dominated by microRNA-320a, microRNA-193b-3p, and microRNA-93-5p. Finally, many regulatory drugs (Copper, Artenimol, Indomethacin, etc.) with targeted pathogenic genes were predicted. Overall, this study explores the underlying pathogenesis of acute myeloid leukemia and predicts the related regulatory drugs. The results of this study provide a new method and theoretical basis for follow-up mechanism research and treatment strategy.

Keywords: dysfunction module, acute myeloid leukemia, KEGG

1. Introduction

Acute myeloid leukemia (AML) is a hereditary disease, which is regulated by hereditary pathogenesis [1]. On the other hand, the environment is also a critical factor in the pathogenesis of acute myeloid leukemia [2]. Therefore, it is urgent to study the therapeutic mechanism of acute myeloid leukemia. Fortunately, many biologists and medical researchers have invested in the pathogenesis, physiological process, and treatment of acute myeloid leukemia, and achieved excellent results. For example, one report has shown that haploid hematopoietic cell transplantation plays an essential role in the treatment of acute myeloid leukemia [3].

In many cases of this kind of AML, there is a resistance against chemotherapy caused mainly by factors that are involved in the oxidative stress process only for the production of reactive oxygen species (ROS) and also the existence of the hypodiploid cells. Among available antioxidants, vitamins A, C, and E, glutathione (GSH), and the enzymes superoxide dismutase (SOD), catalase (CAT), glutathione peroxidase (GPx), and glutathione reductase (GRx) are the most important [4].

The acute myeloid leukemia has been proved to inhibit DNMT3A, IDH 1/2, and TET 2 by targeted molecular therapy [4]. Previous studies have shown that ubiquitin-binding enzyme E2E1 (UBE2E1) inhibits the growth of acute myeloid leukemia by microarray analysis, thus becoming a new prognostic factor [5]. Therefore, the useful prognostic tools such as cytogenetics and mutation detection for acute myeloid leukemia provide a new direction for medical treatment [6].

*email: 34733760@qq.com
Interestingly, in the study of acute myeloid leukemia, gene mutations were also closely related to prognosis [7]. More and more studies have shown that the low expression of NKD-1 in cancer cells has significantly reduced the incidence of acute myeloid leukemia and predicted the adverse prognosis of cytogenetics [8]. However, the low expression of microRNA-215 increases the risk of acute myeloid leukemia and is closely related to poor clinical prognosis [9]. It is worth noting that exon 12 of NPM 1 gene was amplified by next-generation sequencing to detect and remove minimal residues in acute myeloid leukemia [10]. Also, Noguera NI et al. found that high doses of ascorbic acid (Asc) and arsenic trioxide (ATO) combined to induce apoptosis, thereby killing the pathogenic cells of acute myeloid leukemia and acute promyelocytic leukemia [11].

It has been confirmed that the sensitivity of serum endostatin and fibroblast growth factor 19 (FGF-19) is predicted by ROC curve analysis, thus affecting the occurrence of acute myeloid leukemia [12]. Studies have shown that the treatment of acute lymphoblastic leukemia (ALL) by bispecific T cell antigen-antibody isocitrate dehydrogenase (IDH) has made significant achievements in clinical practice [13]. Xu et al. confirmed that Ara-C could increase the frequency of rs2293157T/T through primary chemotherapy, thus regulating the STAT 5 gene polymorphism and acute myeloid leukemia [14]. On the other hand, IGK light chain is expressed in acute myeloid leukemia cells so that myeloid cell-derived IGK may play a key role in leukocytes [15]. Besides, abnormal activity of nuclear factor kappa B (NF-kappa B) promotes proliferation and inhibits apoptosis of acute myeloid leukemia cells, which indicates that nuclear factor kappa B signaling pathway plays an essential role in acute myeloid leukemia [16]. It has been confirmed that the analysis of gene expression profile of acute myeloid leukemia by gene chip technology shows that the low expression of SMC1A protein in prognosis, thus predicting the occurrence of acute myeloid leukemia [17].

Generally speaking, these critical results provide valuable guidance for the research work and much inspire our thinking. Therefore, in order to better address the pathogenesis of acute myeloid leukemia mediated by a transcription factor (TF) and ncRNA, we conducted multiple data integration. It not only provides new insights into the underlying pathogenesis of acute myeloid leukemia but also provides a new way of thinking for the study of other diseases.

2. Materials and methods

2.1. Data resources

First of all, we collected many published results on acute myeloid leukemia from the NCBI-Gene database [18], and the Human Online Mendelian Heredity (OMIM) database [19]. In order to systematically investigate the molecular mechanism of AML-related genes, 851 and 287 psoriasis-related genes were obtained from Gene and OMIM databases, respectively. The protein interaction (String) database [20] was used to search for the protein with a score of more than 900. A total of 3 900 potential pathogenic genes of acute myeloid leukemia were obtained. Based on the expression microarray matrix of acute myelogenous leukemia (AML) samples, we constructed the expression profiles of 3 900 potential pathogenic genes of AML.

2.2. Coexpression Module Based on Cohesion Recognition

To explore the co-expression of these potential pathogenic genes in acute myeloid leukemia, we used weighted gene co-expression network analysis (WGCNA) [21] to analyze the gene expression profile matrix of the related genes in acute myeloid leukemia and to find the co-expression gene module. Unlike other clustering methods, WGCNA clustering criteria have biological significance. It takes the n power of the correlation coefficient of gene expression so that the distribution of the correlation coefficient values gradually conforms to the scale-free distribution. Then, a hierarchical clustering tree is constructed based on cohesion and correlation coefficient between genes. Genes with similar patterns are classified as the same branch, while different branches of the clustering tree represent different gene modules. Therefore, the results obtained by this method have higher reliability. In this study, the branches with the most significant number of genes (2 590 genes) may imply weak cohesion, and their
synergistic expression is not apparent. Therefore, in the next study, it was removed, and the remaining 14 branches were identified as functional modules.

2.3. Functional and Pathway Enrichment Analysis and Identification of Dysfunctional Modules
Exploring the function and signal pathway of genes is often an effective means to study the molecular mechanism of diseases. The function and pathway of module gene participation can characterize the dysfunction mechanism of the module in the process of disease occurrence. Therefore, we use R language Cluster Profiler package [22] for enrichment analysis of Go function (p-value cutoff = 0.05, q-value cutoff = 0.05) and KEGG pathway (p-value cutoff = 0.05, q-value cutoff = 0.05). According to the function and pathway of module gene participation, it was identified as a module of acute myeloid leukemia related dysfunction. Also, we use Cytoscape’s CLUGO application [23] to analyze the function of the integrated module network.

2.4. Exploring the Interaction of Modular Genes and Identifying Introductory Genes
Potential pathogenic genes of acute myeloid leukemia (AML) are clustered into functional modules based on co-expression relationship. However, besides co-expression relationships, there are also target interactions among genes. Therefore, it is helpful to understand the core molecules of the driving module function and dysfunction by observing the target interaction of genes in the module. Based on the String database, we constructed a protein interaction subnet for each module. Cytoscape is then used for presentation and network analysis (including connectivity calculations). Finally, screening the genes with the highest connectivity is considered as the core molecule to regulate the progress of the module and identified as the intrinsic gene. These endogenous genes may represent pathogenic molecules of acute myeloid leukemia.

2.5. Regulatory factors and potential drugs for predicting dysfunction module
ncRNA and TF are the potential drivers of gene transcription and post-transcriptional regulation, which often mediate the occurrence of diseases. In order to explore the driving forces of the co-expression module of genes related to acute myeloid leukemia, pivot analysis was performed using the score >= 0.5 ncRNA-RNA (protein) interaction pair in RAID v2.0 database [24] and all human transcription factor target data in TRRUST V2 database [25]. Pivot analysis refers to finding at least two interacting drivers with the module in target pairs and calculating the significance of the interaction between the driver and the module based on the hypergeometric test. Screening ncRNA or TF with P-value < 0.01 is a pivot of the significant regulatory module. Finally, the pivot is identified as the core pivot by statistical analysis.

Similarly, the prediction of potential drugs is also based on pivot analysis. The pivot analysis was performed using the Drugbank database [26] drug-target gene data as the background set. The pivot, which significantly regulates the dysfunction module, was considered as a potential drug for the treatment of acute myeloid leukemia. Pivot analysis results count Copper, Indomethacin, and other drugs. They can regulate more modules and have more obvious therapeutic or toxic effects. On the other hand, searching for drug target genes in drug regulatory modules is a potential target for the treatment of the disease.

3. Results and discussions
3.1. Regulation of genes associated with acute myeloid leukemia and co-expression in dysfunction module
Biologists have carried out a series of studies on the process of acute myeloid leukemia and identified many pathogenic genes of acute myeloid leukemia-related diseases. However, the regulatory and synergistic relationship between these genes is not clear. Therefore, in order to systematically study the mechanism of the genes associated with acute myeloid leukemia, we have integrated a large number of data. First, we collected 985 genes related to acute myeloid leukemia from multiple databases. Secondly,
Based on the human protein interaction data, we seek the interacting genes of these related genes. Combining the two gene sets, we obtained 3900 genes, potentially regulating acute myeloid leukemia. Also, based on weighted gene co-expression network analysis (WGCNA), we observed that these genes showed significant grouping co-expression in 151 samples of acute myeloid leukemia. Finally, we obtained 14 sets of co-expressed gene sets and identified them as modules. We obtained 14 modules co-expressed module, and mainly observed the clustering phenomenon of module gene expression behavior in patient samples (Figures 1A & 1B). These modules may participate in different functional pathways, representing different regulatory mechanisms to mediate the occurrence and development of acute myeloid leukemia.

![Cluster Dendrogram](image1.png)

**Figure 1.** Clustering of disordered genes into co-expression modules (A & B)

3.2. Multiple factors of dysfunction module may characterize pathogenic molecules of acute myeloid leukemia

Further study on the biological characteristics of the dysfunction module shows that the dysfunction module of acute myeloid leukemia is composed of many potential genomes of acute myeloid leukemia and clustered into modules according to co-expression relationship. Seventy endogenous genes, including UBB, RAC1, TP53, and PIK3CA, were screened from 14 dysfunction modules, and their connectivity was calculated (Figure 2A). On the other hand, besides the significant regulation of intrinsic genes on modules, there are also key drivers such as ncRNA and transcription factors. Therefore, based on the hypergeometric test, we performed a pivot analysis of ncRNA and obtained 1841 ncRNA-Module interaction pairs (Figure 2B). These ncRNAs are involved in various physiological processes, such as cell proliferation and cycle progression, apoptosis, invasion, angiogenesis, and tumorigenesis. Also, statistical analysis of the results showed that small endogenous RNA (microRNA) microRNA-320a drives six modules of acute myeloid leukemia. The imbalance of transcription factor expression may induce an imbalance of the immune system, which directly leads to the activation of inflammation in acute myeloid leukemia. Therefore, according to the regulation relationship of transcription factors to genes, we carried out pivot analysis and prediction on the module and obtained 156 TF-Module regulatory pairs (Figure 2C). Statistical analysis of these transcription factors revealed that TP53, a tumor suppressor gene, significantly regulated four dysfunction modules. It has a potential role in maintaining the homeostasis of the immune system and regulating acute myeloid leukemia. Studies have shown that TP53 mutation plays a crucial role in complex karyotype acute myeloid leukemia [27]. These important
transcription factors have been identified as the core transcription factors in the regulation of hematological tumorigenesis.

![Figure 2](image)

**Figure 2.** The pivot that regulates modular genes (Excerpt); A. Degree in co-expression module network. The larger the node, the higher the degree. B. ncRNA that regulates modular genes. Yellow nodes represent ncRNA, and green nodes represent modules. C. TF that regulates modular genes. Green nodes represent transcription factors, and yellow nodes represent modules.

3.3. Functional and Pathway Identification of Acute Myeloid Leukemia Modules Participated by Modules

Function and pathway are essential means to reveal pathogenesis. Therefore, in order to study the possible dysfunction caused by gene disorder, we performed GO function and KEGG pathway enrichment analysis on 14 interaction modules. We collected a wealth of GO terms and obtained 3873 cell composition items, 5850 molecular functional terms, 36188 biological processes, and 1746 functional pathways. On the one hand, based on functional analysis, we found that these functions mainly focused on DNA damage response, cell death, apoptotic regulation, and others (Figure 3A). On the other hand, the enrichment analysis of the KEGG pathway reflects that cancer-related genes are mainly involved in p53, NF-kappa B, and Wnt signaling pathways (Figure 3B). Besides, statistical analysis of functions and pathways revealed that 9 of these modules significantly enriched and regulated the biological processes related to metastasis, such as DNA metabolism.

![Figure 3](image)

**Figure 3.** Enrichment results of modular genes; A. Modular genes are involved in GO function. B. Modular genes are involved in KEGG signaling pathway.
3.4. Potential drugs for acute myeloid leukemia

The ultimate goal of pathogenesis research is to explore effective treatment methods and potential treatment mechanisms. Exploring potential drugs and drug targets based on in-depth and comprehensive pathogenesis is a scientific and systematic way to achieve this ultimate goal. Therefore, we use this factor-mediated dysfunction module to predict potential drugs. According to the drug-target relationship between drugs and their target genes, we performed pivot analysis on the module and set P-value < 0.05. A total of 778 Drug-Module regulatory pairs involving 843 drugs were obtained (Figure 4). These drugs have significant pharmacological effects on dysfunction modules and are potential drugs for the treatment of acute myeloid leukemia. Statistical results show that Copper can play a regulatory role by targeting six dysfunction modules. These drugs have laid a theoretical foundation for further exploration of drug R&D personnel.

Figure 4. Drugs with potential therapeutic effects on modular genes (Excerpt);
Yellow nodes represent drugs, and green nodes represent modules.

Acute myeloid leukemia (AML) is caused by many complications and pathogenic factors due to the complicated relationship between environment and heredity. As a result, the progress of medical research is plodding. However, its social pressure and economic losses are also increasing day by day, so it is urgent to explore acute myeloid leukemia. Therefore, in order to fully explore the pathogenesis of acute myeloid leukemia mediated by transcription factors and ncRNA, we first integrated these potential pathogenic genes and interacting genes and observed their expression in disease samples. Thus, we obtained 14 expression modules and participated in functions and pathways. At the same time, we analyzed the functions and pathways and found that up to 9 modules significantly enriched and regulated the biological processes related to cancer metastasis, such as DNA metabolism. The seven modules can significantly regulate the biological processes related to tumor metastasis, such as mitotic cell cycle checkpoints. Previous studies have shown that CHFR gene inactivation induced by mitosis has become a critical factor in the treatment of acute myeloid leukemia and a molecular marker for predicting sensitivity [18,16029562]. These functions and pathways involved by module genes have an overall network effect, which comprehensively regulates the pathogenesis of acute myeloid leukemia.

Also, we identified the factors that regulate these dysfunction modules, including RAC1, TP53, PIK3CA, CDC42, and other endogenous genes, ncRNA mainly composed of microRNA-320a, microRNA-193b-3p, microRNA-93-5p, and transcription factors mainly composed of TP53, JUN, E2F1, MYC, MYCN, NFKB1, RUNX1. Introgenous genes are the most active genes interacting with other genes in each dysfunction module. These genes often have the characteristics of pull-in and move the whole body, and have a significant regulatory effect on the module. RNA sequencing analysis showed that significant fusion of NUP214-RAC1 and RAC1-COL12A1 genes was associated with poor prognosis of acute myeloid leukemia [28]. Welch JS et al. showed that the combination of TP53 and decitabine could regulate acute myeloid leukemia[20]. More and more studies have shown that the
A combination of PI3K P110 alpha subtype (PIK3CA) and MEK may be used to treat acute myeloid leukemia [29, 30]. Other endogenous genes UBB, CDC42, and POLR2A are also involved in the treatment of myeloid leukemia to varying degrees.

On the other hand, microRNA-320a is involved in the apoptosis of HL-60 cells induced by SNS-032 [31]. Although there is no direct literature support, mir-193b-3p and mir-93-5p have potential regulatory effects on acute myeloid leukemia. Also, studies have shown that TP53 transcription factor mutations play a crucial role in complex karyotype acute myeloid leukemia [32]. Interestingly, the MEK signaling pathway promotes the unfolded protein response in acute myeloid leukemia by activating the transcription factor JUN to promote endoplasmic reticulum (ER) stress response [33]. However, the imbalance of E2F1 can inhibit the expression of Trib2, thereby improving the survival rate of leukemia patients [34]. Rishi L et al. confirmed that MYC's involvement in FLT3 inhibitors promoting hTERT and down-regulation of Telomerase contributes to the treatment of acute myelogenous leukemia [35].

On the other hand, MYCN promotes the high expression of MYCN by regulating N-Myc downstream regulatory gene 1 (NDRG 1), thereby altering the hematopoietic function of acute myeloid leukemia [36]. Previous studies have shown that NFKB 1 (P50) promoter participates in the induced expression of C/EBP alpha or C/EBP alpha leucine zipper mutants. C/EBP alpha can reduce the activity of histone deacetylase 1 (HDAC 1). It is concluded that C/EBP alpha is a therapeutic target for acute myeloid leukemia [37]. It is worth mentioning that the mutation of RUNX-1 is related to complex change clusters and heredity through the analysis of RUNX-1 transcription factor mutation in acute myeloid leukemia [38]. These regulatory factors play an essential regulatory role in the dysfunction module and constitute the overall regulatory effect of regulating the occurrence and development of acute myeloid leukemia. Therefore, the analysis of multifactorial dysfunction modules and drug relationships can effectively predict the potential therapeutic effects of drugs on acute myeloid leukemia.

Drug resistance and drug toxicity have always existed. In order to scientifically formulate treatment strategies and design effective treatment methods for diseases. Therefore, we attempt to predict drugs that have regulatory effects on AML dysfunction modules. According to the statistical analysis, it was found that Copper had a significant regulatory effect on the six dysfunction modules and had a robust regulatory effect. Several studies have shown that Copper can treat acute myeloid leukemia by targeting leukemia stem-like cells to induce cytotoxic stimulation of the ROS-JNK pathway while inactivating NF-kappa B and Nrf2 [39]. Besides, Indomethacin significantly regulated three dysfunction modules. Lodi A et al. have shown that Indomethacin can still promote oxidative stress response of KG1a and K562 under hypoxic environment, and also cause mitochondrial dysfunction [40]. Finally, other drugs predicted in this study have different regulatory effects on dysfunction modules and are potential drugs for treating or activating acute myeloid leukemia. These results provide valuable references for drug researchers to relocate and identify drug side effects.

4. Conclusions

In conclusion, our work provides a comprehensive and in-depth insight into the pathogenesis of acute myeloid leukemia. Multifactor-mediated dysfunction module not only characterizes the molecular mechanism of acute myelogenous leukemia to a great extent but also lays a theoretical foundation for predicting potential drugs of acute myelogenous leukemia. With the increasing diversity of high-throughput data, this multi-dimensional integration strategy of dysfunction modules helps to elucidate the underlying biological and therapeutic mechanisms of diseases from a global and in-depth perspective.

Reference
2014, 807-818.
32. FERNANDEZ-POL, S., MA, L., OHGAMI, R. S., Immunohistochemistry for p53 is a useful tool to identify cases of acute myeloid leukemia with myelodysplasia-related changes that are TP53 mutated, have complex karyotype, and have poor prognosis. *Mod. Pathol.*, 30, 2017, 382-392.

Manuscript received: 31.07.2020