Am J Transl Res 2020;12(4):1415-1427
www.ajtr.org /ISSN:1943-8141/AJTR0099656

Original Article

Proteomic profiling and bioinformatics analysis identify key regulators during the process from fanconi anemia to acute myeloid leukemia

Hui Hou[1,2], Dan Li[3], Junyan Gao[1], Li Gao[1], Qin Lu[1], Yixin Hu[1], Shuiyan Wu[1], Xinran Chu[1], Yanhua Yao[1], Lin Wan[1], Jing Ling[1], Jian Pan[1], Guoqiang Xu[3], Shaoyan Hu[1]

[1]*Department of Hematology & Oncology, Children's Hospital of Soochow University, Suzhou 215025, Jiangsu, China;* [2]*Inner Mongolia People's Hospital, Hohhot 010010, Inner Mongolia Autonomous Region, China;* [3]*Jiangsu Key Laboratory of Neuropsychiatric Diseases and College of Pharmaceutical Sciences, Soochow University, Suzhou 215123, Jiangsu, China*

Received July 15, 2019; Accepted March 22, 2020; Epub April 15, 2020; Published April 30, 2020

Abstract: Fanconi anemia (FA) is a congenital aplastic anemia, characterized as congenital bone marrow failure, developmental malformation, and the malignant tendency, which may develop into acute myeloid leukemia (AML). However, few studies have been conducted on the progression from FA to AML. In this study, we used proteomic profiling, together with bioinformatics analyses, to explore the molecular mechanisms by which FA progresses to AML. Quantitative proteomic analyses of bone marrow samples identified 168 differentially expressed proteins (DEPs), including 7 upregulated proteins and 161 downregulated proteins in the bone marrow of the FA patient compared with the healthy people. The upregulated proteins were enriched in response to stress, oxygen transport, and hydrogen peroxide catabolic process. The downregulated proteins were enriched in myeloid leukocyte mediated immunity, response to interleukin-12, platelet degranulation and regulation of ATPase activity. Based on these results, we discovered 155 DEPs (142 upregulated and 13 downregulated) in the bone marrow samples between FA and AML patients, of which HIST1H1D, HIST1H3A, PSME1 and THRAP3 may play important roles in the progression of FA to AML and may be used as markers for AML early diagnosis. Finally, cell-line based experiments confirmed that PSME1 had an important effect on the proliferation of leukemia cells.

Keywords: Fanconi anemia, acute myeloid leukemia, bone marrow samples

Introduction

Fanconi anemia (FA) is a congenital aplastic anemia, an autosomal recessive genetic syndrome, which is characterized by congenital bone marrow failure, developmental deformity and malignant tendency [1, 2]. Hematopoietic stem cell genes in FA patients are constantly changing throughout their lives, and FA patients can develop into acute myeloid leukemia (AML), which is characterized by genomic imbalance and chromosome translocation, resulting in copy number abnormality [3-6]. However, how and whether FA develops into AML is unknown. Quantitative proteomics can be used to analyze the differentially expressed proteins in different disease states and explore the potential dysregulated functions and pathways involved in disease progression. Omics analysis of healthy people, FA and AML patient samples may discover key proteins altered during disease development and provide new insights in the development of FA and AML.

Currently, the progression of the disease from FA to AML remains unclear. There is still no good method to determine whether FA patients will further deteriorate into AML. Proteomics is a high-throughput method to study the protein changes in the process of disease. By analyzing the protein changes in bone marrow samples of FA and AML patients, the evolution process from FA to AML can be revealed, providing research ideas for disease diagnosis and treatment. At the same time, potential diagnostic biomarkers in the disease process can also be found.

In this study, we performed quantitative proteomics analysis for the bone marrow of healthy people, FA and AML patients, and identified the differentially expressed proteins (DEPs). The functions and pathways of DEPs were analyzed and visualized using the ClueGO plug-in in CytoScape. The protein-protein interaction network of DEPs was built using STRING online data analysis. The core interaction was extracted by using MCODE in Cytoscape. Through comprehensive analysis, we identified 12 key proteins that may play important roles in the progression of FA to AML. Analysis of the overall patient survival showed that the high expression of HIST1H1D, HIST1H3A, PSME1, and THRAP3 was associated with poor prognosis of AML, which may be used as diagnostic markers from FA malignant progression to AML. This study provides molecular and cellular information in the disease progression of leukemia from FA and may provide a molecular basis for the elucidation of the occurrence of leukemia in the FA patients.

Methods

Samples collection from healthy people, FA and AML patients

Normal bone marrow samples were taken from healthy people. The FA and AML bone marrow samples were taken from an unidentified patient. Patient information was from the primary pathological examination report. In this study, the use of human bone marrow tissue samples was approved by the ethics committee of the Children's Hospital of Soochow University.

Protein sample preparation and LC-MS/MS analysis

Bone marrow cells were centrifuged under 12000 rpm at 4°C for 10 min. The collected sediment was washed three times with ice-cold PBS (pH 7.4). The samples were ground into powder under liquid nitrogen and proteins were extracted with cell lysis buffer. Cell lysate was centrifuge at 12000 rpm and 4°C for 15 min. The supernatant was collected and proteins were precipitated with cold acetone. The precipitated proteins were dissolved in the lysis buffer and digested with trypsin in gel. The resulting peptides from each sample were then labeled with 6-plex TMT labeling reagents for 4 h at room temperature. Labeled samples were mixed at 1:1:1 ratio, fractionated, desalted and analyzed on an Orbitrap mass spectrometry with a previously described procedure [7].

Database search for protein identification

The raw MS/MS data were searched with Proteome Discoverer (Version 1.4) against the UniProt human database [8], which included 20,417 protein sequences. The false discovery rate was set to 1%. The reporter ions from TMT reagents were used for peptide quantification and the relative abundance was normalized with the median ratio. The results were exported to Microsoft Excel for further analysis. The fold changes of > 1.50 were considered as differentially expressed proteins (DEPs).

GSEA analysis

Gene set enrichment analysis was performed by using the Java GSEA Desktop application [9]. The expression matrix (gct file), the phenotype label file and the CLS file were prepared. The enrichment statistic was selected weighted and metric for ranking genes was selected signal to noise. The parameter of number of permutations was set to 1000. The maximal size (exclude larger sets) was set to 500 and minimal size (exclude smaller sets) was set to 5. Then, the file was imported into the GSEA software and the calculation was performed.

GO and KEGG analysis of DEPs

The functions of DEPs were analyzed using the ClueGO [10] plug-in application in Cytoscape [11]. KEGG pathway enrichment analysis was carried out and visualized using ClueGO and CluePedia [12] with $P < 0.05$ as the cut-off value.

Integration of protein-protein interaction (PPI) network

The online database STRING (http://string-db.org) [13] was used to build the PPI network of DEPs. The protein interaction network table was downloaded and visualized using Cytoscape software. MCODE plug-in of Cytoscape was used to perform sub-module analysis.

Survival analysis of AML patients

Survival analysis of AML patients were carried out using the online gene expression profiling

interactive analysis tool (GEPIA, http://gepia.cancer-pku.cn/index.html) [14]. Hazard ratio and 95% confidence were selected and datasets used for the analysis were acute myeloid leukemia.

Results

Sample information and gene test results of FA patients

The bone marrow samples of the control group were taken from healthy people. The FA and AML samples were taken from a female patient who was diagnosed as FA at 8 years old and again AML at 14 years old. Next generation gene sequencing discovered that FANCA gene had a loss of function homozygous mutation in exon34 c.3348+ 1G > A. Because she had a fever for 2 months, the patient was given bone marrow aspiration which identified that blast cells was 20% and she was diagnosed as AML-M5a. Immunotyping confirmed the diagnosis with blast cells expressed aberrant myeloid phenotype.

Statistical analysis of proteome in the bone marrow of the FA patient and healthy person

To explore the potential biomarkers for the development of FA, we sought to identify the differentially expressed proteins in the bone marrow of the healthy person and FA patient using quantitative proteomics according to the TMT labeling procedure (**Figure 1A**). The analysis of physicochemical properties of proteins reflects the quality of the data obtained from mass spectrometry analysis. In this study, the sequence coverage of the most identified proteins was greater than 5%, indicating that proteins identified by proteomics were credible (**Figure 1B**). The molecular weight of most identified proteins is within the range of 10-90 kDa (**Figure 1C**). Only three proteins are less than 10 kDa and 47 proteins are more than 90 kDa. Peptide spectrum matches (PSMs) roughly reflects the protein abundance. The PSMs of most proteins in this experiment were greater than 2, indicating that the identified results were reliable (**Figure 1D**). Most of the identified proteins have an isoelectric point between 5 and 10. The isoelectric point distribution of the identified proteins was also reasonable (**Figure 1E**).

Function and pathway analysis of differentially expressed proteins in the bone marrow of the FA patient compared with healthy person

Compared with normal samples from healthy people, quantitative proteomics identified 168 differentially expressed proteins (DEPs), including 7 upregulated and 161 downregulated proteins in the FA samples. The GSEA analysis (**Figure 2A**) showed that adipogenesis, glucolysis and MYC targets were decreased in FA patients. The enrichment results indicated that the normal physiological function of FA patients was affected compared with healthy people. Enrichment analysis of GO and KEGG pathways was conducted with CytoScape ClueGo plug-in. GO functional enrichment analysis includes biological processes (BP), molecular functions (MF) and cellular compartment (CC). The enrichment analysis (**Figure 2B**) showed that the upregulated proteins were related to the response to stress, oxygen transport, gas transport and hydrogen peroxide catabolic processes. Downregulated proteins were mainly involved in myeloid leukocyte mediated immunity, cellular response to interleukin-12, platelet degranulation, cellular oxidant detoxification and plasminogen activation. ClueGO analysis of KEGG pathways of down-regulated proteins showed that these proteins were associated with regulation of actin cytoskeleton, systemic lupus erythematous, glycolysis/gluconeogenesis and longevity regulating pathway (**Figure 2C**).

The construction and modularization analysis of the PPI of the DEPs in the bone marrow of the FA patient and the healthy person

To further explore the possible connection among the DEPs, we conducted the protein-protein interaction (PPI) network analysis for the 160 downregulated proteins by STRING [15]. The cutoff score for the PPI network analysis was set as 0.4 (medium confidence). The final PPI network includes 152 nodes and 1051 edges. We used the MCODE plug-in module to select the most important submodules in PPI to further understand the biological significance of DEPs resulted from the FA patient. The most significant module (MCODE score = 17.579) consisted of 20 nodes and 167 edges (**Figure 3A**). KEGG analysis showed that these proteins were mainly concentrated in protein process-

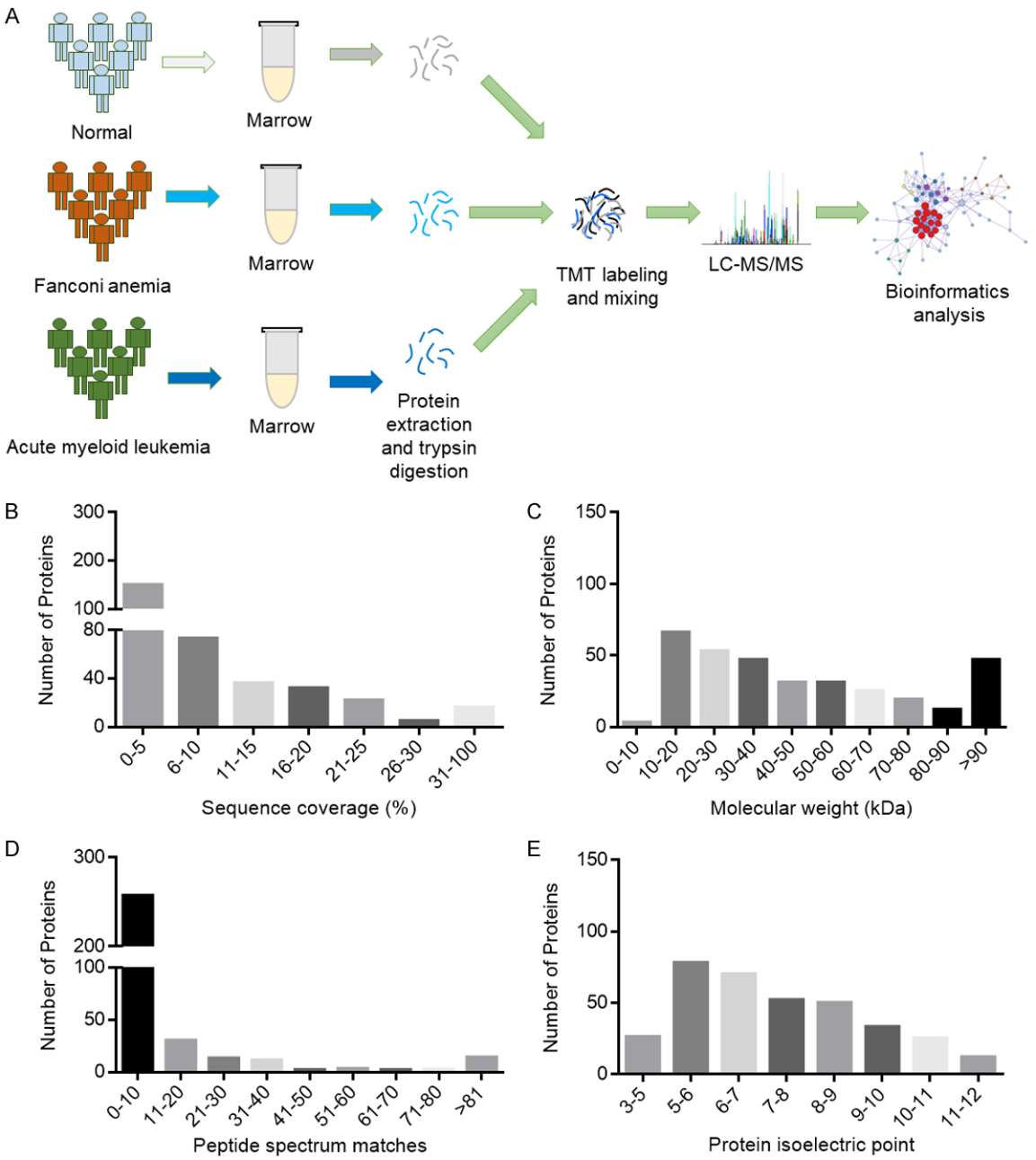


Figure 1. Statistical analysis of proteins identified from the bone marrows of the healthy person, FA and AML patients. A. Procedure for sample preparation, TMT labeling, LC-MS/MS and data analysis. B. Sequence coverage distribution of the identified proteins. The results showed that protein coverage was high and protein identification was reliable. C. Protein mass distribution of the identified proteins. The results showed that the molecular weight distribution of the protein was reasonable and there was no obvious abnormal distribution of the proteins. D. The distribution of peptide spectrum matches (PSMs) of the identified proteins. E. Protein isoelectric point distribution of the identified proteins. The results showed that the distribution of isoelectric points of proteins was reasonable and the identification results were reliable.

ing in endoplasmic reticulum, ribosome biogenesis in eukaryotes and AMPK signaling pathway (**Figure 3B**). Further enrichment analysis was carried out on the proteins in these submodules, and the enrichment results showed that proteins in the most significant modules were mainly related to nuclear-transcribed mRNA catabolic process, nonsense-mediated decay, cytoplasmic translation, positive regulation of DNA biosynthetic process, telomere mainte-

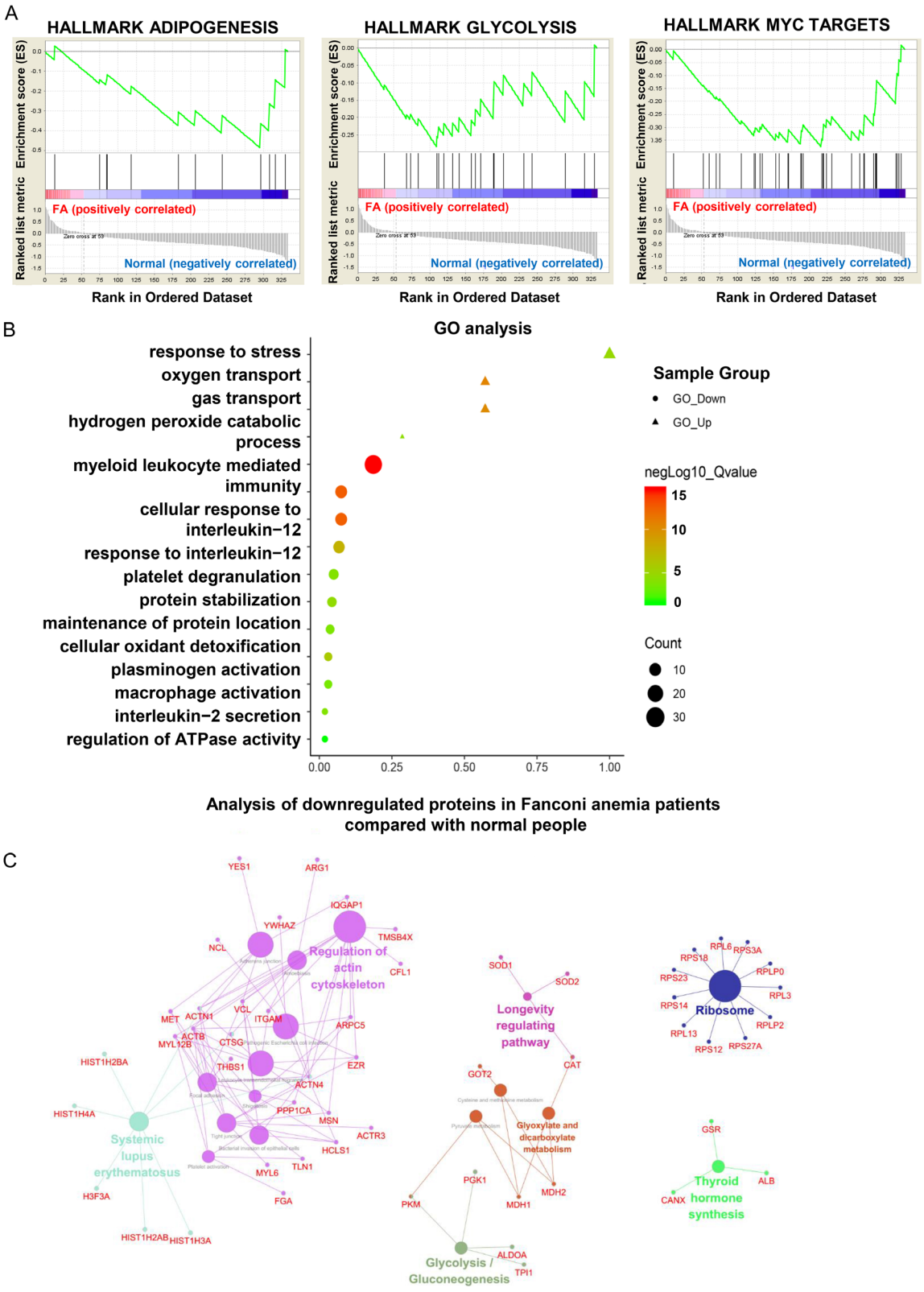
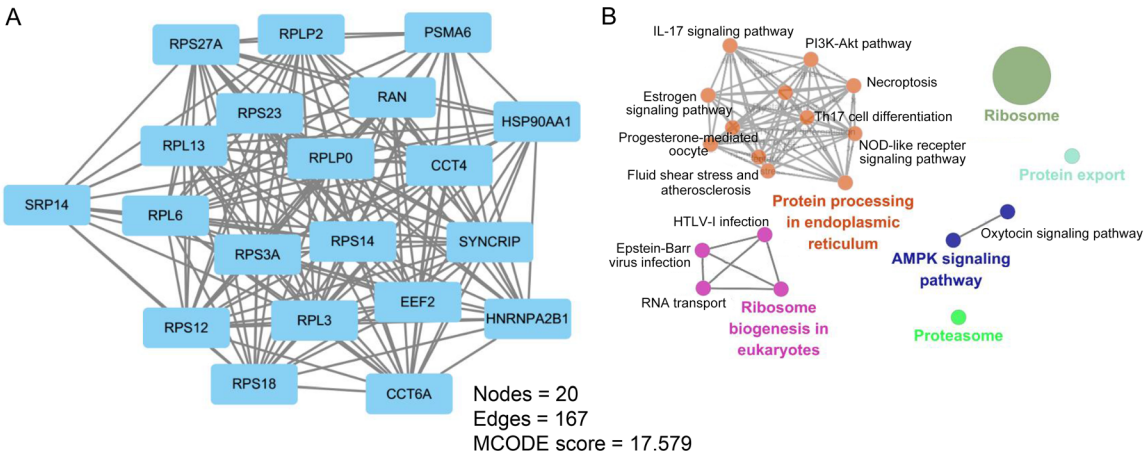


Figure 2. HALLMARK, Gene oncology and signaling pathway enrichment analysis of DEPs in the bone marrow of healthy person and FA patients. A. HALLMARK analysis of DEPs. Results showed that adipogenesis, glycolysis and myc targets were decreased in FA patients. B. GO analysis of DEPs. The results showed that compared with healthy people, patients with FA had increased oxidative stress, oxygen transport and gas transport capacity, but decreased cellular immune response. C. KEGG pathways of downregulated proteins. The downregulated proteins were associated with regulation of actin cytoskeleton, glycolysis/gluconeogenesis and longevity regulating pathway. Function

and signaling pathway enrichment of DRPs were conducted using the plug-in application in Cytoscape. DEPs, differentially expressed proteins; KEGG, Kyoto Encyclopedia of Genes and Genomes.



C

Term	*P*-value	Number of genes	Associated proteins found
Nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	1.43524E-19	11	[RPL13, RPL3, RPL6, RPLP0, RPLP2, RPS12, RPS14, RPS18, RPS23, RPS27A, RPS3A]
Cytoplasmic translation	6.13344E-07	4	[EEF2, RPL6, RPLP2, RPS3A]
Positive regulation of DNA biosynthetic process	9.21913E-07	4	[CCT4, CCT6A, HNRNPA2B1, HSP90AA1]
Telomere maintenance via telomere lengthening	1.78299E-06	4	[CCT4, CCT6A, HNRNPA2B1, HSP90AA1]
Positive regulation of telomerase activity	8.18261E-06	3	[CCT4, HNRNPA2B1, HSP90AA1]
Positive regulation of telomere maintenance via telomere lengthening	9.5704E-06	3	[CCT4, CCT6A, HNRNPA2B1]
Regulation of telomerase activity	2.00242E-05	3	[CCT4, HNRNPA2B1, HSP90AA1]
Positive regulation of telomere maintenance	2.38023E-05	3	[CCT4, CCT6A, HNRNPA2B1]
Ribosome assembly	4.35603E-05	3	[RPL3, RPL6, RPS14]
Regulation of telomere maintenance via telomere lengthening	4.35603E-05	3	[CCT4, CCT6A, HNRNPA2B1]
Telomere maintenance via telomerase	6.38924E-05	3	[CCT4, CCT6A, HSP90AA1]
Ribosome	1.64927E-18	12	[EEF2, RPL13, RPL3, RPL6, RPLP0, RPLP2, RPS12, RPS14, RPS18, RPS23, RPS27A, RPS3A]
rRNA binding	5.7656E-07	4	[EEF2, RPL3, RPS14, RPS18]

Figure 3. Protein-protein interaction (PPI) network of the most significant modules. A. Most significant module of the PPI network. The PPI network contains 20 nodes and 167 interacting relationships. PPI network selection score is greater than 0.4. B. Enriched pathways of the most significant modules. The results showed that protein processing in endoplasmic reticulum, AMPK signaling pathway, ribosome biogenesis in eukaryotes and proteasome pathways were significantly enriched. C. Enriched biological processes in the most significant modules. The results showed that cytoplasmic translation, positive regulation of DNA biosynthetic process and regulation of telomerase activity functions were significantly enriched.

nance via telomere lengthening, positive regulation of telomerase activity, positive regulation of telomere maintenance via telomere lengthening and regulation of telomerase activity (**Figure 3C**). Cellular compartment analysis showed that these proteins were enriched in the ribosome. Molecular function analysis indicated that the downregulated proteins were mainly related to the function of rRNA binding.

GO and KEGG enrichment analysis of DEPs between FA and AML

To further understand the biological process of FA deterioration into AML, we conducted GO and KEGG enrichment analysis on DEPs. A total of 155 DEPs were identified (142 upregulated and 13 downregulated) in the bone marrow of the FA and AML patient. The GSEA analysis (**Figure 4A**) showed that compared with FA patients, the epithelial mesenchymal transition, MYC targets, PI3K/AKT/mTOR signaling, oxidative phosphorylation and mitotic spindle were increased while the heme metabolism function was decreased in the bone marrow of the AML patient. The upregulated DEPs were particularly abundant in leukocyte activation involved in immune response, myeloid leukocyte activation, actin cytoskeleton reorganization, and tumor necrosis factor superfamily cytokine production. Meanwhile, downregulat-

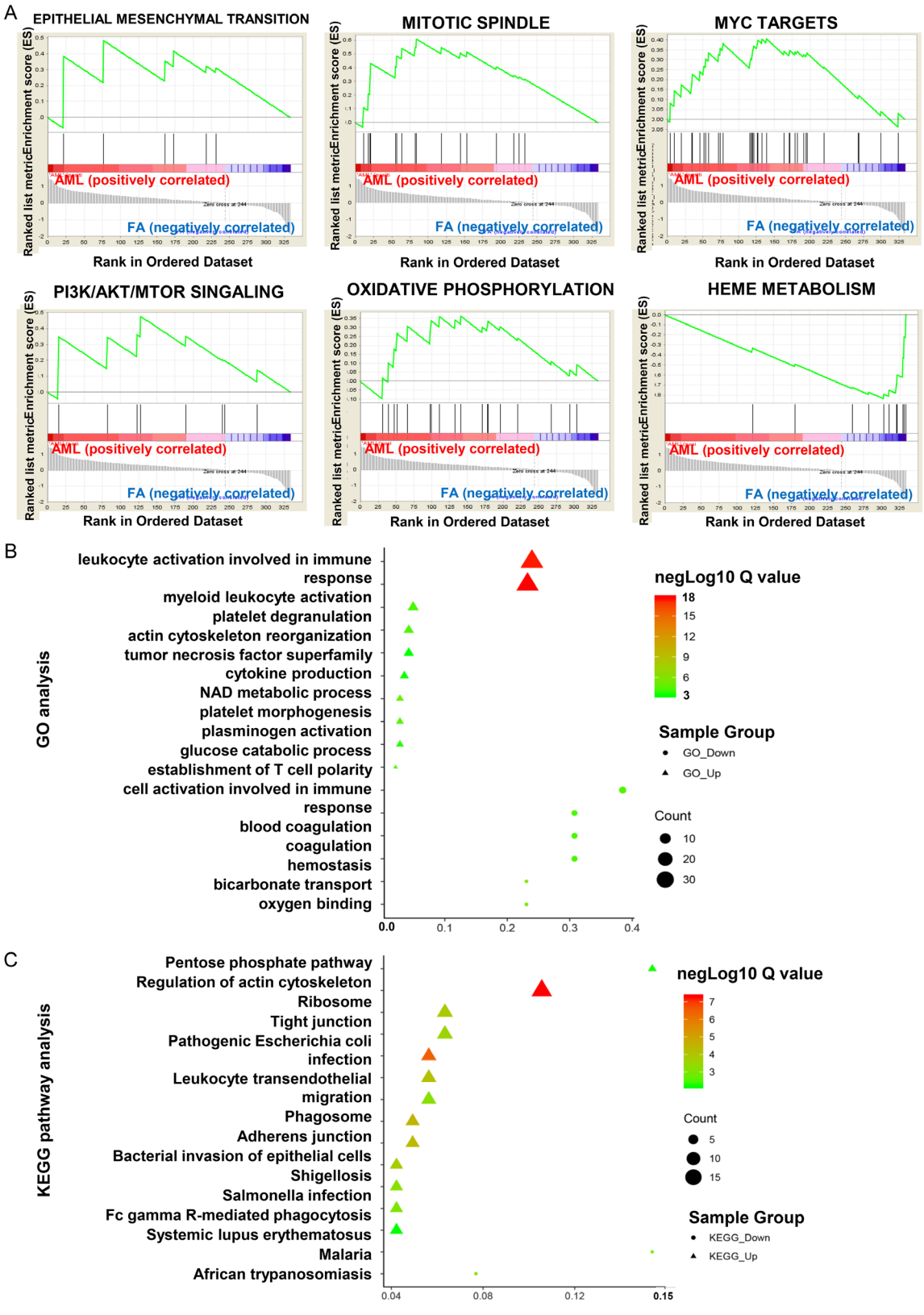


Figure 4. HALLMARK, Gene oncology and signaling pathway enrichment analysis of DEPs in the bone marrow of FA and AML patients. (A) HALLMARK analysis of DEPs. The PI3K/Akt/mTOR signaling pathway was activated and oxidative phosphorylation was increased in AML patients compared with FA patients (B) GO analysis of DEPs. The leukocyte activation involved in immune response and myeloid leukocyte activation were increased in AML patients. (C) KEGG pathways. The dots represent the upregulated functions and pathways while the triangles represent the

downregulated functions and pathways. The size of dots and triangles indicates the number of proteins accumulated and the color represents degree of enrichment.

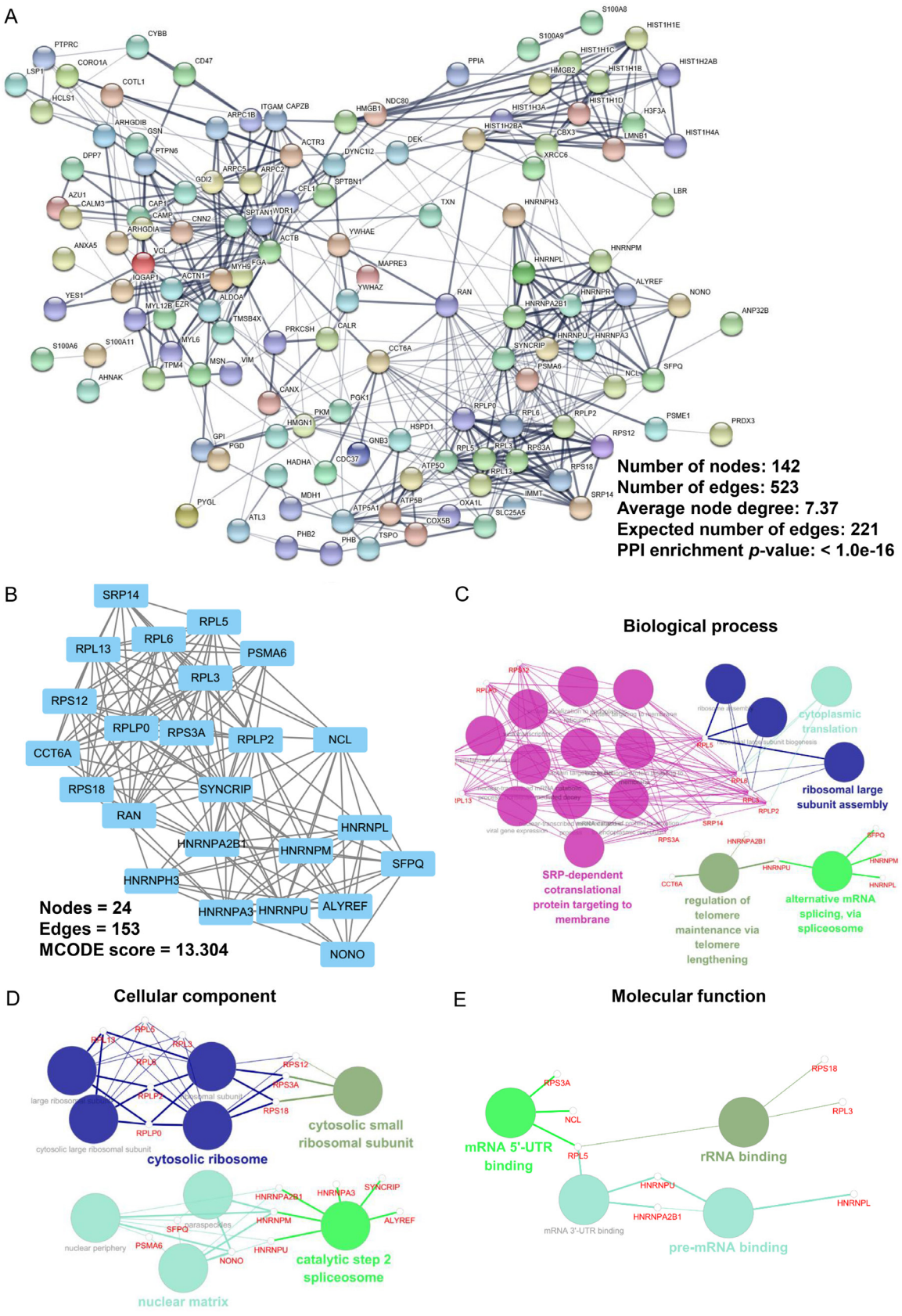


Figure 5. Protein-protein interaction network and GO enrichment analysis of the upregulated proteins in the bone marrow of FN and AML patients. A. PPI network of upregulated proteins. The PPI network contains 142 nodes and 523 interacting relationships. B. Most significant module of the PPI network. C. Biological process enrichment analysis of most significant modules of the PPI network. The cytoplasmic translation, SRP-dependent cotranslational protein targeting to membrane and regulation of telomere maintenance functions were increased. D. Cellular component enrichment analysis. The differences in cell composition were mainly concentrated in cytosolic ribosome and nuclear matrix. E. Molecular function enrichment analysis. The rRNA binding, pre-mRNA binding and mRNA binding functions were enriched.

ed proteins were mainly involved in cell activation involved in immune response, blood coagulation, hemostasis bicarbonate transport and oxygen binding (**Figure 4B**). KEGG pathway enrichment analysis showed the upregulated DEPs were enriched in the pentose phosphate pathway, regulation of actin cytoskeleton, tight junction, leukocyte transendothelial migration, salmonella infection and system lupus erythematosus and other aspects (**Figure 4C**).

Modularization analysis of the PPI of the DEPs in the bone morrow of the AML and FA patient

The protein-protein interaction network of 142 upregulated proteins was established with STRING. The total PPI network includes 142 nodes, 523 edges and 221 expected number of edges. The PPI enrichment *P*-value was less than 1.0×10^{-16} (**Figure 5A**). The most significant module (MCODE score = 13.304) consisted of 24 nodes and 153 edges (**Figure 5B**). Further enrichment analysis was carried out on the proteins in this submodule, and the biological process enrichment analysis showed that proteins in the most significant module were mainly related to signal recognition particle (SRP)-dependent cotranslational protein targeting to membrane, ribosomal large subunit assembly and regulation of telomere maintenance via telomere lengthening (**Figure 5C**). Cell composition analysis showed that these proteins were enriched in the cytosolic ribosome and nuclear matrix (**Figure 5D**). Molecular function analysis showed that the downregulated proteins were mainly related to the mRNA 5'-UTR binding, rRNA binding and pre-mRNA binding (**Figure 5E**).

Key genes modulating the overall survival rate of AML

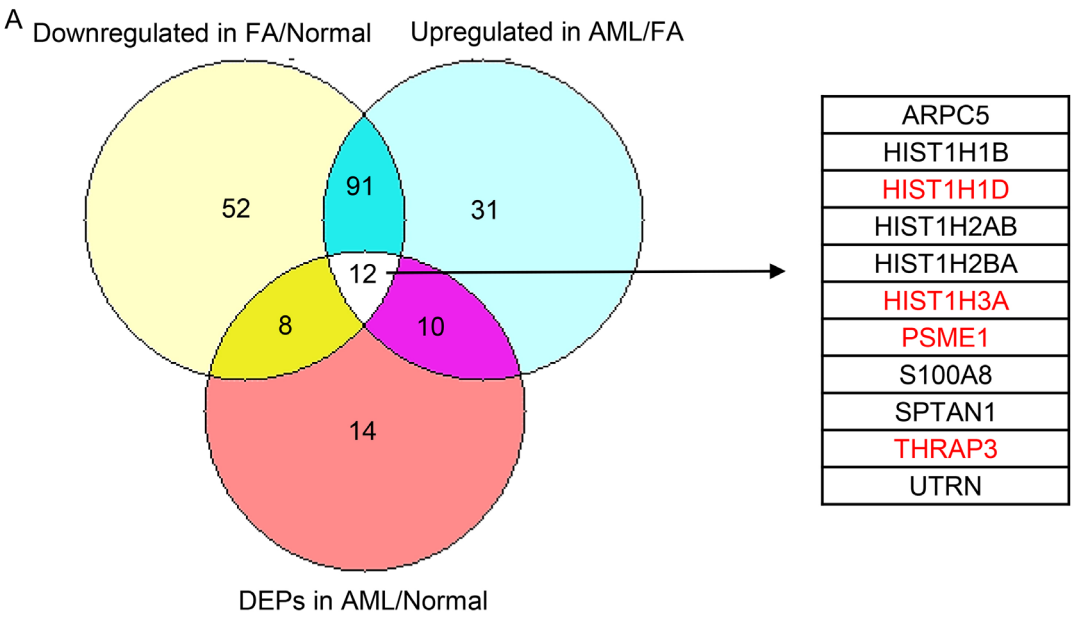
Through the analysis of proteomics data, we found that a series of proteins decreased during the process from healthy people to FA and increased during the process from FA to AML. To further investigate the altered proteins during the progression of FA into AML, Venn diagram analysis was performed on proteins downregulated by FA/Normal, upregulated by AML/FA, and differentially expressed proteins of AML/Normal. This analysis revealed 12 common proteins, including ARPC5, HIST1H1B, HIST1H1D, HIST1H2AB, HIST1H2BA, HIST1H3A, PSME1, S100A8, SPTAN1, THRAP3, and UTRN (**Figure 6A**). In our previous study, we have found that high level of S100A8 expression had a worse outcome in pediatric AML patients [16]. In this study, we further analyzed the effect of other proteins on the outcome of pediatric AML in mRNA level with GEPIA (http://gepia.cancer-pku.cn/index.html). We chose AML (acute myeloid leukemia) in the database and set a 95% confidence interval for this analysis. The analysis found that the high expression of four proteins HIST1H1D, HIST1H3A, PSME1, and THRAP3 was related to the poor prognosis of leukemia (**Figure 6B-E**). These data suggest that these four proteins may be key proteins in the progression of the disease and may serve as early markers for the evolution to AML.

PSME1 affects the biological function of leukemia cells

To further investigate the biological functions of key proteins derived from proteomics, we investigated the effects of PSME1 on leukemia cells (K562). The results showed that all three knockdown shRNA could significantly reduce the expression of PSME1 protein (**Figure 7A**). Further experiments showed that PSME1 knockdown could significantly inhibit the proliferation of K562 cells (**Figure 7B**). The results indicate that PSME1 plays an important role in leukemia cells.

Discussion

The development of proteomic techniques provides more possibilities for the study of disease progression. Only a small size of samples is required to detect important regulatory pro-


A
Downregulated in FA/Normal
Upregulated in AML/FA
52
91
31
12
8
10
14
DEPs in AML/Normal
ARPC5
HIST1H1B
HIST1H1D
HIST1H2AB
HIST1H2BA
HIST1H3A
PSME1
S100A8
SPTAN1
THRAP3
UTRN

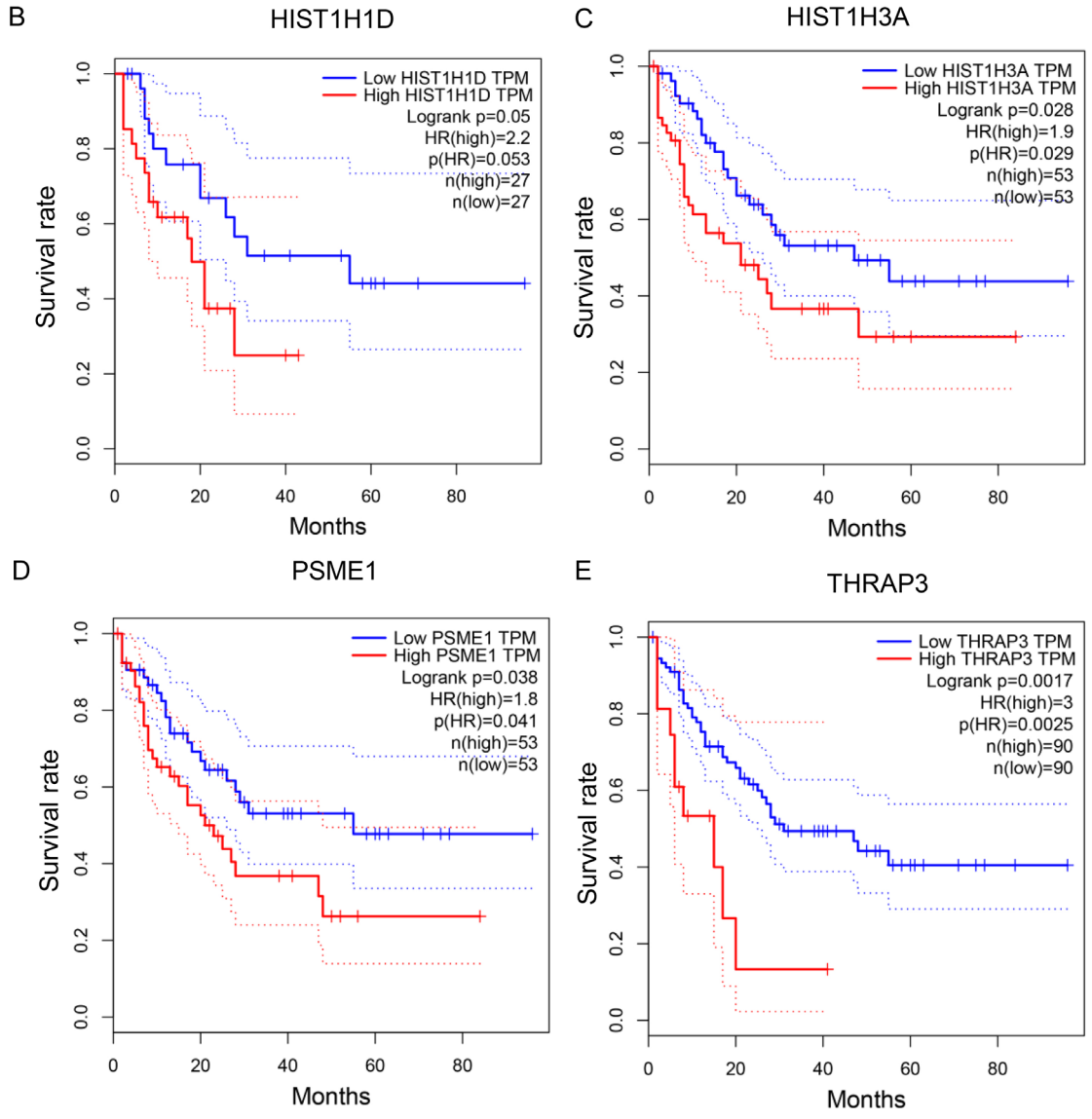

B
HIST1H1D
Low HIST1H1D TPM
High HIST1H1D TPM
Logrank p=0.05
HR(high)=2.2
p(HR)=0.053
n(high)=27
n(low)=27
Survival rate
Months
C
HIST1H3A
Low HIST1H3A TPM
High HIST1H3A TPM
Logrank p=0.028
HR(high)=1.9
p(HR)=0.029
n(high)=53
n(low)=53
Survival rate
Months
D
PSME1
Low PSME1 TPM
High PSME1 TPM
Logrank p=0.038
HR(high)=1.8
p(HR)=0.041
n(high)=53
n(low)=53
Survival rate
Months
E
THRAP3
Low THRAP3 TPM
High THRAP3 TPM
Logrank p=0.0017
HR(high)=3
p(HR)=0.0025
n(high)=90
n(low)=90
Survival rate
Months

Figure 6. A. Venn diagram of proteins that are downregulated in FN/Normal, upregulated in AML/FN, and DEPs in AML/Normal bone marrow samples. Twelve common proteins were listed. B. Overall survival curve between patients with low (green line) and high (red line) HIST1H1D mRNA expression. Compared with AML patients with low expression of HIST1H1D, patients with high expression of HIST1H1D had a significantly shorter survival time. C. Overall survival curve between patients with low (green line) and high (red line) HIST1H3A expression. The expression of HIST1H3A protein was negatively correlated with the survival time of patients. D. Overall survival curve between patients with low (green line) and high (red line) PSME1 expression groups. The higher the expression of PSME1 protein, the shorter the survival time of patients. E. Overall survival curve between patients with low (green line) and high (red line) THRAP3 expression groups. The higher the expression of THRAP3 protein, the shorter the survival time of patients. The X axis indicates overall survival time (months), and the Y axis indicates the present survival (%). The dotted line represents the confidence interval.

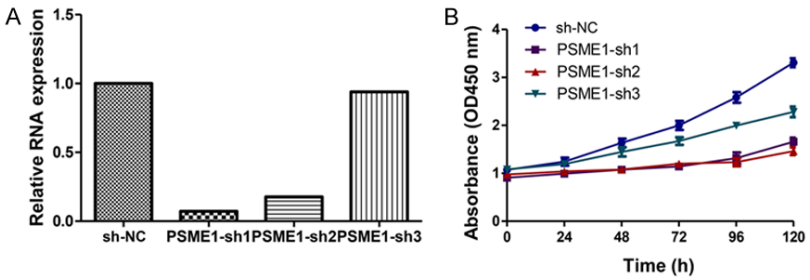


Figure 7. Biological function of PSME1 in K562 cells. A. After shRNA PSME1 transfection, the expression level of PSME1 was detected. B. The effect of PSME1 on the proliferation of leukemia cells. The results showed that knockdown of PSME1 protein could significantly inhibit cell proliferation.

teins and to analyze the associated biological functions. Using proteomic profiling and bioinformatics analysis, we comprehensively analyzed the pathogenesis, disease progression, and functions and pathways that differ between FA and AML in a small size of patient samples. In this study, high-throughput proteomics was used to study the whole proteome in the bone marrow of healthy person, FA, and AML patients. The differences in functions and pathways between FA and healthy person were revealed by analyzing the DEPs using various databases. A series of specific proteins with high expression level were discovered during the process from FA to AML. Analysis of overall survival rate revealed that HIST1H1D, HIST1H3A, PSME1, and THRAP3 were important regulators for leukemia. Therefore, this work provided valuable information and basis for the high susceptibility of FA patients to AML.

The rate for the gene mutation of genetic diseases in the chromosomal repair system of FA patients is relatively high (14.4%) while the DNA repair ability is defective, leading to the increased chromosome aberration and organ malformation. There are defects in pluripotent stem cell function caused by humoral inhibitory factors in the bone marrow microenvironment [17]. Studies have also shown that some FA patients with fibroblasts continue to produce excessive amounts of cell growth factors, suggesting that they may be prone to develop leukemia [18-20]. In this study, the results of key MCODE analysis (**Figure 3C**) showed that many telomerase-related functions were attenuated in patients with FA compared with normal subjects. Apoptosis dysregulation and telomere maintenance abnormalities were also discovered in FA cells [21, 22]. The spontaneous apoptosis of FA cells was increased, which may be associated with the ability of FA cells to repair damage. Telomeres are involved in chromosome stability and cell proliferation and telomeres shorten when cells divide. Telomere shortening is accelerated in the FA cells and is more pronounced in severe FA cases. It had high breaks for the telomere sequence in FA cells. This information suggests that telomeres remain defective in FA. In another aspect, the hematopoietic dysfunction of cannabinemia is probably due to high apoptotic rate and defects in telomere maintenance.

In this study, abnormal oxygen metabolism was found in FA patients. Dysfunction in oxygen metabolism is an important factor in the defective pathological DNA repair in FA cells [23]. FA cells are damaged by accumulated oxygen free radicals that are produced by mutagens such as high oxygen tension, gamma rays, attractants, and drugs that produce active hydroxide [24-27]. The cell cycle progression of FA is also abnormal. FA cells grow slowly, G2 phase is prolonged, and G2/M phase transition is delayed or completely arrested, which further increase

the exposure of cells to high oxygen concentration, resulting in an increase in oxidative damage [28-30]. This oxidative damage also increases the possibility that FA deteriorates into AML. Abnormal oxygen metabolism may be correlated with the pathogenesis of AML and oxygen metabolism disorder is an early event in FA patients that develop into leukemia.

Histones regulate gene expression as they interact with DNA to form chromatin and protect or unwrap DNA. Histones mainly affect the epigenetic function and epigenetic alteration, such as DNA methylation and histone modifications, can modulate gene expression and play important roles in the occurrence and development of tumors [31]. Our study suggests that histones may be involved in the deterioration of FA into AML [32].

Conclusions

FA Patients are at increased risk of myelopoietic cell carcinoma-acute myeloid leukemia (AML). Proteomics analysis can detect proteins that are differentially regulated and thus bioinformatics analysis can be used to discover abnormalities in the functions and pathways between the bone marrow of FA and AML. This work would provide key markers for the future analysis of the progression of FA to AML. Our study revealed the functions and pathways of the differentially regulated proteins in the bone marrow of FA and healthy people, and also explored the biological processes that are dysregulated during the development of FA o AML. These results suggest that HIST1H1D, HIST1H3A, PSME1, and THRAP3 may be useful markers for the diagnosis of FA-induced AML. Further functional experiments confirmed that PSME1 could significantly affect the proliferation of leukemia cells and play an important role in the development of leukemia, which is worthy of further study.

Acknowledgements

Hui Hou, Dan Li, Junyan Gao, Li Gao, Qin Lu, Yixin Hu, Shuiyan Wu, and Xinran Chu contributed to the work equally. The work was supported by the Natural Science Foundation of China (NO. 81770193, 81700165, 81770145, 81800185), Jiangsu Province Projects (NO. BE2017659, BE2017658, BRA2017541, BK20180053 and CXTDA2017014), and Suzhou projects (SZS201615, SZZX201504, SS201809 and SYS201643).

Disclosure of conflict of interest

None.

Address correspondence to: Shaoyan Hu, Department of Hematology & Oncology, Children's Hospital of Soochow University, Suzhou 215025, Jiangsu, China. Tel: +86-512-80692929; E-mail: hushaoyan@suda.edu.cn; Guoqiang Xu, College of Pharmaceutical Sciences, Soochow University, Suzhou 215123, Jiangsu, China. E-mail: gux2002@suda.edu.cn

References

[1] Crawford D and Dearmun A. Fanconi anaemia. Nurs Child Young People 2017; 29: 17.
[2] Tischkowitz MD and Hodgson SV. Fanconi anaemia. J Med Genet 2003; 40: 1-10.
[3] Mathew CG. Fanconi anaemia genes and susceptibility to cancer. Oncogene 2006; 25: 5875-5884.
[4] Minton K. Inflammation: inflammatory pathology of Fanconi anaemia. Nat Rev Immunol 2016; 16: 336-337.
[5] Wrighton KH. DNA repair: a new fan of the Fanconi anaemia pathway. Nat Rev Mol Cell Biol 2010; 11: 603.
[6] Nalepa G and Clapp DW. Fanconi anaemia and cancer: an intricate relationship. Nat Rev Cancer 2018; 18: 168-185.
[7] Zhu Y, Lei Q, Li D, Zhang Y, Jiang X, Hu Z and Xu G. Proteomic and biochemical analyses reveal a novel mechanism for promoting protein ubiquitination and degradation by UFBP1, a key component of ufmylation. J Proteome Res 2018; 17: 1509-1520.
[8] UniProt Consortium. UniProt: a hub for protein information. Nucleic Acids Res 2015; 43: D204-212.
[9] Subramanian A, Kuehn H, Gould J, Tamayo P and Mesirov JP. GSEA-P: a desktop application for gene set enrichment analysis. Bioinformatics 2007; 23: 3251-3253.
[10] Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z and Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009; 25: 1091-1093.
[11] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.

[12] Bindea G, Galon J and Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics 2013; 29: 661-663.
[13] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ and Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-613.
[14] Tang Z, Li C, Kang B, Gao G, Li C and Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017; 45: W98-W102.
[15] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ and Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-D613.
[16] Yang XY, Zhang MY, Zhou Q, Wu SY, Zhao Y, Gu WY, Pan J, Cen JN, Chen ZX, Guo WG, Chen CS, Yan WH and Hu SY. High expression of S100A8 gene is associated with drug resistance to etoposide and poor prognosis in acute myeloid leukemia through influencing the apoptosis pathway. Onco Targets Ther 2016; 9: 4887-4899.
[17] Lecourt S, Vanneaux V, Leblanc T, Leroux G, Ternaux B, Benbunan M, Chomienne C, Baruchel A, Marolleau JP, Gluckman E, Socie G, Soulier J and Larghero J. Bone marrow microenvironment in fanconi anemia: a prospective functional study in a cohort of fanconi anemia patients. Stem Cells Dev 2010; 19: 203-208.
[18] Hess CJ, Ameziane N, Schuurhuis GJ, Errami A, Denkers F, Kaspers GJ, Cloos J, Joenje H, Reinhardt D, Ossenkoppele GJ, Zwaan CM and Waisfisz Q. Hypermethylation of the FANCC and FANCL promoter regions in sporadic acute leukaemia. Cell Oncol 2008; 30: 299-306.
[19] Langmuir PB, Aplenc R and Lange BJ. Acute myeloid leukaemia in children. Best Pract Res Clin Haematol 2001; 14: 77-93.
[20] Cavenagh JD, Richardson DS, Gibson RA, Mathew CG and Newland AC. Fanconi's anaemia presenting as acute myeloid leukaemia in adulthood. Br J Haematol 1996; 94: 126-128.
[21] Sarkar J and Liu Y. Fanconi anemia proteins in telomere maintenance. DNA Repair (Amst) 2016; 43: 107-112.
[22] Brosh RM Jr, Bellani M, Liu Y and Seidman MM. Fanconi anemia: a DNA repair disorder characterized by accelerated decline of the hematopoietic stem cell compartment and other features of aging. Ageing Res Rev 2017; 33: 67-75.
[23] Uziel O, Reshef H, Ravid A, Fabian I, Halperin D, Ram R, Bakhanashvili M, Nordenberg J and Lahav M. Oxidative stress causes telomere damage in Fanconi anaemia cells - a possible predisposition for malignant transformation. Br J Haematol 2008; 142: 82-93.
[24] Du W, Adam Z, Rani R, Zhang X and Pang Q. Oxidative stress in Fanconi anemia hematopoiesis and disease progression. Antioxid Redox Signal 2008; 10: 1909-1921.
[25] Lyakhovich A. Damaged mitochondria and overproduction of ROS in Fanconi anemia cells. Rare Dis 2013; 1: e24048.
[26] Jomova K and Valko M. Importance of iron chelation in free radical-induced oxidative stress and human disease. Curr Pharm Des 2011; 17: 3460-3473.
[27] Degan P, Bonassi S, De Caterina M, Korkina LG, Pinto L, Scopacasa F, Zatterale A, Calzone R and Pagano G. In vivo accumulation of 8-hydroxy-2'-deoxyguanosine in DNA correlates with release of reactive oxygen species in Fanconi's anaemia families. Carcinogenesis 1995; 16: 735-741.
[28] Sala-Trepat M, Rouillard D, Escarceller M, Laquerbe A, Moustacchi E and Papadopoulo D. Arrest of S-phase progression is impaired in Fanconi anemia cells. Exp Cell Res 2000; 260: 208-215.
[29] Centurion SA, Kuo HR and Lambert WC. Damage-resistant DNA synthesis in Fanconi anemia cells treated with a DNA cross-linking agent. Exp Cell Res 2000; 260: 216-221.
[30] Sabatier L and Dutrillaux B. Effect of caffeine in Fanconi anemia. I. Restoration of a normal duration of G2 phase. Hum Genet 1988; 79: 242-244.
[31] Kim HM, Beese-Sims SE and Colaiacovo MP. Fanconi anemia FANCM/FNCM-1 and FANCD2/FCD-2 are required for maintaining histone methylation levels and interact with the histone demethylase LSD1/SPR-5 in caenorhabditis elegans. Genetics 2018; 209: 409-423.
[32] Renaud E, Barascu A and Rosselli F. Impaired TIP60-mediated H4K16 acetylation accounts for the aberrant chromatin accumulation of 53BP1 and RAP80 in Fanconi anemia pathway-deficient cells. Nucleic Acids Res 2016; 44: 648-656.