Urinary Proteomics Study of Patients with Transfusion-Dependent Thalassemia

Authors: Chen Yuzhen, Chen Weijie, Xinhua Zhang, Li Wang, Huang Binbin, Ye Yuhua, Xiangmin Xu, Gao Youhe, Ye Yuhua, Xiangmin Xu, Gao Youhe

Date: 2026-04-07T14:47:38+00:00

Abstract

Thalassemia is a severe hereditary hemoglobinopathy that poses a significant threat to human health, particularly in southern China, where its clinical treatment still faces numerous limitations. This study is the first to meticulously explore the systemic impact of thalassemia through urinary proteomics. By comparing the urinary proteins of transfusion-dependent thalassemia patients with those of healthy controls, a total of 816 differentially expressed proteins were identified, 19 of which exhibited an "all-or-none" change pattern. Simultaneously, random permutation test results indicated that at least 98.92% of these differentially expressed proteins were not generated by chance. Several differentially expressed proteins have been reported to be associated with thalassemia or its therapeutic mechanisms; for instance, TFR2 has been reported as a novel potential therapeutic target for β-thalassemia, and SMAD3 is related to the mechanism of luspatercept, a drug approved by the U.S. Food and Drug Administration for the treatment of transfusion-dependent thalassemia. Furthermore, multiple biological pathways previously reported to be associated with thalassemia were enriched, suggesting that the urinary proteome has the potential to provide references and clues for identifying potential therapeutic targets for thalassemia. Additionally, several significantly altered proteins and biological pathways not previously reported in association with thalassemia may provide new insights for the research and optimization of diagnosis and treatment strategies for transfusion-dependent thalassemia.

Full Text

Preamble

Urinary Proteomics Study of Patients with Transfusion-Dependent Thalassemia

Yuzhen Chen [1†], Weijie Chen [2†], Xinhua Zhang [3], Li Wang [4], Binbin Huang [5], Yuhua Ye [2*], Xiangmin Xu [2*], Youhe Gao [1*]

1．北京师范大学生命科学学院基因工程药物及生物技术北京市重点实验室，中国北京

2. Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, Guangdong Province, China
3. Department of Hematology, The 923rd Hospital of the People's Liberation Army, Nanning 530021, Guangxi Province, China
4. Department of Pediatrics, The 923rd Hospital of the People' s Liberation Army, Nanning 530021, Guangxi Province, China
5. Department of Internal Medicine, Nanning Sixth People' s Hospital, Nanning 530022, Guangxi Province, China

† These authors contributed equally to this work. * Corresponding authors: Youhe Gao, E-mail: gaoyouhe@bnu.edu.cn; Xiangmin Xu, E-mail: xixm@smu.edu.cn; Yuhua Ye, E-mail: yeyuhua_{genetics}@163.com.

摘要

Thalassemia is a severe hereditary hemoglobinopathy that poses a significant threat to human health, particularly in southern China. Currently, clinical treatments for this condition remain subject to numerous limitations. This study represents the first comprehensive exploration of the systemic effects of thalassemia using urinary proteomics. By comparing the urinary protein profiles of patients with transfusion-dependent thalassemia (TDT) against a healthy control group, we identified a total of 816 differentially expressed proteins (DEPs). Notably, 19 of these proteins exhibited qualitative "on/off" changes, appearing or disappearing entirely between the two groups. Furthermore, random permutation testing demonstrated that at least 98.92% of these DEPs were not generated by chance.

Several identified DEPs have been previously reported to be associated with thalassemia or its therapeutic mechanisms. For instance, TFR2 has been identified as a novel potential therapeutic target for β-thalassemia, while SMAD3 is linked to the mechanism of action of luspatercept, a drug recently approved by the U.

S. Food and Drug Administration (FDA) for the treatment of transfusion-dependent thalassemia. Our analysis also revealed enrichment in multiple biological pathways previously documented in thalassemia research. These findings suggest that the urinary proteome holds significant potential to provide references and clues for identifying novel therapeutic targets. Additionally, several significantly altered proteins and biological pathways identified in this study have not yet been linked to thalassemia in existing literature, potentially offering new insights for the research and optimization of diagnostic and therapeutic strategies for transfusion-dependent thalassemia.

关键词

Urine; Proteomics; Transfusion-Dependent Thalassemia

Urine Proteome Analysis in Patients with Transfusion-Dependent Thalassemia

Yuzhen Chen[1†], Weijie Chen[2†], Xinhua Zhang[3], Li Wang[4], Binbin Huang[5], Yuhua Ye[2*], Xiangmin Xu[2*], Youhe Gao[1*]

[1] Gene Engineering Drug and Biotechnology Beijing Key Laboratory, College of Life Sciences, Beijing Normal University, Beijing 100875, China [2] Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, Guangdong, China. [3] Department of Hematology, 923rd Hospital of the People' s Liberation Army, Nanning 530021, Guangxi, China. [4] Department of Pediatrics, 923rd Hospital of the People' s Liberation Army, Nanning 530021, Guangxi, China. [5] Department of Internal Medicine, Sixth People' s Hospital of Nanning, Nanning 530022, Guangxi, China.

[†] These authors contributed equally to this work. Supported by Beijing Normal University (11100704).

- Corresponding authors: Youhe Gao, E-mail: gaoyouhe@bnu.edu.cn. Xiangmin Xu, E-mail: xixm@smu.edu.cn. Yuhua Ye, E-mail: yeyuhua_{genetics}@163.com.

Abstract

Thalassemia is a hereditary hemoglobinopathy that poses a serious threat to human health and is particularly prevalent in southern China. At present, its clinical treatment still faces many limitations. In this study, the effects of thalassemia on the body were first meticulously explored via the urine proteome. Urinary proteins from patients with transfusion-dependent thalassemia (TDT) were compared with those from healthy controls. A total of 816 differential proteins were identified, among which 19 exhibited changes from presence to absence or vice versa. Additionally, results from a randomized grouping test indicated that at least 98.92% of these proteins were not randomly generated. Several of these proteins have been reported to be associated with thalassemia or its therapeutic mechanisms. For example, TFR2 has been reported as a novel

potential therapeutic target for β-thalassemia, and SMAD3 is associated with luspatercept, a drug approved by the U.

S. Food and Drug Administration for the treatment of TDT. Moreover, multiple biological pathways enriched by these differential proteins have been reported to be related to thalassemia, suggesting that the urine proteome has the potential to provide references and clues for identifying potential therapeutic targets of thalassemia. Furthermore, some significantly changed proteins and enriched biological pathways, not yet reported to be related to thalassemia, may provide clues for the research and optimization of diagnosis and treatment strategies of TDT.

Keywords

urine; proteomics; transfusion-dependent thalassemia

1 引言

Thalassemia is the most widely distributed monogenic inherited disease globally, affecting the largest number of people. Its pathogenesis lies in the imbalance of globin chain production, which leads to ineffective erythropoiesis, increased hemolysis, and disordered iron homeostasis. It is estimated that 1.5% of the global population carries the β-thalassemia allele, while 5% carry the α-thalassemia allele, with over 90% of affected individuals distributed across tropical and subtropical regions [**?**]. In China, approximately 30 million people carry thalassemia-related mutations, of whom about 300,000 have major or intermediate thalassemia requiring medical intervention [**?**].

Based on transfusion requirements and clinical characteristics, patients can be classified into transfusion-dependent thalassemia (TDT) and non-transfusion-dependent thalassemia (NTDT). This clinical classification has been widely applied in clinical practice and research. TDT refers to patients who require lifelong blood transfusions to sustain life; treatment for these patients necessitates maintaining pre-transfusion hemoglobin levels at 9–10 g/dL while managing secondary iron overload. NTDT refers to patients who do not depend on regular transfusions for survival, although they may occasionally require transfusions during special circumstances—such as surgery, pregnancy, or infection—where hemoglobin levels may drop acutely or sub-acutely. They may also receive short-term transfusions to improve anemia-related issues. In NTDT, iron overload stems from increased intestinal iron absorption caused by ineffective erythropoiesis [**?**].

Thalassemia is a preventable but difficult-to-treat hereditary disease. Implementing premarital, preconception, and prenatal screening is the primary measure for effective prevention and control. Regarding treatment, the focus is on managing chronic anemia while closely monitoring and addressing iron overload, as well as complications related to the disease or its treatment. Most patients re-

ceive conservative treatment regimens consisting of blood transfusions and iron chelation therapy [**?**]. However, challenges remain, including high treatment costs, complex resource allocation, and poor patient

compliance [**?**], and these methods do not provide a fundamental cure for the disease. Furthermore, due to limitations such as the lack of suitable donors, age, health status, and treatment costs, only a small minority of patients undergo curative treatments such as allogeneic hematopoietic stem cell transplantation or gene therapy [**?**, **?**, **?**]. Therefore, in-depth exploration of diagnosis and treatment strategies is of great significance for improving patient quality of life and reducing the disease burden.

Proteomics reveals the composition and variation of proteins within cells or organisms by analyzing protein structure, expression, post-translational modifications, and protein-protein interactions [**?**]. Urine is a filtrate of blood that does not require and is not subject to the strict regulation of homeostatic mechanisms. Consequently, it can accommodate and accumulate more significant changes without harming the body, reflecting changes in all organs and systems earlier and more sensitively. This makes it an excellent source for biomarker discovery [**?**]. This study aims to leverage the advantages of urine in comprehensively, systematically, and sensitively reflecting the physiological state of the body [**?**]. By analyzing the urinary proteome of TDT patients, we seek to explore the systemic impact of the disease in detail, thereby providing clues for optimizing TDT diagnosis and treatment protocols.

2.1 尿液样品的收集与处理

In this study, urine samples were collected from 29 patients with transfusion-dependent thalassemia (TDT) aged 3–18 years (Ethics Review Approval No.: Guokeyiban Shenzi (2021) CJ1028) and stored in a freezer at −80°C. Mass spectrometry data for 25 healthy control children (aged 2–15 years) were obtained from a previous study conducted by our laboratory [**?**].

The collected urine samples were centrifuged at 12,000 ×g for 30 min at 4°C, and the resulting supernatant was transferred to new centrifuge tubes. Dithiothreitol (DTT, Sigma) was added to a final concentration of 20 mmol/L. After vortexing, the mixture was heated in a metal bath at 37°C for 1 h and then cooled to room temperature. Subsequently, iodoacetamide (IAA, Sigma) was added to a final concentration of 50 mmol/L; the mixture was vortexed, briefly centrifuged, and allowed to react in the dark at room temperature for 40 min. Pre-cooled absolute ethanol (four times the volume of the supernatant) was added, and after thorough mixing, the proteins were precipitated at −20°C for 36

h. Following centrifugation at 10,000 ×g for 30 min at 4°C, the supernatant was discarded. The protein pellet was resuspended in an appropriate volume of lysis buffer (8 mol/L urea, 2 mol/L thiourea, 25 mmol/L DTT, and 50 mmol/L Tris) to obtain the urinary protein extract. Protein concentration was determined using the Bradford assay.

Using the filter-aided sample preparation (FASP) method [**?**], 100 μg of urinary protein sample was transferred to a 1.5 mL centrifuge tube, and UA solution (8 mol/L urea, 0.1 mol/L Tris-HCl, pH

8.

5) was added to reach a total volume of 200 μL. A 10 kD ultrafiltration tube (Pall, Port Washington, NY, USA) was pre-washed twice by adding 200 μL of UA solution to the filter membrane and centrifuging at 14,000 $\times$g for 10 min at 18°C. The processed protein sample was then loaded and centrifuged at 14,000 $\times$g for 40 min at 18°C. Subsequently, 200 μL of UA solution was added, vortexed, and centrifuged at 14,000 $\times$g for 40 min at 18°C; the filtrate was discarded, and this step was repeated once. The membrane was then washed twice with 25 mmol/L NH_4HCO_3 solution, vortexed, and centrifuged at 14,000 $\times$g for 40 min at 18°C, with the filtrate discarded each time. The collection tube was replaced, and 100 μL of NH_4HCO_3 solution was added. Trypsin (Trypsin Gold, Promega, Madison, WI, USA) was added at a mass ratio of 1:50 (trypsin:protein) for digestion. After overnight incubation in a 37°C water bath, the peptide solution was collected. Finally, the peptides were desalted using an HLB column (Waters, Milford, MA, USA), dried using a vacuum concentrator, and stored at $-$80°C.

2.2 液相色谱-串联质谱分析

The enzymatically digested samples were reconstituted in 0.1% formic acid, and the resulting peptides were quantified using a BCA kit. The peptide concentration was then adjusted to 0.5 μg/μL. A pooled peptide sample was prepared by taking 6.2 μL from each individual sample and fractionated using a High pH Reversed-Phase Peptide Fractionation Kit (Thermo Fisher Scientific, Rockford, IL, USA). Ten fractions were collected, dried using a vacuum concentrator, and subsequently reconstituted in 0.1% formic acid. Finally, iRT reagent (Biognosis, Schlieren, Switzerland) was added to the samples at a volume ratio of 10:1 (sample:iRT).

To generate the spectral library, mass spectrometry data for the ten collected fractions were acquired in Data Dependent Acquisition (DDA) mode. For each fraction, 1 μg of the sample was separated using an EASY-nLC

1200 chromatography system (Thermo Fisher Scientific, Waltham, MA, USA). Samples were loaded onto a C18 reversed-phase trap column (75 μm $\times$ 2 cm, 3 μm) and a reversed-phase analytical column (75 μm $\times$ 25 cm, 2 μm) at a flow rate of 0.3 μL/min. Gradient elution was performed over 90 minutes using mobile phase A (0.1% formic acid) and mobile phase B (80% acetonitrile + 0.1% formic acid).

Mass spectrometry data were analyzed and acquired using an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, Waltham, MA,

USA). The MS operating parameters were configured as follows: a spray voltage of 2.25 kV; a primary full scan (MS1) performed across a range of 350–1550 m/z at a resolution of 120,000; and MS/MS scans acquired in Orbitrap mode at a resolution of 30,000 with a Higher-energy Collisional Dissociation (HCD) energy of 30%. The top 20 precursor ions with the highest intensity were selected for fragmentation, with a dynamic exclusion duration of 30 seconds. The DDA acquisition results were imported into Proteome Discoverer software (version 2.1, Thermo Fisher Scientific) for database searching. The PD search results were then utilized to establish the Data Independent Acquisition (DIA) method, with window widths and quantities calculated based on the m/z distribution density.

Mass spectrometry data for individual samples were acquired in DIA mode, with each sample analyzed in triplicate. The liquid chromatography settings for the DIA mode were identical to those used for the DDA mode. The MS operating parameters were set as follows: a spray voltage of 2.3 kV; a primary full scan (MS1) across a range of 350–1500 m/z at a resolution of 60,000; and MS/MS scans acquired in Orbitrap mode across a range of 200–2000 m/z at a resolution of 30,000 with an HCD energy of 32%. To ensure quality control throughout the analytical process, a single DIA analysis of the pooled peptides was performed every 16 injections.

2.3 数据处理

The DIA acquisition results were imported into Spectronaut Pulsar software (version 19, Biognosys AG, Schlieren, Switzerland) for data processing and analysis. Peptide intensities were calculated by summing the peak areas of the respective MS2 fragment ions, and protein intensities were subsequently determined by aggregating the intensities of their constituent peptides. For statistical analysis, the mean values from three technical replicates per sample were utilized. Protein identification was performed using a threshold of at least two unique peptides per protein and a protein-level false discovery rate (FDR) of less than 1%.

2.4 数据分析

Proteins identified in the TDT group were compared with those in the healthy control group. Differentially expressed proteins (DEPs) were screened based on a fold change (FC) ≥ 2 or ≤ 0.5, with statistical significance determined by a two-tailed unpaired t-test ($p < 0.01$).

Hierarchical cluster analysis (HCA) was performed using the Wukong platform (https://www.omicsolution.com/wkomics/main/). Biological analyses were conducted using the UniProt website (https://www.uniprot.org/) and the DAVID database (https://david.ncifcrf.gov/). Functional analysis of the differentially expressed proteins was further supported by a literature search in the PubMed database (https://pubmed.ncbi.nlm.nih.gov/).

3.1 尿液蛋白质鉴定情况

LC-MS/MS analysis was performed on a total of 54 samples from the TDT (Thalassemia Major) group and the healthy control group. Based on the criteria of at least two specific peptides per protein and a protein-level false discovery rate (FDR) of less than 1%, a total of 3,424 proteins were identified. By comparing the urinary proteins of TDT patients with those of healthy controls using the criteria of a fold change (FC) ≥ 2 or ≤ 0.5 and a two-tailed unpaired t-test $p < 0.01$, 816 differentially expressed proteins (DEPs) were identified. Detailed information regarding these DEPs is provided in Supplementary Table 1 .

Among these, 19 DEPs exhibited an "all-or-nothing" expression pattern (Table 1). Specifically, "presence-to-absence" refers to proteins that were identified in the urine samples of the healthy control group but were not detected in any of the 29 samples from the TDT group. Conversely, "absence-to-presence" refers to proteins that were identified in the urine samples of the TDT group but were not detected in any of the 25 samples from the healthy control group.

UniProt ID

Protein name

P-value

Q8NI35 Q9Y5G0 Q9Y5F0 P62701 Q9NQC3 Q9Y219 Q9NVA2 Q02750 P84022 O75110 O43301 Q96FL9 Q9BXB1

InaD-like protein Protocadherin gamma-B5 Protocadherin beta-13 Small ribosomal subunit protein eS4, X isoform Reticulon-4 Protein jagged-2 Septin-11 Dual specificity mitogen-activated protein kinase kinase 1 Mothers against decapentaplegic homolog 3 Probable phospholipid-transporting ATPase IIA Heat shock 70 kDa protein 12A Polypeptide N-acetylgalactosaminyltransferase 14 Leucine-rich repeat-containing G-protein coupled receptor 4 Leucine-rich repeat and fibronectin type III domain-containing protein 1 Semaphorin-3B D-amino-acid oxidase Long-chain specific acyl-CoA dehydrogenase, mitochondrial Alpha-ketoglutarate dehydrogenase component 4 Nucleoside diphosphate kinase 3

8.83$\times10 - 051.53\times10 - 048.54\times10 - 049.54\times10 - 049.91\times10 - 041.25\times10 - 031.46\times10 - 031.54\times10 - 031.65\times10 - 032.54\times10 - 032.58\times10 - 035.07\times10 - 036.49\times$10-03

6.74$\times$10-03

8.10$\times10 - 037.79\times10 - 038.54\times10 - 038.87\times10 - 039.19\times$10-03

Q9P244 Q13214 P14920 P28330 P82909 Q13232

Hierarchical clustering analysis was performed on the total identified proteins [Figure 1: see original paper], which effectively distinguished the TDT group from the healthy control group.

3.2 随机分组检验

To determine the probability that the identified differentially expressed proteins (DEPs) were generated by chance, we performed a permutation test using the total proteins identified in the comparison between the Thalassemia Major (TDT) group and the healthy control group. Due to the computational constraints of the laboratory hardware, it was not feasible to perform a permutation test on the entire sample set. Consequently, we developed a script to randomly select 16 samples from the healthy group and 16 samples from the TDT group. This sample size represents the maximum capacity supported by our current computational resources, ensuring that we obtained the most reliable data possible within existing technical constraints.

The 32 selected samples were shuffled and randomly reassigned into two new groups. This resulted in a total of 300,540,195 possible combinations. Differential expression analysis was then conducted for these combinations using the same criteria as the original study ($FC \geq 2$ or ≤ 0.5, $p < 0.01$). To further accommodate computational limits, we sampled the total combinations at an interval of 100,000, resulting in 3,006 specific combinations for which the number of randomly generated DEPs was calculated.

This entire sample extraction and permutation process was repeated 10 times. After each iteration, the proportion of credible DEPs (those not arising by chance) was calculated, and the final average was determined. The results indicate that at least 98.92% of the identified DEPs were not generated by chance, with a standard deviation of 0.1% .

Sampling group

Number of differential proteins

Number of differential proteins generated randomly

Percentage of reliable differential proteins

98.79% 98.99% 99.07% 98.94% 98.76% 98.88% 98.94% 99.03% 98.87% 98.90%

3.3 差异蛋白分析

Functional analysis of the differentially expressed proteins (DEPs) was conducted using the UniProt website and the PubMed database, revealing several proteins previously reported to be associated with thalassemia. The DEP transferrin receptor protein 2 (TfR2) (FC = 2.70, $p = 1.85 \times 10^{-03}$) has been reported to regulate erythropoiesis based on systemic iron availability. The underlying mechanism likely involves modulating the sensitivity of erythroid progenitor cells to erythropoietin (EPO), thereby maintaining a balance between erythropoiesis and available iron stores [**?**]. TFR2 has also been identified as a novel potential therapeutic target for β-thalassemia [**?**, **?**]. Notably, TFR2-targeted therapy enhances EPO-mediated effects specifically within erythroid cells, offering an advantage over erythropoietin-stimulating agents and representing a

potential treatment option for other diseases characterized by anemia and/or ineffective erythropoiesis [**?**]. Additionally, the DEPs CD81 antigen (FC = 0.12, $p = 1.64 \times 10^{-08}$) and transferrin receptor-trafficking protein (FC = 0.05, $p = 4.19 \times 10^{-03}$) were identified in this study.

The CD81 antigen possesses functions such as transferrin receptor binding. Transferrin receptor 1 (TfR1) plays a critical role in cellular iron uptake; furthermore, hepatic TfR1 can interact with HFE to regulate hepcidin production, thereby influencing systemic iron homeostasis [**?**]. Other studies have demonstrated that hepatic TfR1 promotes hepcidin suppression and subsequent iron overload in β-thalassemia [**?**]. Additionally, TfR1 is significantly upregulated in patients with β-thalassemia [**?**].

Mothers against decapentaplegic homolog 3 (SMAD3) exhibited an "all-or-nothing" expression pattern among the identified DEPs (FC = 0, $p = 1.65 \times 10^{-03}$). It was detected in 18 out of 25 samples in the healthy control group but was not identified in any of the 29 samples in the transfusion-dependent thalassemia (TDT) group. SMAD3 is a key molecule in the TGF-β signaling pathway [**?**].

SMAD3 is involved in various biological processes, including cellular responses to TGF-β stimulation, positive regulation of TGF-β3 and TGF-β2 production, the TGF-β receptor signaling pathway, and SMAD protein signal transduction. Its molecular functions include TGF-β receptor binding, formation of SMAD protein complexes, and R-SMAD binding. Ligands of the TGF-β superfamily have been confirmed to act as inhibitors of late-stage erythropoiesis [**?**, **?**, **?**]. Research indicates that luspatercept and sotatercept can improve anemia caused by diseases characterized by ineffective erythropoiesis, including β-thalassemia, by reducing SMAD2/3 signaling capacity [**?**]. Among these, luspatercept has been approved by the U.

S. Food and Drug Administration (FDA) for the treatment of TDT [**?**]. This drug significantly reduces the transfusion burden in TDT patients [**?**] and increases hemoglobin levels in non-transfusion-dependent thalassemia (NTDT) patients [**?**].

The DEP mitogen-activated protein kinase 14 (MAP kinase p38 alpha) was identified (FC = 6.08, $p = 9.49 \times 10^{-04}$). Studies have shown that the traditional Chinese medicine *Carapax Testudinis* (tortoise shell) is a potential novel drug for treating β-thalassemia. It activates the p38 MAPK signaling pathway, mediating histone epigenetic modifications in the γ-globin gene promoter region, which in turn activates the fetal hemoglobin (HbF) gene [**?**]. Furthermore, other research suggests that thalidomide induces γ-globin gene expression during adult erythropoiesis through reactive oxygen species (ROS)-mediated activation of the p38 MAPK signaling pathway and histone H4 acetylation [**?**].

The DEPs vascular cell adhesion protein 1 (VCAM-1) (FC = 8.24, $p = 3.45 \times 10^{-08}$) and thrombomodulin (FC = 0.22, $p = 2.74 \times 10^{-05}$) were also identified.

Previous studies have reported elevated levels of VCAM-1 and thrombomodulin in the serum and plasma of thalassemia patients [**?**].

The DEP hemopexin (FC = 0.49, $p = 6.16 \times 10^{-03}$) is involved in biological processes such as hemoglobin metabolism, heme metabolism, heme transport, and intracellular iron ion homeostasis. It possesses molecular functions including heme transmembrane transporter activity and metal ion binding. Research has shown that hemopexin is significantly downregulated in patients with β-thalassemia. Additionally, the DEP clusterin (FC = 0.50, $p = 2.29 \times 10^{-04}$) was identified; clusterin in plasma extracellular vesicles has been reported to have significant potential for the diagnosis of β-thalassemia [**?**].

The DEP catalase (FC = 3.21, $p = 1.05 \times 10^{-05}$) is involved in biological processes such as hemoglobin metabolism, aerobic respiration, and response to hypoxia. Its functions include heme binding, metal ion binding, antioxidant activity, catalase activity, peroxidase receptor oxidoreductase activity, and NADP binding.

The DEP PGI2-stimulating factor (FC = 0.37, $p = 2.93 \times 10^{-07}$) was identified. Studies have indicated that levels of PGI2 metabolites are elevated in the urine of thalassemia patients [**?**].

3.4 生物学通路分析

Biological process enrichment analysis of the identified differentially expressed proteins (DEPs) was performed using the DAVID database ([Figure 2: see original paper]A; detailed information is provided in Supplementary Table 2). These DEPs are primarily involved in biological processes such as cell adhesion, telomere organization, negative regulation of megakaryocyte differentiation, cell migration, angiogenesis, immune system processes, multivesicular body assembly, the multivesicular body sorting pathway, cell surface receptor signaling pathways, vascular endothelial growth factor (VEGF) signaling pathways, regulation of platelet activation, and blood coagulation. Research indicates that patients with thalassemia exist in a lifelong, chronic hypercoagulable state. The procoagulant effects of thalassemic erythrocytes may contribute to this hypercoagulable state by enhancing thrombin generation and initiating platelet activation [**?**].

Pathway enrichment analysis conducted via the Kyoto Encyclopedia of Genes and Genomes (KEGG) database ([Figure 2: see original paper]B; detailed information is provided in Supplementary Table

2) revealed significant enrichment in pathways including hematopoietic cell lineage, lysosome, cell adhesion molecules (CAMs), regulation of actin cytoskeleton, phagosome, efferocytosis, and complement and coagulation cascades.

A: Results of biological process enrichment analysis (Top 20; $p < 0.01$); B: Results of KEGG pathway enrichment analysis (Top 20; $p < 0.01$).

4 讨论

The clinical manifestations of thalassemia exhibit high heterogeneity, ranging from mild cases that are nearly asymptomatic and free of complications to severe forms requiring lifelong blood transfusion support. This study is the first to utilize urinary proteomics to meticulously explore the systemic impact of thalassemia. By comparing the urinary proteins of patients with transfusion-dependent thalassemia (TDT) and healthy controls, we identified several differentially expressed proteins (DEPs) previously reported to be associated with thalassemia or its therapeutic mechanisms. Furthermore, multiple biological pathways known to be linked to thalassemia were significantly enriched.

In addition, this study collected urine samples from three patients with non-transfusion-dependent thalassemia (NTDT). Comparing the urinary proteins of NTDT patients with those of healthy controls—using criteria of fold change (FC) ≥ 2 or ≤ 0.5 and a two-tailed unpaired t-test $p < 0.01$—we identified a total of 602 DEPs. Detailed information is provided in . Hierarchical clustering analysis of these identified DEPs (Figure 3 [Figure 3: see original paper]) clearly distinguished the samples of the NTDT group from those of the healthy control group.

Similarly, to determine the probability that the identified DEPs were generated by chance, a permutation test was performed on the total proteins identified in the comparison between the NTDT and healthy groups. The results indicate that at least 82.86% of the DEPs were not generated randomly.

Biological process enrichment analysis of the identified DEPs was performed using the DAVID database (Figure 4 [Figure 4: see original paper]A; detailed information in). These DEPs are primarily involved in biological processes such as cell adhesion, multivesicular body (MVB) assembly, the MVB sorting pathway, cell migration, angiogenesis, cell surface receptor protein tyrosine kinase signaling pathways, regulation of MVB-lysosome fusion, platelet activation, blood coagulation, and cellular response to TGF-β stimulation. KEGG pathway enrichment analysis (Figure 4B; detailed information in) showed significant enrichment in endocytosis, the PI3K-Akt signaling pathway, hematopoietic cell lineage, regulation of the actin cytoskeleton, cell adhesion molecules (CAMs), complement and coagulation cascades, phagosomes, efferocytosis, and the MAPK signaling pathway.

A: Results of biological process enrichment analysis (Top 20; $p < 0.01$); B: Results of KEGG pathway enrichment analysis (Top 20; $p < 0.01$).

A Venn diagram was used to illustrate the overlap of DEPs identified in the TDT vs. healthy control comparison and the NTDT vs. healthy control comparison (Figure 5 [Figure 5: see original paper]A). A total of 398 DEPs were commonly identified, including proteins previously reported to be associated with thalassemia, such as SMAD3, transferrin receptor-trafficking protein, CD81 antigen, thrombomodulin, and PGI2-stimulating factor. Notably, SMAD3—which

is associated with the mechanism of luspatercept, a drug approved by the U.

S. Food and Drug Administration for the treatment of TDT—showed a complete disappearance (from presence to absence) in both the TDT and NTDT groups compared to healthy controls. Venn diagrams also displayed the overlap of biological processes and enriched KEGG pathways between the two comparison groups (Figure 5B, C). These results collectively demonstrate that the urinary proteome can reflect the systemic effects of thalassemia and distinguish between TDT and NTDT. However, this study has certain limitations: the sample size for NTDT patients was small, which poses a risk of statistical bias in the NTDT vs. healthy control comparison. Additionally, the healthy control group consisted entirely of males. These findings require further validation through large-scale clinical studies with well-matched cohorts.

A: Differentially expressed proteins (FC ≥ 2 or ≤ 0.5, $p < 0.01$); B: Biological processes ($p < 0.01$); C: KEGG pathways ($p < 0.01$).

In summary, the use of the urinary proteome facilitates a comprehensive and systematic exploration of the systemic impact of thalassemia. The urinary proteome has the potential to provide clues for identifying latent therapeutic targets and offers a foundation and reference for related research and the optimization of diagnosis and treatment strategies.

参考文献

[1] Kattamis A, Kwiatkowski JL, Aydinok

Y. Thalassaemia. Lancet. 2022;399(10343):2310-2324. doi:10.1016/s0140-6736(22)00536-0. [2] Wang WD, Hu F, Zhou DH, Gale RP, Lai YR, Yao HX, Li C, Wu BY, Chen Z, Fang JP, Chen SJ, Liang

Y. Thalassaemia in China. Blood Rev. 2023;60:

101074. doi:10.1016/j.blre.

101075.

101076. [3] Musallam KM, Cappellini MD, Viprakasit V, Kattamis A, Rivella S, Taher AT. Revisiting the nontransfusion-dependent (NTDT) vs. transfusion-dependent (TDT) thalassemia classification 10 years later.

Am J Hematol. 2021;96(2):E54-e56. doi:10.1002/ajh.

26056. [4] El-Beshlawy A, El-Ghamrawy

M. Recent trends in treatment of thalassemia. Blood Cells Mol Dis. 2019;76:53-58. doi:10.1016/j.bcmd.

2019.

2020.

2021. [5] Taher AT, Weatherall DJ, Cappellini MD. Thalassaemia. Lancet. 2018;391(10116):155-167. doi:10.1016/s0140-6736(17)31822-6. [6] Motta

I, Bou-Fakhredin R, Taher AT, Cappellini MD. Beta Thalassemia: New Therapeutic Options Beyond Transfusion and Iron Chelation. Drugs. 2020;80(11):1053-1063. doi:10.1007/s40265-020-

[7] Manzoni C, Kia DA, Vandrovcova J, Hardy J, Wood NW, Lewis PA, Ferrari

R. Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences. Brief Bioinform. 2018;19(2):286-302. doi:10.1093/bib/bbw114.

[8] Gao

Y. Urine—an untapped goldmine for biomarker discovery? Sci China Life Sci. 2013;56(12):1145-1146. doi:10.1007/s11427-013-4574-1.

[9] Gao YH. Personalized medicine is based on urinary omics data [J]. Journal of Chinese Mass Spectrometry Society, 2025, 46(05): 531.

[10] Meng WS. Application of urinary proteome in early diagnosis, disease progression monitoring, and exercise monitoring [D]. Beijing: Beijing Normal University, 2022.

[11] Wiśniewski JR, Zougman A, Nagaraj N, Mann

M. Universal sample preparation method for proteome analysis. Nat Methods. 2009;6(5):359-362. doi:10.1038/nmeth.1322.

[12] Nai A, Lidonnici MR, Rausa M, Mandelli G, Pagani A, Silvestri L, Ferrari G, Camaschella

C. The second transferrin receptor regulates red blood cell production in mice. Blood. 2015;125(7):1170-1179.

doi:10.1182/blood-2014-08-596254. [13] Di Modica SM, Tanzi E, Olivari V, Lidonnici MR, Pettinato M, Pagani A, Tiboni F, Furiosi V, Silvestri L, Ferrari G, Rivella S, Nai

A. Transferrin receptor 2 (Tfr2) genetic deletion makes transfusionindependent a murine model of transfusion-dependent β-thalassemia. Am J Hematol. 2022;97(10):

13241336. doi:10.1002/ajh.

13241337. [14] Artuso I, Lidonnici MR, Altamura S, Mandelli G, Pettinato M, Muckenthaler MU, Silvestri L, Ferrari G, Camaschella C, Nai

A. Transferrin receptor 2 is a potential novel therapeutic target for βthalassemia: evidence from a murine model. Blood. 2018;132(21):2286-2297. doi:10.1182/blood-201805-852277. [15] Schmidt PJ, Toran PT, Giannetti AM, Bjorkman PJ, Andrews NC. The transferrin receptor modulates Hfe-dependent regulation of hepcidin expression. Cell Metab. 2008;7(3):205-214. doi:10.1016/j.cmet.

2007.

2008.
2009. [16] Xiao X, Moschetta GA, Xu Y, Fisher AL, Alfaro-Magallanes VM, Dev S, Wang CY, Babitt JL.

Regulation of iron homeostasis by hepatocyte TfR1 requires HFE and contributes to hepcidin suppression in β-thalassemia. Blood. 2023;141(4):422-432. doi:10.1182/blood.

2022017811. [17] Li N, An P, Wang J, Zhang T, Qing X, Wu B, Sun L, Ding X, Niu L, Xie Z, Zhang M, Guo X, Chen X, Cai T, Luo J, Wang F, Yang

F. Plasma proteome profiling combined with clinical and genetic features reveals the pathophysiological characteristics of β-thalassemia. iScience. 2022;25(4):

104091. doi:10.1016/j.isci.

104092.

104093. [18] Hata A, Chen YG. TGF-β Signaling from Receptors to Smads. Cold Spring Harb Perspect Biol. 2016;8(9). doi:10.1101/cshperspect.a022061. [19] Söderberg SS, Karlsson G, Karlsson

S. Complex and context dependent regulation of hematopoiesis by TGF-beta superfamily signaling. Ann N Y Acad Sci. 2009;1176:55-69. doi:10.1111/j.17496632.2009.04569.

x. [20] Blank U, Karlsson

S. The role of Smad signaling in hematopoiesis and translational hematology.

Leukemia. 2011;25(9):1379-1388. doi:10.1038/leu.

2011.
2012. [21] U.S.

Administration.

Reblozyl®

information[EB/OL].

(luspatercept-aamt)

prescribing

(2020-04)[2026-03-10].

https://www.accessdata.fda.gov/drugsatfda_{docs}/label/2020/761136orig2lbl.pdf. [22] Cappellini MD, Viprakasit V, Taher AT, Georgiev P, Kuo KHM, Coates T, Voskaridou E, Liew HK, Pazgal-Kobrowski I, Forni GL, Perrotta S, Khelif A, Lal A, Kattamis A, Vlachaki E, Origa R, Aydinok

Y, Bejaoui M, Ho PJ, Chew LP, Bee PC, Lim SM, Lu MY, Tantiworawit A, Ganeva P, Gercheva L, Shah F, Neufeld EJ, Thompson A, Laadem A, Shetty

JK, Zou J, Zhang J, Miteva D, Zinger T, Linde PG, Sherman ML, Hermine O, Porter J, Piga

A. A Phase 3 Trial of Luspatercept in Patients with TransfusionDependent β-Thalassemia. N Engl J Med. 2020;382(13):1219-1231. doi:10.1056/NEJMoa1910182. [23] Taher AT, Viprakasit V, Kattamis A, Perrotta S, Ricchi P, Porter JB, Coates TD, Forni GL, Musallam KM, Esposito O, Pilot R, Kuo WL, Lai Y, Reverte M, Wei R, Bueno LM, Cappellini MD. Efficacy and safety of luspatercept in non-transfusion-dependent β-thalassemia: long-term results from the BEYOND study. Blood Adv. 2025;9(23):6108-6119. doi:10.1182/bloodadvances.

2025016554. [24] Qian X, Chen J, Zhao D, Guo L, Qian

X. Plastrum testudinis induces γ-globin gene expression through epigenetic histone modifications within the γ-globin gene promoter via activation of the p38 MAPK signaling pathway. Int J Mol Med. 2013;31(6):1418-1428. doi:10.3892/ijmm.

2013.

2014. [25] Aerbajinai W, Zhu J, Gao Z, Chin K, Rodgers GP. Thalidomide induces gamma-globin gene expression through increased reactive oxygen species-mediated p38 MAPK signaling and histone H4

acetylation in adult erythropoiesis. Blood. 2007;110(8):2864-2871. doi:10.1182/blood-2007-01-065201. [26] Eldor A, Rachmilewitz EA. The hypercoagulable state in thalassemia. Blood. 2002;99(1):36-43. doi:10.1182/blood.v99.

1.

2. [27] Li N, Wu B, Wang J, Yan Y, An P, Li Y, Liu Y, Hou Y, Qing X, Niu L, Ding X, Xie Z, Zhang M, Guo X, Chen X, Cai T, Luo J, Wang F, Yang

F. Differential proteomic patterns of plasma extracellular vesicles show potential to discriminate β-thalassemia subtypes. iScience. 2023;26(2):

106048. doi:10.1016/j.isci.

106049.

106050. [28] Eldor A, Lellouche F, Goldfarb A, Rachmilewitz EA, Maclouf

J. In vivo platelet activation in betathalassemia major reflected by increased platelet-thromboxane urinary metabolites. Blood. 1991;77(8):1749-1753.

Note: Figure translations are in progress. See original paper for figures.