Annals of Hematology (2019) 98:1813–1826
https://doi.org/10.1007/s00277-019-03710-1

ORIGINAL ARTICLE

Colla corii asini might upregulate ZNF471 and THOC5 by KRAB domain–containing zinc-finger protein pathway and THO complex subunit 5 pathway to improve anemia of pregnant women with β-thalassemia

Yanfang Li[1] · Zhanfeng Zhang[2] · Lilin Yang[3] · Xiangyi Li[1] · Jingwen Zhou[1] · Daocheng Li[1] · Songping Luo[4]

Received: 18 January 2019 / Accepted: 5 May 2019 / Published online: 16 May 2019

Abstract

Pregnant patients with β-thalassemia are more likely to have progressive anemia which expose them to risk of adverse pregnancy outcomes, blood transfusion, and iron overload. Results from our previous study indicated that *Colla corii asini* (CCA, E'jiao), a natural ingredient of traditional Chinese medicine, could significantly increase hemoglobin level of pregnant women with β-thalassemia, but the underlying molecular mechanism was unclear. Thus, we applied high-throughput transcriptome sequencing to study the transcriptomic change before and after the CCA treatment. Twenty eligible pregnant women were recruited and randomized to either the CCA treatment group or the blank control group in a 3:1 ratio. Patients in the treatment group orally received daily 15 g CCA powder for 4 weeks. We analyzed the therapeutic effect indexes and the transcriptomic change in subjects' peripheral blood before and after treatment. We found that β CD 41-42(-TTCT)/β^A was the main genotype of the subjects. The regulatory impact of CCA treatment became more evident among the subjects of genotype β CD 41-42(-TTCT)/β^A. Gene ontogenesis analysis revealed that the top five molecular functions of differentially expressed genes were involved in membrane functionality and cellular structure. We further identified two consistent upregulated genes ZNF471 and THOC5 in the effective treatment group, which were engaged in Kruppel-associated box (KRAB) domain-containing zinc-finger protein pathway and THOC5 pathway, respectively. Based on our current findings, we hypothesize that the anti-anemia effect of CCA on pregnant women with β-thalassemia might be related to translation regulation of spectrin synthesis, membrane stability, and eventually prolonged the life span of erythrocytes.

Keywords β-thalassemia · Anemia · Pregnancy · Transcriptome sequencing · *Colla corii asini*

Background

β-thalassemia is a common type of hemolytic anemia disease caused by mutations in the β-globin gene. Thalassemia is prevalent in provinces in southern China particularly in the population of Guangdong, Guangxi, and Yunnan. Epidemiological study showed that in Guangdong alone, approximately 17.83% of the 14,332 pregnant women across 21 regions examined were diagnosed as carriers of thalassemia, the prevalence of β- and α-globin gene mutations among the pregnant women were 4.53% and 13.31%, respectively [1].

Although β-thalassemia is not as common as α-thalassemia, previous studies revealed that as pregnancy progresses, women with β-thalassemia are more likely to have progressive anemia, especially between 24 and 32 gestational weeks. Approximately 60 to 80% of patients with β- thalassemia

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00277-019-03710-1) contains supplementary material, which is available to authorized users.

✉ Songping Luo
gynspluo@hotmail.com

[1] Department of Obstetrics, the First Affiliated Hospital, Guangzhou University of Chinese Medicine, Guangzhou, China

[2] Department of Clinical Laboratory, the First Affiliated Hospital, Guangzhou University of Chinese Medicine, Guangzhou, China

[3] Department of Gynecology, the First Affiliated Hospital, Guangzhou University of Chinese Medicine, Guangzhou, China

[4] First School of Clinical Medicine, Guangzhou University of Chinese Medicine, No.12 Jichang Road, Guangzhou 510405, Guangdong, China

require blood transfusion during pregnancy, of which 30% have no previous transfusion history. In spite of extended genotype and antibody screening as well as fully phenotype-matched blood given is recommended in the clinic, transfusion-related alloimmunization may still occur more frequently during gestation owing to the immune physiological characteristics of pregnancy. Lack of available extended matching transfusion technology to patients in developing regions also contributes to the risk of developing alloimmunization. In studies of pregnant patients with β-thalassemia, rates of alloimmunization are exceedingly high, ranging from 5~30% depending on the country where the study was performed [2]. The immune hemolysis will aggravate anemia in turn and require for repeated blood transfusion, and then further heighten the risk of transfusion-related infectious diseases and iron overload [3, 4].

Existing evidence have confirmed that compared with healthy pregnant women, thalassemia patients are associated with a wide range of adverse pregnancy outcomes which can severely impact the fetal growth and maternal health. Moreover, the severity of adverse pregnancy outcomes is closely related to a low level of hemoglobin (Hb) and high level of fetal hemoglobin (HbF) during pregnancy [3–5]. Therefore, it is recommended to maintain Hb level above 100 g/L during pregnancy [5].

Despite the progress in thalassemia management during the last three decades, blood transfusion remains the only available treatment. Also, no consensus has been reached in treating anemia of pregnant thalassemia patients. Early studies indicated that gene-targeting drugs like γ-chain activator and α-chain inhibitor improve hemolytic anemia by induction of HbF in patients with thalassemia. Yet both categories are potentially teratogenic, carcinogenic, and also bone marrow suppressive. These side effects render them unsuitable for pregnant patients [6].

Compared with conventional medicine, traditional Chinese medicine (TCM) has lower toxicity in certain prescriptions where natural ingredient instead of concentrated and synthetic compound is used. In TCM, *Colla corii asini* (CCA, E'jiao) is a gelatin-like preparation derived from donkey hide and has been widely used in anti-anemic therapy for more than 2000 years. Result from our previous study indicated that CCA treatment can significantly increase the levels of Hb and major adult hemoglobin (HbA) ($\alpha_2\beta_2$) without affecting the level of serum ferritin (SF) [7]. These observations lead to the question whether CCA improves anemia through induction of β-globin chain. To address this question, we applied high-throughput transcriptome sequencing to study the transcriptomic change in the peripheral blood cell of pregnant women with β-thalassemia before and after the CCA treatment. While the transcriptomic changes induced by CCA treatment in β-thalassemia patients was not substantial in general, we observed a correlation between treatment efficacy and a certain genotype in the subgroup analysis that might provide insight into the role of CCA treatment and its molecular mechanism in improving pregnant patients with β–thalassemia.

Materials and methods

Subjects

Pregnant women diagnosed of minor or intermediate β-thalassemia at the First Affiliated Hospital of Guangzhou University of Chinese Medicine between January 2018 and June 2018 were recruited as study subjects based on inclusion and exclusion criteria. Subjects and clinical characteristics for human study protocols were approved by the Institutional Medical Research Ethics Committee (Approval Letter Number: ZYYECK [2017] 074) of the First Affiliated Hospital, Guangzhou University of Chinese Medicine. Informed consent was obtained from all subjects before enrollment.

Diagnosis, inclusion, exclusion and withdrawal criteria

Diagnosis criteria

Patients were diagnosed as carriers of thalassemia based on genetic test. Classification of mild and intermediate β-thalassemia was based on the guidelines published by the Thalassemia International Federation in 2008 [8]. Mild anemia was classified as 60 g/L ≤ Hb < 110 g/L during pregnancy in accordance with the diagnostic criteria of the World Health Organization [9].

Inclusion criteria

1. Pregnant women diagnosed as thalassemia carriers by genetic test with clinical presentation of minor or intermediate β-thalassemia
2. Patients with mild anemia (70 g/L ≤ Hb < 100 g/L) prior to study enrollment
3. Singleton pregnancy
4. Patients who have not received blood transfusion or any forms of anti-anemia treatment in conventional medicine or TCM in the last 12 weeks
5. Obtained informed consent

Exclusion criteria

1. Patients with any of the following abnormalities: immunodeficiency, primary diseases involving cardiovascular

system, liver, kidney, gastrointestinal tract, endocrine system, and hematological system
2. Allergic to two or more drugs
3. Patients with mental illness or poor compliance with medical treatment

Withdrawal criteria

1. Patients who did not take drugs in compliance with the requirements of the study protocol
2. Patients who were lost to follow-up or exited the study voluntarily
3. Patients who took anti-anemia drugs and received blood transfusion during the study period

Treatment scheme

Once diagnosed as thalassemia and enrolled in the study, patients would be required to cease medications that contain iron and folic acid. In total, 20 patients were enrolled and randomized to either the treatment group or the control group in a 3:1 ratio by statistical package for social sciences (SPSS) 21.0 random number generator. Patients in the treatment group received daily dosage of 15 g oral CCA (Shandong Dong-E E-Jiao Co., Ltd.) in powder form for 4 consecutive weeks. The powder was dissolved in warm water and taken orally. Patients in control groups do not receive any intervention.

Blood sample collection

Ten milliliters of venous blood samples were drawn from 20 subjects after overnight fasting of 8 h for the following analysis: (1) Hb concentration measured by automated hematology analyzer XE5000 (Sysmex, Japan); (2) levels of adult hemoglobin (HbA), minor adult hemoglobin (HbA2), and fetal hemoglobin (HbF) in peripheral blood measured by automated capillary electrophoresis (Sebia, France); and (3) level of serum ferritin (SF) measured by full automatic immune analyzer I 2000 SR (Abbott, USA).

At the same time, 5 mL of venous blood samples were collected by EDTA tubes for further transcriptome analyses. The blood was first mixed by horizontal shaker and centrifuged at 1700 r/min for 10 min. After discarding the upper layer of the plasma, the lower layer of blood cells was topped up with 5 mL of saline and then centrifuged at 2500 r/min for 20 min. A mononuclear cell layer located between the plasma and the lymphocyte separation layer was transferred to a 15-mL centrifuge tube and topped up with 10 mL saline. The cell suspension was washed by centrifugation at 1800 r/min for 10 min twice. The supernatant was discarded after centrifuging and the remaining mononuclear cells were preserved in 1 mL Trizol (Invitrogen, USA) and stored frozen at − 80 °C refrigerator until RNA extraction.

RNA transcriptome analysis of peripheral blood mononuclear cells

RNA transcriptome sequencing analysis was performed by the Beijing Genomics Institute (BGI). The main processes of transcriptome resequencing study were described as follows.

Experiment workflow After total RNA extraction and DNase I treatment, Agilent 2100 Bioanalyzer (Agilent RNA 6000 Nano Kit) was used to do the total RNA sample quality control (QC). mRNAs were isolated from total RNA with oligo(dT) method. In a mixture with the fragmentation buffer, the mRNA was fragmented into 200–700-nt short fragments. The cDNA was then synthesized using mRNA fragments as templates. cDNA fragments are purified and resolved with elution buffer (EB) for end reparation and single-nucleotide A (adenine) addition. After that, the cDNA fragments are linked with adapters. Those cDNA fragments with suitable size are selected for the PCR amplification. Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System are used in quantification and qualification of those libraries.

Bioinformatics workflow Our primary sequencing data were produced using Illumina HiSeqTM and have been deposited in the NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/Traces/sra) under accession number SRP175968. These data, denoted as raw reads, were subjected to quality control (QC) to determine whether a resequencing step was needed. Specifically, the raw reads were cleaned by removal of reads containing primer/adapter sequences, low-quality reads (i.e., a percentage of low-quality bases greater than 50% in a read, with low-quality bases defined as bases with a sequencing quality of no more than 10), and reads with more than 10% unknown bases.

After filtering, the remaining reads were called "clean reads," which were aligned to the reference sequences with Bowtie2 [10] (Version: v2.2.5, Parameters: -q –phred64 –sensitive –dpad 0 –gbar 99999999 –mp 1,1 –np 1 –score-min L,0,-0.1 -I 1 -X 1000 –no-mixed –no-discordant -p 1 -k 200) and then calculated gene expression level with RSEM [11](Version: v1.2.12, Parameters: default).

Statistical analysis Statistical analysis was performed using SPSS 21.0 software (SPSS Inc., Chicago). All continuous variables were analyzed for normal distribution. The data that fulfilled normal distribution were presented as mean ± standard deviation and analyzed by *t* test, while the data of non-normal distribution were presented as median (quartile1-quartile3) and analyzed by Mann-Whitney *U* test.

Categorical variables were expressed as rates or proportions and analyzed with Fisher's exact test.

Differentially expressed genes (DEGs) were detected with DEseq2 as requested. DEseq2 is based on the negative binomial distribution, performed as described at Michael I et al. [12]. Fold change ≥ 2.00 and adjusted *P* value ≤ 0.05 were considered having significance.

Gene ontology analysis of DEG GO analysis allows functional association of DEGs using three structured networks of defined terms that describe gene product attributes. With the GO annotation result, we classify DEGs according to official classification, and we also perform GO functional enrichment using phyper, a function of R. See wiki for details (https://en.wikipedia.org/wiki/Hypergeometric_distribution). Then, we calculate false discovery rate (FDR) for each *P* value. In general, the terms which FDR not larger than 0.01 are defined as significant enriched. The *P* value cut-off was set at 0.05.

Pathway analysis of DEG KEGG pathway analysis for significant DEGs was also performed. With the KEGG annotation result, we classified DEGs according to official classification and performed pathway functional enrichment using phyper, a function of R. See wiki for details (https://en.wikipedia.org/wiki/Hypergeometric_distribution). Then, we calculate false discovery rate (FDR) for each *P* value, in general, the terms which FDR not larger than 0.01 are defined as significant enriched. The *P* value cut-off was set at 0.05.

Results

Patient characteristic

All 20 patients completed the study. However, 1 patient with genotype of β CD17(A-T)/β^A in the treatment group ceased treatment for 11 days due to contraction of a cold during the treatment period. Therefore, her result was excluded from RNA transcriptome analysis. In order to obtain more data information, we further conducted subgroup analysis based on the efficacy and genotype. Those patients in the treatment group whose Hb level increased greater than or equal to 5 g/L after treatment were assigned to the effective treatment group, while those with Hb level increased less than 5 g/L were assigned to the ineffective treatment group.

The predominant β-thalassemia mutation was β CD 41-42(-TTCT)/β^A expressed in 45.0% (9/20) of the participants. Among them, 44.4% (4/9) had responded to CCA treatment and were assigned to the effective treatment group. The rest of the participants was of genotype β IVS-II-654(C-T)/β^A (4/20), β-28(A-G)/β^A (2/20), β CD71-72(+A)/β^A (1/20), and β CD17(A-T)/β^A (2/20). All subjects were married Chinese Han female.

The comparison of baseline characteristics among different groups was shown in Tables 1 and 2. No significant differences were detected between the control and treatment groups in terms of age, preconception body mass index (BMI), gestational week at the time of enrollment, pregnancy history, and educational level as well as the baseline level of HbA, HbA2, HbF, and SF. However, when the effective and ineffective groups were further analyzed, it was shown that the gestational week at the time of enrollment and educational level were statistically different ($P = 0.002$ and $P = 0.028$, respectively). Pregnant women in the ineffective treatment group were likely to have anemia earlier and require treatment and tended to have higher educational level, compared with those in the effective treatment group. We speculated that this might be because the earlier anemia occurs and the more serious condition it is, the more difficult it is to overcome. We did not have an adequate explanation of the difference of educational level between the two groups, which requires future research with a larger sample size to verify.

Moreover, the baseline of Hb level was significantly lower in the treatment group than the one in the control group, but the basal Hb level had no significant difference between the effective treatment group and the ineffective treatment group. Because the enrollment was random and not subject to selection bias, the variability may be due to the difference in the number of participants in the control group and the treatment group.

Clinical therapeutic index

Next, we examined the red blood cell (RBC) count, Hb concentration, and the proportion of each Hb component before and after treatment. Data were shown in Table 3. After treatment, no increase of RBC count and Hb level were observed in the treatment group compared with the control group ($P = 0.413$ and $P = 0.527$). Significant change was observed after we further segregated the patients from treatment groups into effective group and ineffective group regardless of the genotype of β CD 41-42(-TTCT)/β^A ($P < 0.001$).

The significant increase of Hb level was consistently accompanied by a significant increase of RBC count. In the effective treatment group, both the RBC count and Hb level increased substantially after treatment while a marginal decrease of RBC count and a slow increase of Hb level were observed in the ineffective treatment group. Similar changes were observed in the thalassemia patients of genotype β –CD 41-42(-TTCT)/β^A.

In addition, no significant differences of three Hb components were detected before and after treatment. Similarly, SF levels were not affected by treatment regardless of the genotype.

Table 1 The characteristics of studied subjects (treatment group vs control group)

Items	Treatment group ($n = 15$)	Control group ($n = 5$)	Test value	P value
Age (year)[T]	30.60 ± 4.35	29.60 ± 4.72	0.436	0.668
Preconception BMI (kg/m^2)[T]	20.02 ± 1.81	20.51 ± 1.51	− 0.531	0.602
Gestational week at enrollment (weeks)[T]	24.53 ± 4.66	23.00 ± 5.79	0.602	0.555
Gravidity[F]				
≥ 3	5(33.33%)	2(40.00%)	/	0.558
2	7(46.67%)	1(20.00%)		
1	3(20.00%)	2(40.00%)		
Parity[F]				
≥ 1	8(53.33%)	2(40.00%)	/	1.000
0	7(46.67%)	3(60.00%)		
Number of induced abortions [F]				
≥ 1	7(46.67%)	3(60.00%)	/	1.000
0	8(53.33%)	2(40.00%)		
Number of spontaneous abortions[F]				
1	2(13.33%)	0(0.00%)	/	1.000
0	13(86.67%)	5(100.00%)		
Educational level[F]				
Secondary school	3(20.00%)	2(40.00%)	/	0.516
Junior college	7(46.67%)	3(60.00%)		
College	5(33.33%)	0(0.00%)		
Genotype[F]				
β CD 41-42(-TTCT)/β^A	9(60.00%)	2(40.00%)	/	0.681
β IVS-II-654(C–T)/β^A	3(20.00%)	1(20.00%)		
β-28(A-G)/β^A	1(6.67%)	1(20.00%)		
β CD71-72(+A)/β^A	1(6.67%)	0(0.00%)		
β CD17(A-T)/β^A	1(6.67%)	1(20.00%)		
RBC level before treatment (× 10^{12}/L)[T]	4.02 ± 0.46	4.53 ± 0.21	− 2.363	0.03*
Hb level before treatment (g/L)[T]	85.07 ± 7.71	93.20 ± 1.79	− 3.790	0.001**
HbA level before treatment (%)[M]	93.20(92.80, 94.00)	93.30(92.55, 94.95)	0.656	0.553
HbA2 level before treatment (%)[M]	5.00(4.90, 5.40)	5.40(5.05, 5.60)	1.563	0.142
HbF level before treatment (%)[T]	1.78 ± 1.32	1.00 ± 1.01	1.200	0.246
SF level before treatment (ng/mL)[M]	48.57(41.54,166.92)	34.15(22.64, 96.37)	− 1.091	0.306

RBC, red blood cell; *Hb*, hemoglobin; *HbA*, major adult hemoglobin; *HbA2*, secondary adult hemoglobin; *HbF*, fetal hemoglobin; *SF*, serum ferritin

The symbol asterisk indicates the significance level; *$P < 0.05$, **$P < 0.01$. The continuous variables marked with superscript letter "T" fulfilled the criteria of normal distribution and were presented as mean ± SD and analyzed by *t* test. The continuous variables marked with superscript letter "M" did not fulfill the criteria of the normal distribution and were presented as median (Q1–Q3) and analyzed by Mann-Whitney *U* test. The categorical variables marked with superscript letter "F" were expressed as *n* (%) and analyzed by Fisher's exact test

Gene expression analysis

Genome mapping

Forty samples of 20 study subjects were sequenced on Illumina HiSeq Platform and generated approximately 7.86 Gb per sample. The average genome mapping rate and the average gene mapping rate were 95.63% and 81.11%, respectively. 22,516 genes were identified, among which 19,842 of them are known genes and the remaining 2754 of them are novel genes. Given the uniformity of the mapping result for each sample, the samples are comparable.

The distribution of gene expression

The distribution of the gene expression level of each sample was analyzed and the dispersion of distribution, change of gene abundance, and concentration of gene expression in the sample interval were shown in box plot and density map, respectively (Fig. 1a, b). The result demonstrated that quality of gene expression in our study is consistent and comparable.

Differentially expressed gene detection

We used DEseq2 algorithms to detect the DEGs (differentially expression genes) between groups and results are shown in Fig. 2. Among all studied subjects, the transcriptome analysis revealed differential gene expression of three genes in the efficient treatment group after treatment (> 2-fold) while no DEGs was observed in the inefficient treatment group (Fig. 2a, b). For studied subjects of genotype β CD 41–42(-TTCT)/β^A, in a total of four genes were differentially expressed in the efficient treatment group after treatment (> 2-fold) while 46 genes with differential expression were found in the inefficient treatment group after treatment (Fig. 2c, d).

Table 2 The characteristics of studied subjects (effective treatment group vs Ineffective treatment group)

Items	Effective group ($n = 6$)	Ineffective group ($n = 9$)	Test value	P value
Age (year)^{T}	31.50 ± 4.04	30.00 ± 4.69	0.639	0.534
Preconception BMI (kg/m^2)^{M}	18.49(18.08, 21.58)	20.31(18.87, 22.03)	1.296	0.224
Gestational week at enrollment (weeks)^{T}	28.50 ± 2.59	21.89 ± 3.76	3.738	0.002**
Gravidity^{F}				
≥ 3	2(33.33%)	3(33.33%)	/	1.000
2	3(50.00%)	4(44.44%)		
1	1(16.67%)	2(22.22%)		
Parity^{F}				
≥ 1	5(83.33%)	3(33.33%)	/	0.119
0	1(16.67%)	6(66.67%)		
Number of induced abortions^{F}				
≥ 1	2(33.33%)	5(55.56%)	/	0.608
0	4(66.67%)	4(44.44%)		
Number of spontaneous abortions^{F}				
1	0(0.00%)	2(22.22%)	/	0.486
0	6(100.00%)	7(77.78%)		
Educational level^{F}				
Secondary school	1 (16.67%)	2(22.22%)	/	0.028*
Junior college	5(83.33%)	2(22.22%)		
College	0(0.00%)	5(55.56%)		
Genotype^{F}				
β CD 41–42(-TTCT)/β^{A}	4(66.67%)	5(55.56%)	/	1.000
β IVS-II-654(C-T)/β^{A}	2(33.33%)	1(11.11%)		
β-28(A-G)/β^{A}	0(0.00%)	1(11.11%)		
β CD71–72(+A)/β^{A}	0(0.00%)	1(11.11%)		
β CD17(A-T)/β^{A}	0(0.00%)	1(11.11%)		
RBC level before treatment ($\times 10^{12}$/L)^{T}	4.12 ± 0.56	3.96 ± 0.41	0.646	0.530
Hb level before treatment (g/L)^{T}	88.50 ± 8.76	82.78 ± 6.44	1.464	0.167
HbA level before treatment (%)^{T}	93.52 ± 0.78	92.78 ± 1.62	1.030	0.322
HbA2 level before treatment (%)^{M}	4.95(4.90, 5.10)	5.40(4.95, 5.50)	1.587	0.145
HbF level before treatment (%)^{T}	1.47 ± 0.74	1.99 ± 1.61	− 0.738	0.474
SF level before treatment (ng/mL)^{M}	44.75(40.62, 65.18)	88.33(32.31,185.61)	1.296	0.224

RBC, red blood cell; *Hb*, hemoglobin; *HbA*, major adult hemoglobin; *HbA2*, secondary adult hemoglobin; *HbF*, fetal hemoglobin; *SF*, serum ferritin

The symbol asterisk indicates the significance level; *$P < 0.05$, **$P < 0.01$. The continuous variables marked with superscript letter "T" fulfilled the criteria of normal distribution and were presented as mean ± SD and analyzed by t test. The continuous variables marked with superscript letter "M" did not fulfill the criteria of the normal distribution and were presented as median (Q1–Q3) and analyzed by Mann-Whitney U test. The categorical variables marked with superscript letter "F" were expressed as n (%) and analyzed by Fisher's exact test

Relationship between DEGs and therapeutic effectiveness

We further analyzed the differentially expressed genes in the efficient and inefficient treatment groups, and it was found that most DEGs were upregulated in the efficient treatment group while they were downregulated in the inefficient treatment group as shown in Fig. 3a. The same trend was observed among the study subjects of genotype β CD 41-42(-TTCT)/β^{A} (Fig. 3b).

Gene ontology of the DEGs

To explore the potential mechanisms of anti-anemia effect by CCA treatment in thalassemia patient, we performed GO analysis on DEGs. Data were filtered on the basis of current publicly available literature and experimental evidence regarding gene and hemoglobin associations with thalassemia. The GO project mainly covers the three domains of biological process,

Table 3 The change of clinical therapeutic indexes before and after treatment

Items	△RBC ($\times 10^{12}$/L)	△Hb (g/L)	△HbA (%)	△HbA2 (%)	△HbF (%)	△SF (ng/mL)
(a) Treatment group vs control group among all subjects						
Treatment group($n = 15$)	0.19 ± 0.32	5.40 ± 5.65	0.35 ± 0.37	− 0.11 ± 0.11	− 0.25 ± 0.36	− 34.03 ± 40.94
Control group($n = 5$)	0.06 ± 0.23	1.60 ± 4.45	0.02 ± 0.29	− 0.06 ± 0.11	0.04 ± 021	− 20.12 ± 25.15
Test value	0.838	0.644	1.820	− 0.815	− 1.653	− 0.709
P value	0.413	0.527	0.086	0.426	0.116	0.488
(b) Effective group vs ineffective group among all subjects						
Effective group($n = 6$)	0.50 ± 0.24	10.00 ± 3.63	0.37 ± 0.22	− 0.15 ± 0.16	− 0.22 ± 017	− 10.47 ± 21.21
Ineffective group ($n = 9$)	− 0.02 ± 0.13	0.67 ± 2.92	0.34 ± 0.46	−0.08 ± 0.04	0.27 ± 0.46	− 49.74 ± 44.25
Test value	5.362	5.516	0.109	−1.052	0.252	2.007
P value	0.000**	0.000**	0.915	0.337	0.805	0.066
(c) Effective group vs ineffective group among the subjects with genotype of β CD 41-42(-TTCT)/β^A						
41-42 effective group($n = 4$)	0.50 ± 0.08	9.00 ± 2.71	035 ± 0.26	− 0.15 ± 0.21	− 0.20 ± 0.22	− 12.28 ± 26.10
41-42 ineffective group ($n = 5$)	0.01 ± 0.10	0.20 ± 3.11	0.06 ± 0.18	− 0.08 ± 0.04	0.02 ± 0.15	− 43.56 ± 56.72
Test value	7.911	4.451	1.956	− 0.743	− 1.817	1.011
P value	0.000**	0.003**	0.091	0.482	0.112	0.346

△, the different level before and after treatment; *RBC*, red blood cell; *Hb*, hemoglobin; *HbA*, major adult hemoglobin; *HbA2*, secondary adult hemoglobin; *HbF*, fetal hemoglobin; *SF*, serum ferritin; *41-42 effective group*, the effective treatment group with genotype of β CD 41-42(-TTCT)/β^A

The symbol asterisk indicate the significance level; *$P < 0.05$, **$P < 0.01$

cellular component, and molecular Function (http://www.geneontology.org).

The gene ontogenesis analysis of our transcriptome data (Table 4) showed that the top five molecular functions of DEGs were correlated with following pathways: carbon dioxide transmembrane transporter activity, tRNA dihydrouridine synthase activity, ribosome binding, ammonium transmembrane transporter activity, and D-glucose transmembrane transporter activity. On the other hand, the top five biological processes included carbon dioxide transmembrane

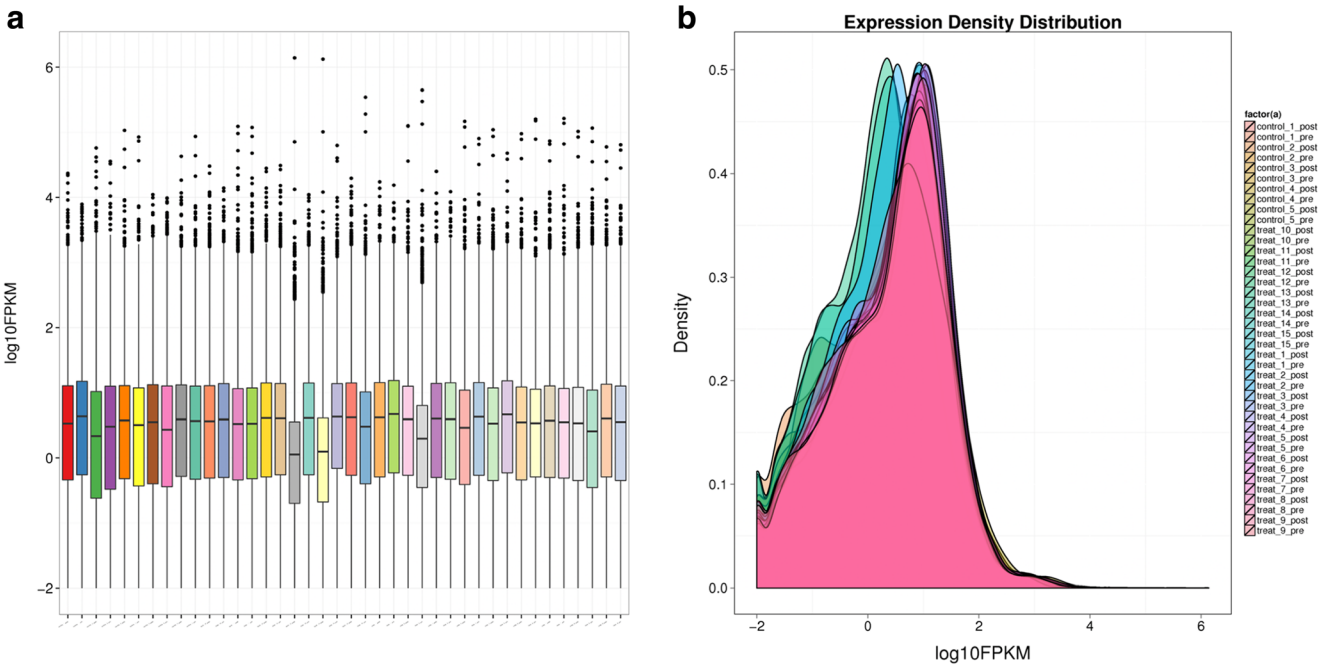

Fig. 1 Gene expression box plot and gene expression density map. **a** Gene expression box plot. *X* axis represents the sample name. *Y* axis represents the log10FPKM value. **b** Gene expression density map. *X* axis represents the log10 FPKM value. *Y* axis represents gene density. FPKM, fragments per kilobase of transcript per million mapped reads

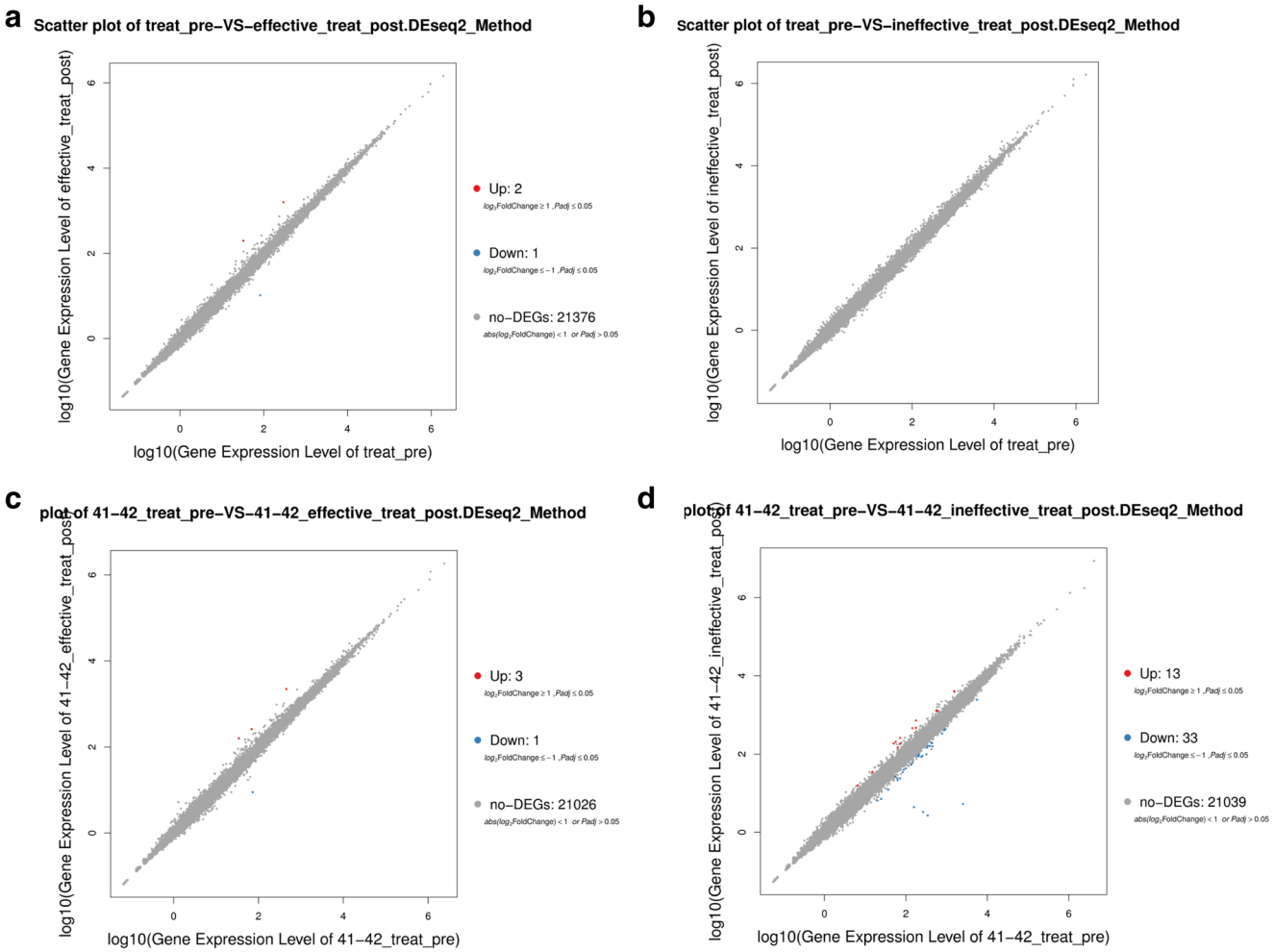

Fig. 2 Scatter plot of DEGs. *X* and *Y* axis represent log10-transformed gene expression level; red color represents the upregulated genes; blue color represents the downregulated genes; gray color represents the non-DEGs. DEGs, differentially expressed genes. **a** Comparison before and after treatment in the efficient treatment group. **b** Comparison before and after treatment in the inefficient treatment group. **c** Comparison before and after treatment in the efficient treatment group with a genotype of β CD 41-42(-TTCT)/β^{A}. **d** Comparison before and after treatment in the inefficient treatment group with a genotype of β CD 41-42(-TTCT)/β^{A}

transport, negative regulation of cell fate commitment, one-carbon compound transport, translational elongation, and tRNA dihydrouridine synthesis. Adding to that, the top five cellular components of DEGs are involved in cytoplasm, spectrin-associated cytoskeleton, mitochondrial envelope, mitochondrial membrane, and organelle envelope.

KEGG pathway of the DEGs

To investigate the biological significance of DEGs and their link to CCA treatment efficacy, we uploaded these genes with correlation coefficient greater than or equal to 0.95 to the KEGG application for biological function and pathway analysis. Our result reveals that three top functional categories included cell processes, environmental information processing, and genetic information processing (Fig. 4). Forty-five percent of the cell process genes were known to be involved in transport and catabolism, and 30.0% genes involved in the cellular community. More than 75% of the environmental information processing genes are functionally related to signal transduction. About 36.8% of genes belonging to the category of genetic information processing are associated with functional pathway of folding, sorting, and degradation. The remaining 26.3%, 21.1%, and 15.8% of genes were associated with the functional pathway of transcription, translation, and replication and repair, respectively.

To best investigate the pathways and molecular interactions of the identified genes, we deployed the latest version of the KEGG database for pathway enrichment analysis to examine DEGs (http://www.genome.jp/kegg). Our results demonstrated that upregulated genes were enriched for the KRAB domain–containing zinc-finger protein (KRAB-ZFPs) pathway and THO complex subunit 5 (THOC5) pathway (Table 5). Among the efficient treatment group of study subjects with gene type of β CD 41-42(-TTCT), two genes were consistently upregulated (Table 6).

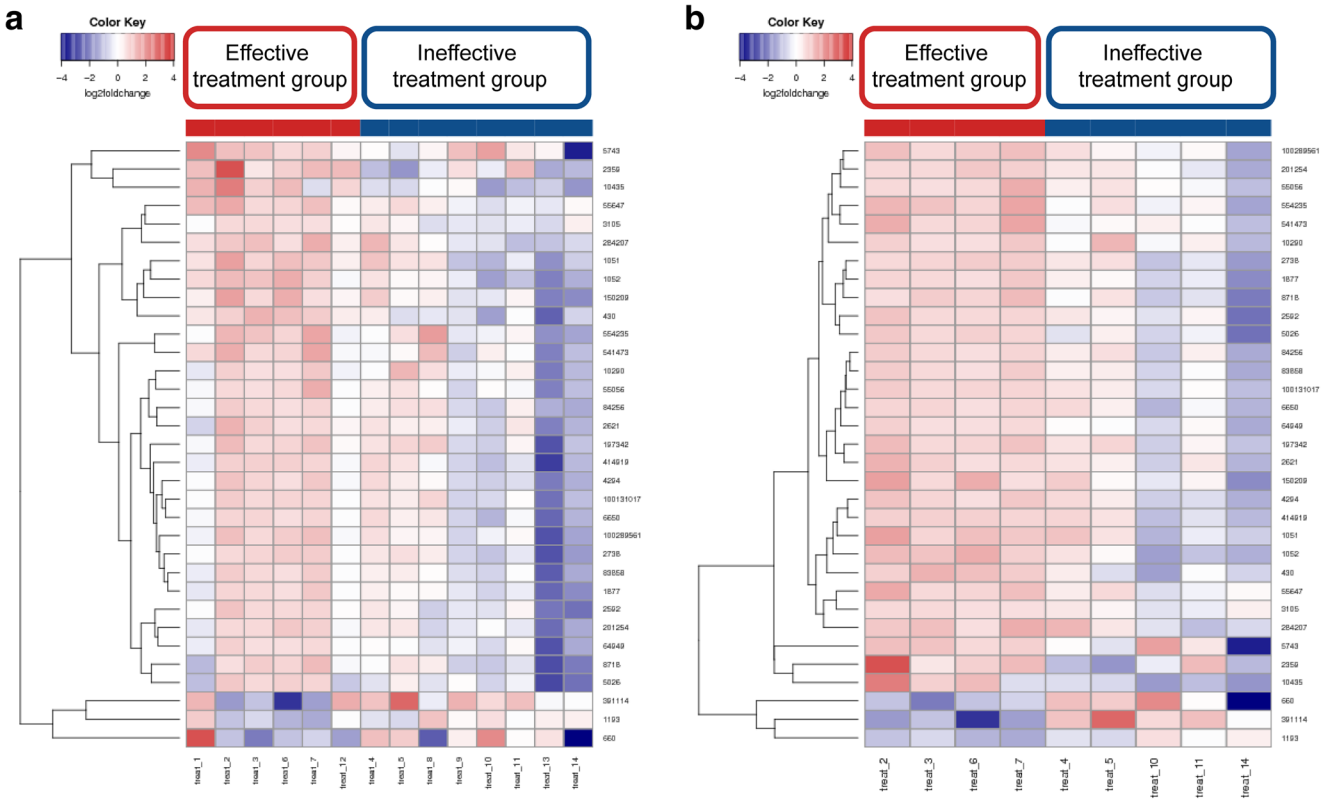

Fig. 3 The expression patterns of upregulated genes and downregulated genes were shown by the heat map. **a** Comparison between the effective group and the ineffective group among all study subjects. **b** Comparison between the effective group and the ineffective group among study subjects whose genotype was β CD 41-42(-TTCT)/β^{A}. In the middle of the figure, red and blue indicate upregulation and downregulation, respectively

Discussion

Ineffective erythropoiesis and excessive erythrocytolysis are central characteristics of β-thalassemia [13, 14]. Theoretically, treatments targeting at promoting erythropoiesis or reducing erythrocytolysis could both increase hemoglobin level and improve anemia in β-thalassemia. Our current findings are consistent with prior study that CCA therapy can

Table 4 Interaction of DEGs associated with top 5 GO (correlation coefficient ≥ 0.95)

Domain	Gene ontology term	Cluster frequency	Genome frequency of use	Corrected *P* value
Molecular function	1. Carbon dioxide transmembrane transporter activity	2 out of 136 (1.5%)	2 out of 16,991 (0.0%)	0.01405
	2. tRNA dihydrouridine synthase activity	2 out of 136 (1.5%)	4 out of 16,991 (0.0%)	0.08344
	3. Ribosome binding	4 out of 136 (2.9%)	50 out of 16,991 (0.3%)	0.15025
	4. Ammonium transmembrane transporter activity	3 out of 136 (2.2%)	29 out of 16,991 (0.2%)	0.34781
	5. D-glucose transmembrane transporter activity	2 out of 136 (1.5%)	9 out of 16,991 (0.1%)	0.48770
Biological process	1. Carbon dioxide transmembrane transport	2 out of 130 (1.5%)	2 out of 17,008 (0.0%)	0.06296
	2. Negative regulation of cell fate commitment	3 out of 130 (2.3%)	11 out of 17,008 (0.1%)	0.07476
	3. One-carbon compound transport	3 out of 130 (2.3%)	11 out of 17,008 (0.1%)	0.07476
	4. Translational elongation	6 out of 130 (4.6%)	114 out of 17,008 (0.7%)	0.26203
	5a. tRNA dihydrouridine synthesis	2 out of 130 (1.5%)	4 out of 17,008 (0.0%)	0.37399
	5b. Carbon dioxide transport	2 out of 130 (1.5%)	4 out of 17,008 (0.0%)	0.37399
Cellular component	1. Cytoplasm	108 out of 139 (77.7%)	11,055 out of 18,127 (61.0%)	0.00453
	2. Spectrin-associated cytoskeleton	3 out of 139 (2.2%)	8 out of 18,127 (0.0%)	0.00550
	3. Mitochondrial envelope	16 out of 139 (11.5%)	707 out of 18,127 (3.9%)	0.02413
	4. Mitochondrial membrane	15 out of 139 (10.8%)	669 out of 18,127 (3.7%)	0.04437
	5. Organelle envelope	20 out of 139 (14.4%)	1128 out of 18,127 (6.2%)	0.08739

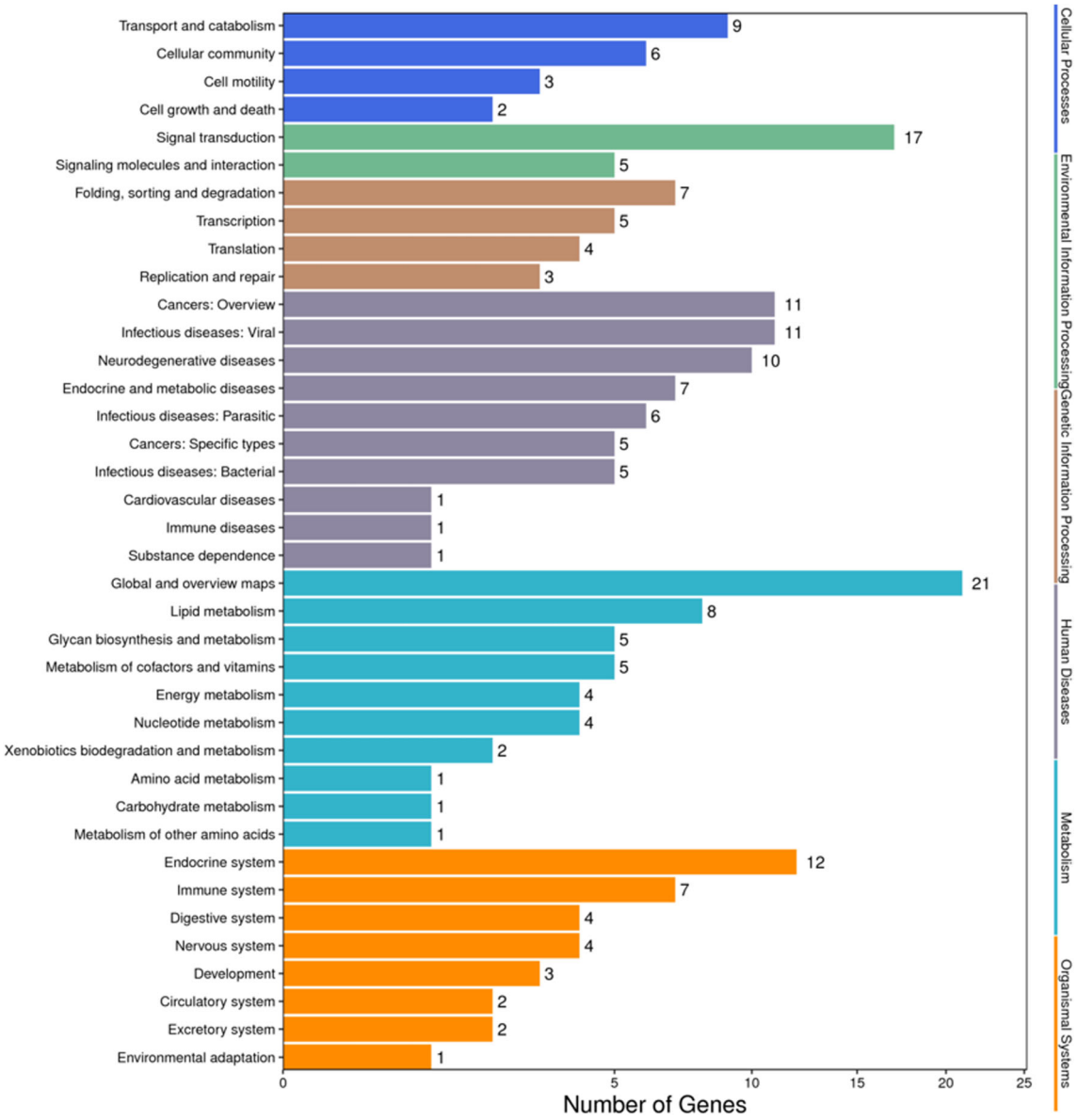

Fig. 4 KEGG of DEGs with correlation coefficient ≥ 0.95. DEGs, differentially expressed genes

effectively increase Hb level in pregnant women with β-thalassemia [7]. The improvement is accompanied by an increase in the erythrocyte count in peripheral blood. Based on these results, we hypothesized that CCA therapy increases erythroid cell numbers and hemoglobin level by improving survival or proliferation of erythrocyte.

However, multiple regulatory pathways are involved in the process of erythrocyte production and maturation as well as the degradation of hemoglobin [14]. To unveil the impact of CCA therapy on the complexity of erythorcytes and hemoglobin production, we applied the high-throughput transcriptome sequencing method to study the changes of transcriptome in the peripheral blood cell of pregnant women with β-thalassemia before and after the CCA treatment in order to explore the potential mechanism, targets, and pathways involved.

Potential mechanisms by CCA in improving anemia

When we compared the DEGs and the enrichments between the treatment group and the control group before and after treatment, neither few DEGs nor clear enrichments were found (not shown). We further segregated the treatment group into efficient treatment group and inefficient treatment group and a clear enrichment and more DEDs became eminent. The findings suggested that CCA played a role in the treatment of anemia among the responders through affecting certain processes. However, genotype diversity in non-responder in the treatment group may compound the findings.

The gene ontogenesis analysis of the transcriptome data reveals five top molecular functions and biological processes of DEGs related to events of gene transcription and protein

Table 5 The DEGs in the efficient treatment group comparing pre- and post-treatment (all study subjects)

Gene ID	Symbol	Regulation	log2 FC	Padj	*P* value	Cellular component	Molecular function	Biological process	Kegg orthology
57,573	ZNF471	Up	0.137	3.01E-07	0.588	Intracellular membrane-bounded organelle (GO:0043231)	Transition metal ion binding (GO:0046914); Nucleic acid binding (GO:0003676)	Transcription, DNA templated (GO:0006351)	KRAB domain–containing zinc-finger protein (K09228)
8563	THOC5	Up	− 0.103	4.64E-05	0.195	Nuclear lumen (GO:0031981); cytoskeleton (GO:0005856); Nuclear chromosome, telomeric region (GO:0000784); Nucleoplasm (GO:0005654); THO complex part of transcription export complex (GO:0000445); Cytoplasm (GO:0005737); Transcription export complex (GO:0000346)	RNA binding (GO:0003723); Protein binding (GO:0005515); Nucleic acid binding (GO:0003676)	mRNA export from nucleus (GO:0006406); RNA processing (GO:0006396); Primitive hemopoiesis (GO:0060215); Negative regulation of myeloid leukocyte differentiation (GO:0002762); Termination of RNA polymerase II transcription (GO:0006369); mRNA 3′-end processing (GO:0031124); Viral mRNA export from host cell nucleus (GO:0046784); Negative regulation of DNA damage checkpoint (GO:2000002); Positive regulation of DNA-templated transcription, elongation (GO:0032786); Embryonic hemopoiesis (GO:0035162); RNA splicing (GO:0008380); Monocyte differentiation (GO:0030224)	THO complex subunit 5 (K13174)
BGI_novel_G001638	NA	Down	0.119	0.000	0.661	NA	NA	NA	NA

log2 FC, log2-transformed fold-change between control and treat samples; *Padj*, statistic of adjusted *P* value; *NA*, not available

Table 6 The DEGs in efficient treatment group comparing pre-treatment and post-treatment (study subjects with gene type of β CD 41-42(-TTCT))

Gene ID	Symbol	Regulation	log2 FC	Padj	*P* value	Cellular component	Molecular function	Biological process	Kegg orthology	Nr description
57,573	ZNF471	Up	0.109	1	0.700	Same as in Table 5				Zinc-finger protein 471
8563	THOC5	Up	− 0.072	1	0.495	Same as in Table 5				HCG2011153, isoform CRA_a
BGI_novel_G001638	NA	Up	0.382	1	0.282	Same as in Table 5				Unnamed protein product
102,724,951	LOC102724951	Down	− 0.494	1	0.239	NA	NA	NA	NA	Uncharacterized protein

log2 FC, log2-transformed fold-change between control and treat samples; *Padj*, statistic of adjusted *P* value

translation, including the tRNA dihydrouridine synthesis, ribosome binding, and ammonium transmembrane transporter activity. Pathways related to erythrocyte function and apoptosis, such as carbon dioxide transmembrane transport and negative regulation of cell fate commitment were also found. The main cellular components of DEGs were involved in cytoplasm, mitochondrial membrane, organelle membrane, and spectrin-associated cytoskeleton.

Given the important role of the mitochondria in cellular function by contributing to 95% of the energy required for cellular activity for immature erythroid cells, changes in mitochondrial membrane and other organelle membrane like Golgi could impact the function of immature erythrocytes in the bone marrow and eventually the lifespan of mature erythrocyte released in circulation [14, 15]. Therefore, CCA treatment might potentially influence clinical outcome via acting on mitochondria function in immature erythrocytes.

Apart from the nucleus and cytoplasm of immature erythroid cells in the bone marrow stage, the membrane of immature or mature erythrocyte also plays an essential role in the life span and function of the erythrocyte. It is not only responsible for maintaining the normal morphology of the erythrocyte but also the interactions between the erythrocyte and the external environment. Cellular functions including oxygen and carbon dioxide exchange as well as drug interaction are exerted via erythrocyte membranes. Erythrocytic membrane is composed of phospholipid, integral protein, and cytoskeleton. Among these biomolecules, spectrin is the most abundant protein with the highest molecular weight in the erythrocytic membrane skeleton, accounting for 75% of the total membrane protein. Therefore, spectrin is central to the structural support for intact cell morphology of erythrocytes. Notably, cellular deformability is the most important characteristic required for the survival of erythrocytes, and changes in the spectrin level will significantly impact the function and life span of erythrocytes and eventually the anemia status in the patient [16].

Taken together, our findings suggest that multiple mechanisms might account for CCA increasing red blood cell count and hemoglobin level. Based on the bioinformatics analysis, we hypothesize that CCA regulates the amino acid transport of tRNA by activating the synthesis of uracil and triggers the translation process of globin by activating ribosome binding. It may also improve the structural stability of the erythrocyte and reduce the risk of deformity by controlling spectrin-associated cytoskeleton. With the increase of red blood cells and hemoglobin, the function of hemoglobin binding and transporting carbon dioxide would be improved.

Potential targets and pathways involved in CCA of improving anemia

From the pathway analysis, we also found that the Kruppel-associated box zinc-finger proteins (KRAB ZFPs) pathway and THOC5 pathway showed consistent enrichment of upregulated genes in multiple comparisons. In humans, KRAB ZFPs account approximately 30% of the 800 different zinc-finger proteins. Previous experimental data have reported that majority of the KRAB ZFPs family have a role in biological processes including regulation of cell differentiation, cell proliferation, apoptosis, neoplastic transformation, and cell cycle regulation [17, 18]. Recent study results also revealed that KRAB ZFPs can have impact on transcription either positively or negatively depending on the pathway involved and thereby regulate the genes in the transcriptional and post-transcriptional regulatory processes [17]. It is noteworthy that several KRAB ZFPs are also involved in the modulation of cell proliferation and differentiation in hematopoiesis, and it is possible that CCA might exert its anti-anemic effect in the upstream of hematopoiesis [17].

Previous researches demonstrated THO complex as a transcription and export complex (TREX). THOC5 is a member of THO complex that is directly or indirectly involved in transcriptional elongation, nuclear RNA export, and genome stability [19]. Recently, increasing experimental evidences showed that THOC5 is also an essential element in the maintenance, differentiation, and proliferation of stem cells. Depletion of THOC5 in adult mice leads to rapid apoptosis

and abnormal differentiation of hematopoietic cells and then the animals became anemic [20, 21].

Therefore, our results suggest that ZNF471 and THOC5 might be the potential targets of CCA and their actions are exerted via KRAB ZFPs pathway and THOC5 pathway on upstream hematopoiesis, respectively.

Potential genotype-dependent response of CCA efficacy

More than 50% of the participants in our present study are of genotype β CD 41-42(-TTCT)/β^A. Interestingly, this group of participants turned out to achieve better response to CCA treatment compared with those with other genotypes. Moreover, among the patients of genotype β CD 41-42(-TTCT)/β^A, the more responsive to CCA treatment, the more consistent upregulation or downregulation of certain genes were observed. The codon 41-42(-TTCT) mutation in β-globin gene is a very common genotype of β-thalassemia in Chinese, caused by the frame shift mutation of codon 41-42. The mutation causes a loss of function of β-globin and resulted in β-thalassemia [14]. Based on the correlation between genotype and treatment efficacy, we suspect that genotype specificity may be influencing factors to the therapeutic effect of CCA. It implies that potentially clinicians can assign the treatment to β-thalassemia patients based on their genotype and up to 45% of the β-thalassemia patient population might achieve better clinical outcome.

Study limitations and future expectations

This is the first transcriptomic study to demonstrate the potential mechanisms, targeted genes, and pathways involved in a single-agent TCM therapy for improving anemia of pregnant women with β-thalassemia. There are, however, certain limitations to this study. Firstly, the current sample size of 20 participants was inadequate to draw a conclusion on the impact of genotype difference in β-thalassemia and its relationship to CCA treatment efficacy. Secondly, the baseline of Hb level was significantly different between the treatment and control groups. The gestational age was also statistically different between the effective treatment and ineffective treatment groups. While the patients were recruited and randomized without bias, the small sample size might have contributed to the observed difference and thereby compromised the comparability of subjects. Last but not the least, we did not include subgroup analysis during the stage of research design for comparison between the effective treatment and ineffective treatment subgroups as well as the different genotype subgroups. Therefore, our subgroup analysis is considered a retrospective grouping analysis which would inevitably introduce a selection bias and would require prospective study to further confirm the findings.

Nevertheless, our results provides insight for future work to further investigate the role of genes KRAB ZFPs and THOC5 for their regulatory roles of Hb levels in pregnant women with β-thalassemia. To conclude on the genotype-specific therapeutic effect by CCA on β-thalassemia, a multicenter and prospective, randomized controlled trial is required to ensure sufficient number of pregnant women with thalassemia enrolled to the study.

Conclusions

In conclusion, our present study suggested that the anti-anemia effect of CCA on pregnant women with β-thalassemia might be related to translation regulation, especially the translation processes associated with spectrin synthesis, subsequently improving the membrane stability and prolonging the life span of the erythrocyte. Moreover, gene ZNF471 and THOC5 are the potential targeted genes by CCA treatment through KRAB ZFPs pathway and THOC5 pathway, respectively. Further verification of these transcriptomic changes at functional level are required for development of an effective and safe treatment strategy to improve clinical outcome of β-thalassemia patient during pregnancy.

Funding information This study was supported by the Science and Technology Project of Guangdong Province (grant no. 2017ZC0150) and by grants from the National Natural Science Foundation of China (grant no. 81704111).

Compliance with ethical standards

The procedures performed in the study were approved by the local ethics committee and were conducted in accordance with the current version of the Declaration of Helsinki.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Yin A, Li B, Luo M, Xu L, Wu L, Zhang L, Ma Y, Chen T, Gao S, Liang J, Guo H, Qin D, Wang J, Yuan T, Wang Y, Huang WW, He WF, Zhang Y, Liu C, Xia S, Chen Q, Zhao Q, Zhang X (2014) The prevalence and molecular spectrum of alpha- and beta-globin gene mutations in 14,332 families of Guangdong Province, China. PLoS One 9(2):e89855
2. Rakhi PN, Sophie L (2012) Baby on board: what you need to know about pregnancy in the hemoglobinopathies. Hematology 208–213
3. Voskaridou E, Balassopoulou A, Boutou E, Komninaka V, Christoulas D, Dimopoulou M, Delaki EE, Loukopoulos D, Terpos E (2014) Pregnancy in beta-thalassemia intermedia: 20-year experience of a Greek thalassemia center. Eur J Haematol 93(6):492–499
4. Leung TY, Lao TT (2012) Thalassaemia in pregnancy. Best Pract Res Clin Obstet Gynaecol 26(1):37–51

5. Amooee S, Samsami A, Jahanbakhsh J (2011) The pregnancy outcome in patients with minor β-thalassemia. Iran J Reprod Med 9(1): 9–14
6. Costa D, Capuano M, Sommese L, Napoli C (2015) Impact of epigenetic mechanisms on therapeutic approaches of hemoglobinopathies. Blood Cells Mol Dis 55(2):95–100
7. Li Y, He H, Yang L, Li X, Li D, Luo S (2016) Therapeutic effect of Colla corii asini on improving anemia and hemoglobin compositions in pregnant women with thalassemia. Int J Hematol 104(5): 559–565
8. Cappellini MD, Cohen A, Eleftheriou A et al (2008) Guidelines for the clinical management of thalassaemia. Thalassaemia International Federation, Nicosia
9. Benoist BD, Mclean E, Egli I et al (2008) Worldwide prevalence of anaemia 1993–2005; WHO global database on anaemia. World Health Organization, Geneva
10. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with bowtie 2. Nat Methods 9:357–359
11. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323
12. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550
13. Muncie HJ, Campbell J (2009) Alpha and beta thalassemia. Am Fam Physician 80:339–344
14. Lichtman MA, Kipps TJ, Seligsohn U et al Williams hematology, 8th edn. http://www.accessmedicine.com. Accessed 12 May 2019
15. Koury MJ (2014) Abnormal erythropoiesis and the pathophysiology of chronic anemia. Blood Rev 28(2):49–66
16. Barminko J, Reinholt B, Baron MH (2016) Development and differentiation of the erythroid lineage in mammals. Dev Comp Immunol 58:18–29
17. Lupo A, Cesaro E, Montano G, Zurlo D, Izzo P, Costanzo P (2013) KRAB-zinc finger proteins: a repressor family displaying multiple biological functions. Curr Genomics 14(4):268–278
18. Yang P, Wang Y, Macfarlan TS (2017) The role of KRAB-ZFPs in transposable element repression and mammalian evolution. Trends Genet 33(11):871–881
19. Mouaikel J, Causse SZ, Rougemaille M et al (2013) High-frequency promoter firing links THO complex function to heavy chromatin formation. Cell Rep 5(4):1082–1094
20. Saran S, Tran DDH, Klebba-Färber S et al (2013) THOC5, a member of the mRNA export complex contributes to processing of a subset of wingless/integrated (Wnt) target mRNAs and integrity of the gut epithelial barrier. BMC Cell Biol 14:51
21. Mancini A, Niemann-Seyde SC, Pankow R, el Bounkari O, Klebba-Färber S, Koch A, Jaworska E, Spooncer E, Gruber AD, Whetton AD, Tamura T (2010) THOC5/FMIP, an mRNA export TREX complex protein, is essential for hematopoietic primitive cell survival in vivo. BMC Biol 8:1