

SCIENTIFIC REPORTS



OPEN

β -Thalassemia Patients Revealed a Significant Change of Untargeted Metabolites in Comparison to Healthy Individuals

Syed Ghulam Musharraf^{1,2}, Ayesha Iqbal², Saqib Hussain Ansari³, Sadia Parveen³,
Ishtiaq Ahmad Khan² & Amna Jabbar Siddiqui¹

Received: 02 June 2016

Accepted: 08 January 2017

Published: 13 February 2017

β -Thalassemia is one of the most prevalent forms of congenital blood disorders characterized by reduced hemoglobin levels with severe complications, affecting all dimensions of life. The mechanisms underlying the phenotypic heterogeneity of β -thalassemia are still poorly understood. We aimed to work over metabolite biomarkers to improve mechanistic understanding of phenotypic heterogeneity and hence better management of disorder at different levels. Untargeted serum metabolites were analyzed after protein precipitation and SPE (solid phase extraction) from 100 β -thalassemia patients and 61 healthy controls using GC-MS. 40 metabolites were identified having a significance difference between these two groups at probability of 0.05 and fold change >1.5 . Out of these 40 metabolites, 17 were up-regulated while 23 were down-regulated. PCA and PLS-DA model was also created that revealed a fine separation with a sensitivity of 70% and specificity of 100% on external validation of samples. Metabolic pathway analysis revealed alteration in multiple pathways including glycolysis, pyruvate, propanoate, glycerophospholipid, galactose, fatty acid, starch and sucrose metabolism along with fatty acid elongation in mitochondria, glycerolipid, glyoxylate and dicarboxylate metabolism pointing towards the shift of metabolism in β -thalassemia patients in comparison to healthy individuals.

β -Thalassemia is a common congenital haematological disorder which is characterized by dysregulation in the synthesis of the β -globin chain, one of the major constituents of adult haemoglobin (HbA)¹. It is anticipated that annually 70,000 children are born with various types of thalassemia, and majority of these births are affected by severe forms of β -thalassemia^{2,3}. Hundreds of mutations in the β -globin gene and/or regulatory elements associated with the β -globin gene are known to be the cause of this genetic haemoglobinopathy⁴. While in Pakistan, the most common mutation responsible for causing this blood disease is IVS1-5⁵.

In β -thalassemia, there is an imbalance in α/β -globin ratio and excessive α -globins possibly causes oxidative damage to membrane lipids and proteins of red cell in the form of irreversible hemichromes and also increases intracellular calcium, causing the significant increase in destruction of RBCs and ultimately anaemia⁶. Anaemia stimulates the erythropoietin production with subsequent intensive but ineffective expansion of the bone marrow (up to 25 to 30 times normal), which sequentially causes the typical described bone deformities. Prolonged and severe anaemia and increased erythropoietic activity also result in hepatosplenomegaly, extramedullary erythropoiesis, iron induced dysfunctions of various organs, thrombosis, diabetes, severe infection, and growth retardation⁷.

Revealing alteration of metabolites in the course of a disease in body fluids and tissues is an emerging application in the field of biomedical research as this area has the possibility to act as an effective tool for predicting disease phenotype at early stage, and also predicting response of a treatment and survival⁸. In recent times,

¹H.E.J. Research Institute of Chemistry, International Center for Chemical and Biological Sciences, University of Karachi, Karachi-75270, Pakistan. ²Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi-75270, Pakistan. ³Department of Pediatric Hematology & Molecular Medicine, National Institute of Blood Diseases and Bone Marrow Transplantation, Karachi-75300, Pakistan. Correspondence and requests for materials should be addressed to S.G.M. (email: musharraf1977@yahoo.com)

metabolomics has been used as potential biomarkers in organ transplantation and immunosuppressant toxicity⁹, assessing pathogenesis of lung diseases^{10,11}, toxicology¹², drug discovery and precision medicine¹³ and cancer biology^{14–16}.

Many techniques have been used for screening and diagnosis of haemoglobin variants and thalassemia¹⁷. Determination of the genetic makeup of the person in question and characterization of human blood using complete blood cell count (CBC) are the most reliable methods for diagnosis of thalassemia. Still there is a limitation in the analysis of data due to a large number of possible candidate characteristics and various types of thalassemia and thalassemia trait¹⁸. Moreover, using such methods, there would be no information about alterations in the patterns of metabolites present in the biological materials that can give valuable phenotypic information and mechanistic insight into the biochemistry of disease processes and related abnormalities. Limited studies for markers identification in the blood or urine of β -thalassemia patients have been done which include the analysis of haemoglobin variants to diagnose thalassemia¹⁹, marker for lipid peroxidation-induced DNA damage²⁰, plasma substance P and soluble P-selectin as biomarkers of β -thalassemia induced hypercoagulability²¹, adipocytokines related to haemolytic and inflammatory biomarkers²², biomarkers of iron and oxidant-antioxidant homeostasis^{23,24}, nuclear magnetic resonance-based screening of thalassemia with quantification of some haematological parameters²⁵ and quantification of the free α -Hb²⁶. Various studies have been done showing that metabolic disorders are common in patients with β -thalassemia^{27,28} but to date, there is a lack of metabolomics based biomarkers that may play role in diagnosing the phenotype of β -thalassemia and convey prognostic approach with various management possibilities²⁶.

This study focuses on the untargeted metabolomic analysis of β -thalassemia to gain insights into the molecular and cellular pathogenesis to aid in understanding the pathophysiology of disease at molecular level and to identify biomarkers/biomarkers pattern. Various powerful data mining and statistical bioinformatics methods were used for identifying, prioritizing and classifying robust and generalizable biomarkers with high discriminatory ability. No paper has been published so far regarding the profiling and identification of serum metabolomics of β -thalassemia patients. It is expected that findings of this study will strengthen the knowledge of molecular mechanisms involved in β -thalassemia that will ultimately help in the identification of the characteristic molecule(s) as well as improvement of possible therapy of the disease.

Material and Method

Patient's Selection. The selection of patients for present study was carried out in National Institute of Blood Disease and Bone Marrow Transplantation (NIBD), Karachi, Pakistan after the ethical approval of the Institutional Review Board (IRB), ethic committee of hospital while the experimental protocols were approved by the primary research institute (International Center for Chemical and Biological Sciences, ICCBS). This study included 100 cases of β -thalassemia with following inclusion criteria: Patients have been registered at NIBD, diagnosed as case of β -thalassemia major (defined as Hb <7 g/dL, high HbF, absent or very low HbA, and more than 8 transfusions a year), sampling was carried out prior to blood transfusion if patient requires blood transfusion, and in 4–8 hours fasting condition. Exclusion criteria included: Patients having any evidence of other chronic illnesses unrelated to thalassemia and unwillingness to enrol in the study. Healthy volunteers were also recruited at Dr. Panjwani Center for Molecular Medicine and Drug Research (PCMD) for this study, basic details of patients and healthy controls are mentioned in supplementary information (Table S1). Written informed consent was obtained from all the participants of this study including control and patients. A thorough questionnaire consisting of questions for information required for study was also filled from all the patients.

Sample Collection. Sample collection was carried out in accordance with relevant guidelines and regulations. Blood was collected from participants after 8 hours fasting but for infants and toddlers sample was collected when they felt need for food. About 5 cc of blood was drawn from participants by venepuncture and collected in gel-based BD vacutainer tubes (BD Franklin Lakes NJ, USA, REF: 367381), interior coated with silicone for clot activation. Serum was separated by centrifugation at 2000 rpm for 10 minutes at 4 °C. And then serum was aliquoted and stored immediately at –80 °C freezer till further processing of sample.

Reagents and Solvents. Analytical grade solvents were used for GC-MS analysis. Reagents and solvents included methanol, hexane (Tedia, Tediaway, Fairfield, USA), myristic-d₂₇ acid, *N,O*-bis(trimethylsilyl) trifluoroacetamide (BSTFA) with trimethylchlorosilane (TMCS) (Sigma-Aldrich, St. Louis, Missouri, USA), methoxylamine hydrochloric (Acros Organic, New Jersey, USA), Pyridine (Lab-Scan, Bangkok, Thailand) and Mill-Q/ deionized water (Millipore, Billerica, MA, USA).

Sample Preparation and Derivatization. Protocol used for preparation of samples in this study for profiling of metabolites has been reported previously in detail and it include methods that were performed in accordance with the relevant guidelines and regulations²⁹. In short, first proteins were precipitated by adding 800 μ L of chilled methanol in 100 μ L serum containing 20 μ L of myristic acid (2 mg/mL) as internal standard. Supernatant was subjected to solid phase extraction using a 96 well plate (Strata C18-E, 55 μ m pore size, 70 Å particle, 100 mg sorbent/1 mL Phenomenex, USA) under vacuum (AHC-7502, Phenomenex, USA). After sample loading, the solid phase was washed with 300 μ L of water and metabolites were eluted with 600 μ L methanol and collected in a 96 well collection plate. The eluent was finally dried in vacuum at room temperature. The dried extract was stored at 4 °C until analysis. Derivatization of dried samples was carried out by addition of 50 μ L methoxylamine hydrochloride (15 μ g/ μ L in pyridine) followed by addition of 50 μ L BSTFA with 1% trimethylchlorosilane for formation of trimethylsilyl derivatives. Then sample was centrifuged and analyzed on GC/MS.

GC-MS Analysis. GC-MS analysis of derivatized samples was executed as mentioned previously with minor modifications in the GC method³⁰. Analysis was carried out on 7890 A GC (Agilent technologies, USA), fitted

S. No.	Compound (CAS No.)	Compound Found in No of Disease Samples/100	Compound Found in No of Healthy Samples/61	Retention Time	p (Corr) ([Disease] Vs [Healthy]) :Normalized	Log FC (abs) ([Disease] Vs [Healthy]) :Normalized
1	Geraniol (106-24-1)	98	26	14.18871	1.65E-19	1.365947
2	Palmitic Acid (55520-89-3)	100	60	23.55566	0.002264404	0.6102581
3	α -Glyceryl Palmitate (1188-74-5)	100	49	27.26245	1.04E-06	0.8656348
4	Lactic Acid (17596-96-2)	70	26	9.07812	0.00308354	0.5949775
5	α -Glyceryl stearate (1188-75-6)	100	61	28.265228	7.23E-06	0.80729926
6	M-Pyrol (872-50-4)	53	1	8.527779	7.22E-11	1.082454
7	Citronellyl formate (105-85-1)	81	17	12.806128	4.35E-11	1.0938886
8	Sucrose (19159-25-2)	73	17	27.751137	3.00E-08	0.9526255
9	Triethanolamine (20836-42-4)	84	3	17.671268	1.88E-30	1.5848455
10	5-Ethyl-5-methyldecane (17312-74-2)	28	1	16.734486	3.86E-04	0.68131626
11	2,3-dimethyl-2,3-butane-diol (76-09-5)	57	6	11.471427	9.78E-09	0.97938067
12	Boric Acid (4325-85-3)	36	0	7.138889	1.26E-06	0.85958195
13	Phosphoric Acid (10497-05-9)	20	0	12.92	0.002640698	0.6035476
14	312.0@21.099983	72	26	21.099983	0.002248546	0.61144674
15	147.0@6.300002	39	0	6.300002	1.91E-07	0.9072522
16	151.0@6.0	32	3	6	6.99E-04	0.6602574
17	71.0@9.700005	45	5	9.700005	1.06E-05	0.7957854

Table 1. List of up-regulated metabolites in β -thalassemia patients in comparison to healthy controls.

with a GC auto sampler 120 (PAL LHX-AG12–Agilent Technology) autosampler and coupled to Agilent 7000 Triple Quad system (Agilent Technologies, USA). A fused-silica capillary GC column, HP-5MS 30 m x 0.25 mm ID (Agilent J&W Scientific, Folsom, CA, USA), chemically bonded with a 95% dimethylpolysiloxane 5% diphenyl cross-linked stationary phase (0.25 mm film thickness) was used. The serum sample was injected in the splitless mode using helium as carrier gas. Initially the oven temperature was fixed at 50 °C for 1 min then temperature was raised in three steps. In first step temperature raised at a rate of 10 °C per minute to 80 °C for 3 min then again 10 °C per min to 180 °C for 3 min and in final step 15 °C per min raise to 300 °C for 5 min. After maintaining the temperature at 300 °C for five minutes, it was further increased to 305 °C for one minute which referred as post run. Retention time was locked to the internal standard at 20.070 min. Electron impact ionization (EI) was used as an ionization source for the GC/MS analysis at 70 eV. Data acquisition was done in full scan mode from 50–650 *m/z* in 0.5 seconds scan time. A blank was run between samples to remove contamination. Mass calibration was done with perfluorotributylamine (PFTBA).

GC-MS Data Preprocessing and Statistical Analysis. Agilent Mass Hunter Qualitative Analysis software (version B.04.00) was used for data processing. Peak integration and deconvolution parameters have been previously reported^{29,30}. Mass spectra of the peaks were compared with NIST mass spectral (Wiley registry NIST 11) library leading to presumptive identification of metabolites with a $\geq 70\%$ similarity index. The GC-MS spectra were uploaded on Mass Profiler Professional (MPP) software 12.5. Filtering of the data involved using all available data and minimum absolute abundance of 5,000 counts with 3 number of ions minimum. Match factor 0.3, retention time tolerance 0.05 and delta MZ (low resolution) 0.2 were set as alignment parameters. External scalar was used for normalization of data. Z transform was selected as base lining option treating all the compounds equally irrespective of their intensity. A total of 711 compounds were detected in the entire samples after alignment of data. Statistical significance analysis was done using student T-test unpaired for healthy versus β -thalassemia patients of fold change (FC) 1.5. A PLSDA model was built for healthy versus β -thalassemia patients using auto scaling, N fold validation type, three number of folds and ten number of repeats. Blind samples (n = 20) were also run for external validation. Sensitivity and specificity of the constructed model were also measured.

Results

Metabolite profiling of total 161 serum samples including healthy volunteers (n = 61) and β -thalassemia patients (n = 100) were analysed by using GC-EI-MS. After performing GC-MS analysis as described above, identification of metabolites was carried out using Agilent Mass Hunter Qualitative Analysis software and NIST library. Statistical and multivariate data investigation i.e. Heatmap, PCA plot and PLSDA plot was carried out using MPP software in order to identify comparative and statistically significant metabolites between healthy and β -thalassemia samples. Significant testing and fold change was carried out on total of 711 entities found in this experiment. Student's T-test unpaired, asymptomatic p-value computation and multiple testing correction by Benjamini Hochberg FDR was applied. A list of 40 compounds was generated at probability of 0.05 and fold change >1.5. Out of these 40 metabolites, 17 metabolites were up-regulated and 23 were down-regulated in β -thalassemia patients in comparison to healthy controls as listed in Tables 1 and 2, respectively with their CAS registry numbers. Among these 40 metabolites, 8 were showing a fold change of 2 among disease and healthy group as shown in Fig. 1. Twenty one out of forty low molecular weight metabolites; geraniol, palmitic acid, α -glyceryl palmitate, lactic acid, α -glyceryl stearate, M-pyrol, citronellyl formate, sucrose, triethanolamine, 5-ethyl-5-methyldecane, 2,3-dimethyl-2,3-butane-diol, boric acid, phosphoric acid, hexadecane, methylbis(phenylmethyl)benzene, dodecane, 4,6-dimethyl, phthalic acid, glycerol, stearic acid, n-pentatriacontane and ethylene

S. No.	Compound (CAS No.)	Compound Found in No of Disease Samples/100	Compound Found in No of Healthy Samples/61	Retention Time	p (Corr) ([Disease] Vs [Healthy]) :Normalized	Log FC (abs) ([Disease] Vs [Healthy]) :Normalized
1	Hexadecane (544-76-3)	10	22	23.89688	9.40E-04	-0.6479106
2	Methylbis(phenylmethyl)benzene (999383-38-3)	0	14	25.5	6.50E-06	-0.8112275
3	Dodecane, 4,6-dimethyl (61141-72-8)	15	30	17.23111	3.91E-05	-0.75676084
4	Phthalic acid (117-81-7)	6	33	27.08717	1.49E-11	-1.11504
5	Glycerol (6787-10-6)	49	50	12.99899	0.001520142	-0.6268371
6	Stearic acid (57-11-4)	0	21	24.59524	1.61E-09	-1.0188797
7	n-Pentatriacontane (630-07-9)	3	14	20.94118	8.19E-04	-0.65371585
8	Ethylene glycol (7381-30-8)	0	14	7.350001	6.50E-06	-0.8115035
9	77.0@19.5	1	27	19.5	3.94E-12	-1.1401991
10	77.0@19.200005	0	22	19.20001	4.89E-10	-1.0456579
11	73.0@8.100001	1	21	8.100001	1.22E-08	-0.9731995
12	55.0@24.399998	0	9	24.4	0.001147105	-0.63986087
13	71.0@22.799995	5	22	22.8	3.09E-06	-0.8342981
14	71.0@25.899998	1	19	25.9	1.43E-07	-0.91528964
15	73.0@12.899999	2	11	12.9	0.002962446	-0.5973833
16	71.0@13.600004	11	25	13.6	1.45E-04	-0.7139679
17	57.0@24.600004	2	11	24.6	0.003284676	-0.59084994
18	69.0@12.799999	2	23	12.8	8.59E-09	-0.9838301
19	57.0@11.699998	1	17	11.7	2.23E-06	-0.8439079
20	73.0@14.199998	0	12	14.2	5.17E-05	-0.74642193
21	73.0@18.5:1	24	32	18.5	0.003341812	-0.5893601
22	73.0@9.199999	5	15	9.199999	0.003284676	-0.5910041
23	57.0@27.100002	0	8	27.1	0.00271492	-0.60159874

Table 2. List of down-regulated metabolites in β -thalassemia patients in comparison to healthy controls.

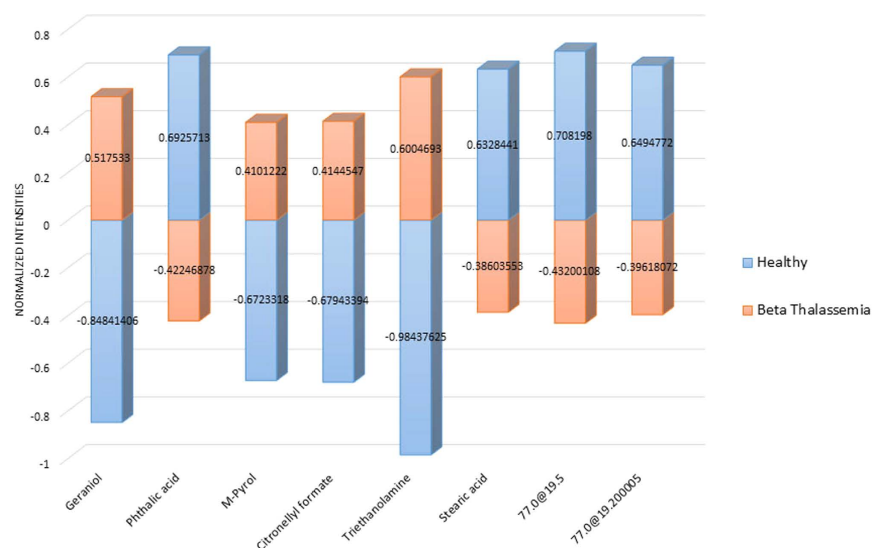


Figure 1. Eight metabolites at FC >2, out of which 4 are down-regulated (bar below the baseline of 0) while 4 are up-regulated (bar above the baseline of 0) in β -thalassemia patients as compared to healthy group.

glycol; were putatively determined by comparing the mass spectra of the peaks with those available in the NIST mass spectral (Wiley registry NIST 11) library at $\geq 70\%$ similarity index while the remaining were not identified at this similarity index. The identified compounds are shown with their name while unidentified with their base peak and retention time in Tables 1 and 2. The EI/MS spectra of remaining nineteen unidentified compounds are shown in supplementary information (Figure S1). Principal component analysis (PCA) was also carried out on our data to make sure that the difference in metabolic pattern is due to difference in health status and not

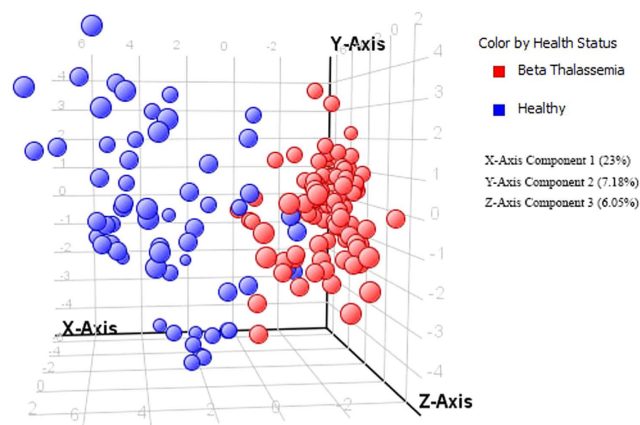


Figure 2. PCA score plot of healthy and β -thalassemia patients based on forty (40) significantly differentiated metabolites i.e. having a fold change >1.5 . Variance on X, Y and Z axis is also shown within brackets.

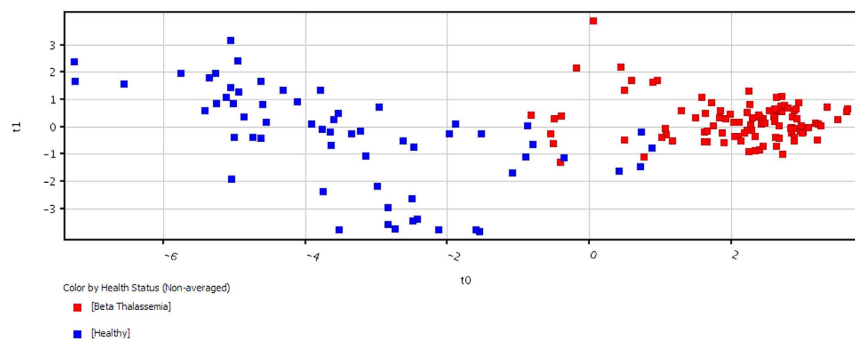


Figure 3. PLS-DA scores scatter plots discerning among healthy controls and β -thalassemia patients based on forty (40) significantly differentiated metabolites i.e. having a fold change >1.5 .

attributable to age or weight. We found that samples were not separating on the basis of age or body mass shown in supplementary information (Figures S2 and S3) therefore excluding these confounders.

PCA was carried out and a model was generated which revealed a vibrant and noteworthy difference between the non-averaged healthy samples and β -thalassemia samples. The PCA scores are shown in Fig. 2 in which each sample is denoted by a single point. The 50% cumulative variance of samples was observed at component 6 and variance of first three components on X, Y and Z axis are found to be 23%, 7.18% and 6.05% respectively. So this indicates that there are many factors responsible for discriminating metabolites between the healthy and disease group. A prediction model of healthy versus disease group was built by multivariate data analysis that include all analysed samples i.e. 100 β -thalassemia and 61 healthy samples and on the basis of forty metabolites having a statistically important difference in expression between these two clusters. Samples were classified into discrete classes also by supervised Partial Least Square Discriminant Analysis (PLSDA). Two parts of the input data were randomly assigned to the training set and remaining into the testing set. Auto-scaling was applied which involves subtracting the variable mean from each variable (data column) and dividing each by its standard deviation. This process was repeated ten times, each time using a different part for testing thus using each row once in training and testing generate a Confusion Matrix, which gives accuracy of prediction for each class. Plots obtained by PLS-DA scores are shown in Fig. 3 exposing an unblemished separation trend between the two sets of our experiment. Sensitivity of the constructed model was calculated from the proportion of β -thalassemia samples that were predicted correctly and referred as true positives, while specificity was determined from the proportion of control samples which were correctly predicted and these are stated as true negatives. Sensitivity of our built model was found to be 92.0% and specificity was 95.0%, respectively, while the overall accuracy of the model was 93.1% as mentioned in Table 3. The predictive capacity (i.e. sensitivity and specificity) of the model was measured also by external validation using 20 serum samples consisting of 10 samples each from healthy controls and β -thalassemia patients. But these samples were decoded prior to preparation and analysis by GC-MS therefore these were an independent or blind test set of samples. External validation correctly predicted the presence of β -thalassemia in 7 out of 10 patients and healthy controls in 10 out of 10 patients resulting in a sensitivity of 70% and specificity of 100%. Sample prediction reports are shown in Figure S4 of supplementary information.

To identify metabolic pathways those are disturbed in β -thalassemia we used web based software MetaboAnalyst 3.0 (www.metaboanalyst.ca/) in which previously mentioned list of identified metabolites was

	B-Thalassemia Predicted	Healthy Predicted	Accuracy
True B-Thalassemia	92	8	92.00
True Healthy	3	58	95.08
Overall Accuracy			93.17

Table 3. Confusion Matrix of Model generated from healthy controls (n = 61) and β -thalassemia patients (n = 100).

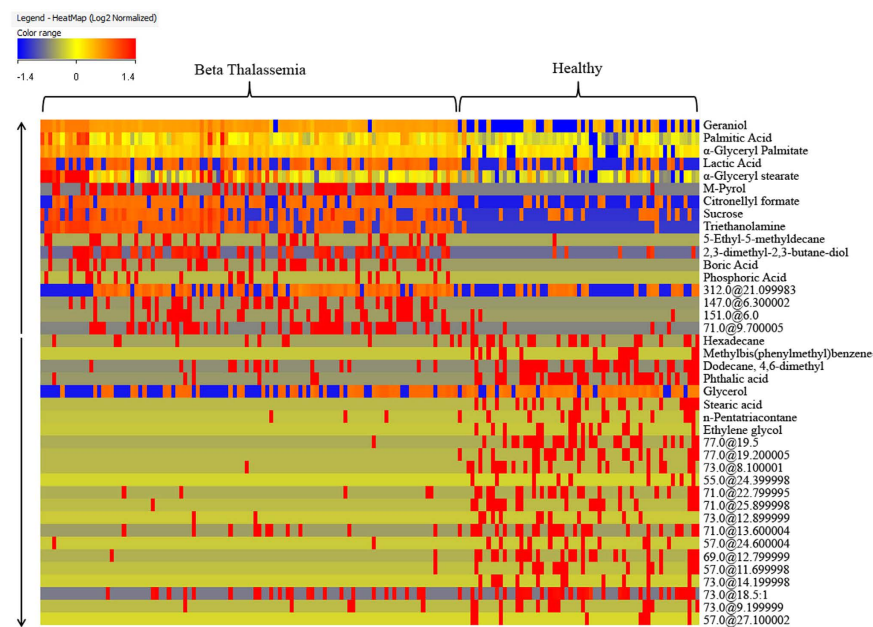


Figure 4. Image of heatmap using non-average samples with normalized intensities of 40 significant metabolites that showed significant difference among healthy & disease group. The arrows indicate the up-regulation and down-regulation of metabolites.

entered. On the basis of several databases such as KEGG (Kyoto Encyclopedia of Genes and Genomes) (<http://www.genome.jp/kegg/>)^{31,32} and HMDB (Human Metabolome Database) (<http://www.hmdb.ca/>) this online software helps in identifying the pathways with significant alterations. The summaries of pathway analysis created on the basis of hypergeometric test and relative-betweenness centrality in pathway topology analysis by this program using up-regulated and down-regulated metabolites are shown in Figures S5 and S6, respectively. The images of distinguishing identified pathways are provided in supplementary data i.e. Figure S7.

Discussion

The innovative omics technology has opened up exciting opportunity for screening and identification of novel biomarker that can acts as an indicator for the physiological alteration of body. Evolving technologies of metabolomics profiling hold potential for lighting biology and human diseases. Metabolites have a wide range of functional groups are present from volatile alcohols, ketones, amines, organic acids to complex lipids, carbohydrates and other secondary metabolites. β -thalassemia is one of the very frequent and extremely disabling genetic disease. Different pathological and environmental stresses change expression level of certain genes and hence concentrations of metabolites of corresponding pathways. Therefore, we aimed to determine these changes in metabolome of β -thalassemia patients for disease prognosis and to understand unclear pathophysiological mechanisms of thalassemia, as field of metabolomics has proven itself a promising technique in understanding pathophysiology of many other various diseases also including genetic diseases such as sickle cell anaemia^{33–36}. However, our results are limited by the fact that the type of mutation was not known in patients.

The comparison of serum metabolites between β -thalassemia patients and normal subjects revealed evident alterations of metabolites in the disease group. The close image of heat map using non-average samples with normalized intensities of forty (40) significant metabolites are shown in Fig. 4 in which the identified metabolites are stated by their name while unidentified with their base peak and retention time. From this heat map it is quite clear that β -thalassemia metabolite profile is totally different from the control group and it can also be observed that the concentration of some of the metabolites is increased while of some is decreased. This change in metabolite profile indicates that in β -thalassemia patients metabolism is shifted from the normal state and it is also reported in literature that metabolism is disturbed in this genetic disease³⁷. Hence, knowledge related to these altered metabolites play an important role in understanding of disease progression at molecular level. Significance

of the transformed metabolite profile in β -thalassemia can be described by referring it to the human metabolome database (HMDB)^{38–40}.

Up-regulated metabolites. Geraniol also named as rhodinol is a monoterpene and an alcohol which occurs in essential oils of several aromatic plants. Its bio-functions include cell signalling, storage and source of fuel or energy and integrity of membrane. It possess anti-cancer, antimicrobial, anti-oxidant, anti-inflammatory and some vascular effects⁴¹, therefore it is possible that geraniol is increased as a result of oxidative stress, inflammation and decreased RBCs membrane integrity in β -thalassemia. Palmitic acid or hexadecanoic acid is one of the most common saturated fatty acids found in animals, a saturated fatty acid found in fats, waxes and body lipids. It is involved in various metabolic pathways in body and its altered levels are also reported in colorectal cancer, breast cancer, eosinophilic esophagitis and gastroesophageal reflux disease. Palmitic acid plays important functions other than providing energy⁴². One of their functions is to induce apoptosis, so their high levels may be responsible for early degradation of RBCs in β -thalassemia patients. α -Glycerol palmitate and α -glyceryl stearate are forms of fatty acid, both are one fatty acid chain containing glycerides, covalently bonded to a glycerol molecule through an ester linkage. Both of these are source of energy as well as required also in maintaining stability of membrane so its increase levels can be due to more destruction of red cells. Lactic acid is a crucial metabolite which is involved in various biochemical processes and its production take place due to extreme activity in muscles. It is a component of various metabolic pathways such as cysteine, propanoate and pyruvate metabolism. As it is a product of anaerobic glycolysis therefore its enhancement in β -thalassemia can be predicted as these patients have low Hb levels which cause less supply of oxygen to tissues and leading to more anaerobic glycolysis. The second reason for more lactic acid levels, that it may be due to poor hepatic function a consequence of iron overload in these patients as abnormal concentrations of lactic acid are found in hepatic biliary malignancies in addition to other cancers. M-pyrol is a product of GABA (γ -aminobutyric acid) a neurotransmitter, its altered levels are seen in bladder infection and urinary tract infections are common in β -thalassemia because of predisposing factors of such as splenectomy, iron overload, anaemia, and granulocyte dysfunctions⁴³. Sucrose is a non-reducing disaccharide of glucose and fructose and is linked to various metabolic pathways of glucose and other sugars. As it is broken down into its constituents fructose and glucose and its increase in blood indicates metabolic syndrome, mostly in β -thalassemia patients glucose homeostasis is abnormal this may result in its elevation²⁸. Citronellyl formate and triethanolamine are metabolites found in cytoplasm as well as extracellularly. Both of them are produced endogenously in addition to dietary source. Phosphoric acid is another important metabolite which we found to be up-regulated in the disease group. It is present in cytoplasm and it act as an osmolyte and enzyme cofactor in biological system. It is also involved in signalling and a list of metabolic pathways such as ammonia recycling, arginine, proline, cysteine, purine, pyruvate, inositol metabolism and various glucose metabolic pathways. Therefore, it can be stated that increase phosphoric acid levels are indicator that various metabolic pathways are up-regulated in this disease.

Down-regulated metabolites. Hexadecane is a 16 carbon atoms chain that has been shown to exhibit anti-inflammatory, anti-bacterial, anti-oxidant and thermogenic functions. Because in β -thalassemia all these activities are increased so low levels of hexadecane can be co-related with its more utilization and less availability in serum freely. Phthalic acid is a toxin or pollutant found in blood and when found in tissues or biofluids, it arises from exposure to phthalate products. Phthalate is an environmental chemical of high public concern because reports of its potential risk to male reproductive health so it can be said that iron load is a major reason of infertility in these patients and not the environmental toxins⁴⁴. Glycerol is a major component of phospholipids and triglycerides that can be converted into glucose by liver to fulfil energy requirements. It is a component of glycerolipid and glycerophospholipid metabolism. And its abnormal levels have been quantified and identified in various disorders^{45,46}, so it is obvious that in β -thalassemia glycerol is more consumed for energy production due to metabolic stress. Octadecanoic acid or Stearic acid is a beneficial saturated fatty acid involved in mitochondrial beta-oxidation of long chain saturated fatty acids and plasmalogen synthesis. These both pathways mainly contribute to maintain dynamics of membrane and cell signalling so low levels of stearic acid in body further contribute to decrease strength of RBCs membrane and altered cell signalling. Ethylene glycol also known as polyethylene oxide (PEO) or polyoxyethylene (POE), depending on its molecular weight is an oligomer or polymer of ethylene oxide. It functions in biosystem as a nutrient, anesthetic, anti-microbial, laxative and radical scavenger and its low levels may further aggravate the oxidative stress and increase susceptibility to infections.

Pathway Analysis. Pathways were produced from MetPA (Metabolomic Pathway Analysis) software that showed dysregulation in β -thalassemia patients (Table 4). Using MetPA identified metabolites were analysed that contains pathways from the KEGG metabolic pathways database and HMDB. Pathway enrichment with topology analysis, and an interactive visualization system is also used to find pathways that are most substantially altered under the conditions of particular experiment. In metabolic networks, more severe effects are produced due to changes in more “vital” locations on the pathway compared to variations occurring in bordering or comparatively isolated positions. In our analysis, we identified several pathways some were generated from list of metabolites that were up-regulated in β -thalassemia, including fatty acid elongation in mitochondria, glycolysis or gluconeogenesis, pyruvate, propanoate, glycerophospholipid, galactose, fatty acid biosynthesis and metabolism, starch and sucrose metabolism that may be amplified in these patients. While the metabolites that were down-regulated in β -thalassemia patients showed abruption in glycerolipid, galactose, glyoxylate and dicarboxylate metabolism and fatty acid biosynthesis. Metabolites involved in dysregulation of these pathways are palmitic acid, lactic acid, sucrose, triethanolamine, glycerol and ethylene glycol. The detail results of pathway analysis are shown in Table 4

Pathway Name	Total Compound Present in Pathway	Hits	Raw p	−log(p)	Holm p	FDR	Impact
Up-regulated:							
Fatty acid elongation in mitochondria	27	1	0.0967	2.3362	1	1	0
Glycolysis or Gluconeogenesis	31	1	0.11029	2.2046	1	1	0
Pyruvate metabolism	32	1	0.11366	2.1745	1	1	0.13756
Propanoate metabolism	35	1	0.1237	2.0899	1	1	0
Glycerophospholipid metabolism	39	1	0.13694	1.9882	1	1	0
Galactose metabolism	41	1	0.14348	1.9415	1	1	0.01716
Fatty acid biosynthesis	49	1	0.16924	1.7764	1	1	0
Fatty acid metabolism	50	1	0.17241	1.7579	1	1	0.02959
Starch and sucrose metabolism	50	1	0.17241	1.7579	1	1	0.05947
Down-regulated:							
Glycerolipid metabolism	32	1	0.064781	2.7367	1	1	0.18847
Galactose metabolism	41	1	0.082382	2.4964	1	1	0
Fatty acid biosynthesis	49	1	0.097804	2.3248	1	1	0
Glyoxylate and dicarboxylate metabolism	50	1	0.099717	2.3054	1	1	0

Table 4. List of up-regulated and down-regulated pathways in β -thalassemia patients in comparison to healthy controls.

illustrating all matched pathways according to p-values from pathway enrichment analysis and pathway impact values from pathway topology analysis.

The pathways with considerable impact include glycerolipid, pyruvate, galactose, starch/sucrose and fatty acid metabolism while metabolites blameable for these deviations are glycerol, lactic acid, sucrose and palmitic acid respectively. We also noted that two pathways are found in both lists of pathways altered either due to increase or decrease of metabolites. One of them is fatty acid biosynthesis so it can be assumed that increase in palmitic acid is compensated by body with decrease in stearic acid to sustain the regulation of fatty acid biosynthesis pathways. Other is galactose metabolism in which increase sucrose is responsible for this pathway alteration which is compensated by decrease in glycerol. But this compensation is not sufficient enough by body because decreased glycerol has no such significant impression on this pathway as compared to increased sucrose. Therefore, it can be anticipated that diabetes an important complication of β -thalassemia is linked to imbalanced and aggravated galactose metabolism.

Conclusion

This study showed that genetic abnormalities in β -thalassemia also give rise to disturbance in metabolism of body that can be observed by alteration in serum metabolomic profile of β -thalassemia patients as compared to the profile of healthy group. Our research demonstrated that metabolite profiling by GC-EL-MS is a reproducible, sensitive and less invasive method that can be used for establishment of a profile distinguishing between β -thalassemia patients and healthy controls with a good sensitivity and specificity. A model was fabricated on forty significantly expressed metabolites precisely classifying β -thalassemia patients and healthy controls on external validation. In addition to this many important pathways are identified that were found to be impaired in β -thalassemia and may play role in disease progression. Moreover, our approach is the first to report differences in the serum metabolome between healthy and β -thalassemia patients, a molecular level understanding that can be used in improving treatment options for the sufferers as well as diagnosing phenotype of patients.

References

- Rivella, S. beta-thalassemias: paradigmatic diseases for scientific discoveries and development of innovative therapies. *Haematologica* **100**, 418–430, doi: 10.3324/haematol.2014.114827 (2015).
- Weatherall, D. J. The challenge of haemoglobinopathies in resource-poor countries. *British journal of haematology* **154**, 736–744, doi: 10.1111/j.1365-2141.2011.08742.x (2011).
- Weatherall, D. J. The inherited diseases of hemoglobin are an emerging global health burden. *Blood* **115**, 4331–4336, doi: 10.1182/blood-2010-01-251348 (2010).
- Giardine, B. *et al.* Updates of the HbVar database of human hemoglobin variants and thalassemia mutations. *Nucleic acids research* **42**, D1063–1069, doi: 10.1093/nar/gkt911 (2014).
- Ansari, S. H. *et al.* Efficacy of hydroxyurea in providing transfusion independence in β -thalassemia. *Journal of pediatric hematology/oncology* **33**, 339–343 (2011).
- Rund, D. & Rachmilewitz, E. Beta-thalassemia. *The New England journal of medicine* **353**, 1135–1146, doi: 10.1056/NEJMra050436 (2005).
- Hatairaktham, S. *et al.* Differential plasma proteome profiles of mild versus severe beta-thalassemia/Hb E. *Annals of hematology* **92**, 365–377, doi: 10.1007/s00277-012-1629-5 (2013).
- Zhang, A., Sun, H., Yan, G., Wang, P. & Wang, X. Mass spectrometry-based metabolomics: applications to biomarker and metabolic pathway research. *Biomedical chromatography: BMC* **30**, 7–12, doi: 10.1002/bmc.3453 (2016).
- Bonneau, E., Tetreault, N., Robitaille, R., Boucher, A. & De Guire, V. Metabolomics: Perspectives on potential biomarkers in organ transplantation and immunosuppressant toxicity. *Clinical biochemistry* **49**, 377–384, doi: 10.1016/j.clinbiochem.2016.01.006 (2016).

10. Stringer, K. A., McKay, R. T., Karnovsky, A., Quemerais, B. & Lacy, P. Metabolomics and Its Application to Acute Lung Diseases. *Frontiers in immunology* **7**, 44, doi: 10.3389/fimmu.2016.00044 (2016).
11. Izquierdo-Garcia, J. L. *et al.* A Metabolomic Approach to the Pathogenesis of Ventilator-induced Lung Injury. *Anesthesiology* **120**, 694–702, doi: 10.1097/aln.000000000000074 (2014).
12. Gao, Y. *et al.* Identifying early urinary metabolic changes with long-term environmental exposure to cadmium by mass spectrometry-based metabolomics. *Environmental science & technology* **48**, 6409–6418, doi: 10.1021/es500750w (2014).
13. Wishart, D. S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nature reviews. Drug discovery*, doi: 10.1038/nrd.2016.32 (2016).
14. Hart, C. D., Tenori, L., Luchinat, C. & Di Leo, A. Metabolomics in Breast Cancer: Current Status and Perspectives. *Advances in experimental medicine and biology* **882**, 217–234, doi: 10.1007/978-3-319-22909-6_9 (2016).
15. Cook, J. A. *et al.* Mass Spectrometry-Based Metabolomics Identifies Longitudinal Urinary Metabolite Profiles Predictive of Radiation-Induced Cancer. *Cancer research* **76**, 1569–1577, doi: 10.1158/0008-5472.can-15-2416 (2016).
16. Deng, L. *et al.* Abstract B50: Detecting colorectal cancer and polyps using nuclear magnetic resonance spectroscopy and mass spectrometry based metabolomics. *Molecular Cancer Research* **14**, B50–B50 (2016).
17. Hartwell, S. K., Srisawang, B., Kongtawelert, P., Christian, G. D. & Grudpan, K. Review on screening and analysis techniques for hemoglobin variants and thalassemia. *Talanta* **65**, 1149–1161, doi: 10.1016/j.talanta.2004.09.013 (2005).
18. Birndorf, N. I., Pentecost, J. O., Coakley, J. R. & Spackman, K. A. An expert system to diagnose anemia and report results directly on hematology forms. *Computers and biomedical research, an international journal* **29**, 16–26 (1996).
19. Edwards, R. L., Griffiths, P., Bunch, J. & Cooper, H. J. Compound heterozygotes and beta-thalassemia: top-down mass spectrometry for detection of hemoglobinopathies. *Proteomics* **14**, 1232–1238, doi: 10.1002/pmic.201300316 (2014).
20. Meerang, M. *et al.* Increased urinary 1,N6-ethenodeoxyadenosine and 3,N4-ethenodeoxycytidine excretion in thalassemia patients: markers for lipid peroxidation-induced DNA damage. *Free radical biology & medicine* **44**, 1863–1868, doi: 10.1016/j.freeradbiomed.2008.02.009 (2008).
21. Ebeid, S. A., Sadek, N. A., El-Benhawy, S. A. & Mahmoud, G. N. Plasma substance P and soluble P-selectin as biomarkers of β -thalassemia induced hypercoagulability. *Alexandria Journal of Medicine* **50**, 107–111 (2014).
22. Makis, A. *et al.* Adipocytokines are related to haemolytic and inflammatory biomarkers in sickle cell beta thalassaemia. *British journal of haematology* **163**, 142–144, doi: 10.1111/bjh.12464 (2013).
23. Kattamis, C. *et al.* Disturbances of biomarkers of iron and oxidant-antioxidant homeostasis in patients with beta-thalassemia intermedia. *Pediatric endocrinology reviews: PER* **8** Suppl 2, 256–262 (2011).
24. Elsayh, K. I., Zahran, A. M., El-Abaseri, T. B., Mohamed, A. O. & El-Metwally, T. H. Hypoxia biomarkers, oxidative stress, and circulating microparticles in pediatric patients with thalassemia in Upper Egypt. *Clinical and applied thrombosis/hemostasis: official journal of the International Academy of Clinical and Applied Thrombosis/Hemostasis* **20**, 536–545, doi: 10.1177/1076029612472552 (2014).
25. Arjmand, M. *et al.* Nuclear magnetic resonance-based screening of thalassemia and quantification of some hematological parameters using chemometric methods. *Talanta* **81**, 1229–1236, doi: 10.1016/j.talanta.2010.02.014 (2010).
26. Abdullah, U. Y. *et al.* The Free Alpha-Hemoglobin: A Promising Biomarker for β -Thalassemia. *Journal of Molecular Biomarkers & Diagnosis* **2014** (2014).
27. De Sanctis, V. *et al.* Growth and endocrine disorders in thalassemia: The international network on endocrine complications in thalassemia (I-CET) position statement and guidelines. *Indian journal of endocrinology and metabolism* **17**, 8–18, doi: 10.4103/2230-8210.107808 (2013).
28. De Sanctis, V. *et al.* Diabetes and Glucose Metabolism in Thalassemia Major: An Update. *Expert review of hematology* **9**, 401–408, doi: 10.1586/17474086.2016.1136209 (2016).
29. Musharraf, S. G., Mazhar, S., Choudhary, M. I., Rizi, N. & Atta ur, R. Plasma metabolite profiling and chemometric analyses of lung cancer along with three controls through gas chromatography-mass spectrometry. *Scientific reports* **5**, 8607, doi: 10.1038/srep08607 (2015).
30. Musharraf, S. G., Mazhar, S., Siddiqui, A. J. & Choudhary, M. I. Metabolite profiling of human plasma by different extraction methods through gas chromatography-mass spectrometry—An objective comparison. *Analytica chimica acta* **804**, 180–189 (2013).
31. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**, 27–30 (2000).
32. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic acids research* **44**, D457–462, doi: 10.1093/nar/gkv1070 (2016).
33. Johnson, C. H., Ivanisevic, J. & Siuzdak, G. Metabolomics: beyond biomarkers and towards mechanisms. *Nature reviews. Molecular cell biology*, doi: 10.1038/nrm.2016.25 (2016).
34. Voge, N. V. *et al.* Metabolomics-Based Discovery of Small Molecule Biomarkers in Serum Associated with Dengue Virus Infections and Disease Outcomes. *PLoS neglected tropical diseases* **10**, e0004449, doi: 10.1371/journal.pntd.0004449 (2016).
35. Cameron, S. J. *et al.* The metabolomic detection of lung cancer biomarkers in sputum. *Lung cancer (Amsterdam, Netherlands)* **94**, 88–95, doi: 10.1016/j.lungcan.2016.02.006 (2016).
36. Darghouth, D. *et al.* Pathophysiology of sickle cell disease is mirrored by the red blood cell metabolome. *Blood* **117**, e57–66, doi: 10.1182/blood-2010-07-299636 (2011).
37. Shamshirsaz, A. A. *et al.* Metabolic and endocrinologic complications in beta-thalassemia major: a multicenter study in Tehran. *BMC endocrine disorders* **3**, 4, doi: 10.1186/1472-6823-3-4 (2003).
38. Wishart, D. S. *et al.* HMDB 3.0—The Human Metabolome Database in 2013. *Nucleic acids research* **41**, D801–807, doi: 10.1093/nar/gks1065 (2013).
39. Wishart, D. S. *et al.* HMDB: a knowledgebase for the human metabolome. *Nucleic acids research* **37**, D603–610, doi: 10.1093/nar/gkn810 (2009).
40. Wishart, D. S. *et al.* HMDB: the Human Metabolome Database. *Nucleic acids research* **35**, D521–526, doi: 10.1093/nar/gkl923 (2007).
41. Chen, W. & Viljoen, A. Geraniol—a review of a commercially important fragrance material. *South African Journal of Botany* **76**, 643–651 (2010).
42. Abramczyk, H. & Brozek-Pluska, B. New look inside human breast ducts with Raman imaging. Raman candidates as diagnostic markers for breast cancer prognosis: Mammaglobin, palmitic acid and sphingomyelin. *Anal Chim Acta* **909**, 91–100, doi: 10.1016/j.aca.2015.12.038 (2016).
43. Wanachiwanawin, W. Infections in E-beta thalassemia. *Journal of pediatric hematology/oncology* **22**, 581–587 (2000).
44. Bajoria, R. & Chatterjee, R. Hypogonadotrophic hypogonadism and diminished gonadal reserve accounts for dysfunctional gametogenesis in thalassaemia patients with iron overload presenting with infertility. *Hemoglobin* **35**, 636–642 (2011).
45. Sjarif, D. R. *et al.* Glycerol kinase deficiency: residual activity explained by reduced transcription and enzyme conformation. *European journal of human genetics: EJHG* **12**, 424–432, doi: 10.1038/sj.ejhg.5201172 (2004).
46. Psychogios, N. *et al.* The human serum metabolome. *PLoS one* **6**, e16957, doi: 10.1371/journal.pone.0016957 (2011).

Acknowledgements

Authors are thankful to the patients and healthy subjects for giving consent for blood sample collection. In addition to that authors also acknowledge KEGG for allowing copyright of metabolic pathway images.

Author Contributions

S.G.M. contributed to the study design, provided conceptual and technical guidance, along with the laboratory equipment and expertise to conduct metabolomic analysis, contributed to the interpretation of the data results, and preparation and writing of the manuscript. A.I. managed the procurement of healthy blood samples, performed preparation of samples for metabolomics and statistical analysis, interpreted data and wrote the manuscript. S.H.A. provided important clinical expertise to this project, helped with the collection of clinical samples as well as in reviewing of manuscript. S.P. managed the procurement of disease group blood samples and data of samples. I.A.K. also contributed in reviewing the manuscript. A.J.S. assisted with the processing of the samples for metabolomic analysis, statistical analysis and reviewing of manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Musharraf, S. G. *et al.* β -Thalassemia Patients Revealed a Significant Change of Untargeted Metabolites in Comparison to Healthy Individuals. *Sci. Rep.* 7, 42249; doi: 10.1038/srep42249 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017