Clinica Chimica Acta 477 (2018) 39–47

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/cca

Serum metabolic profiling of type 2 diabetes mellitus in Chinese adults using an untargeted GC/TOFMS

Huijuan Wang[a,1], Hailong Zhang[b,1], Ling Yao[c], Longzhen Cui[c], Lei Zhang[c], Biao Gao[d], Wen Liu[c], Dongdong Wu[c], Mingliang Chen[c], Xue Li[c], Ailing Ji[a,*], Yanzhang Li[a,*]

[a] Joint National Laboratory for Antibody Drug Engineering, School of Basic Medicine, Henan University, Kaifeng, Henan 475004, China
[b] Joint National Laboratory for Antibody Drug Engineering, Henan Key Laboratory of Cellular and Molecular Immunology, Henan University, China
[c] School of Basic Medicine, Henan University, Kaifeng, Henan 475004, China
[d] Kaifeng Central Hospital, Kaifeng, Henan 475000, China

ARTICLE INFO

Keywords:
Serum
Metabolic profiling
T2DM
Chinese
GC/TOFMS

ABSTRACT

Background: Type 2 diabetes mellitus (T2DM) is a huge burden in China. The Chinese patients with T2DM have several special clinical characteristics. Metabolomics studies predominantly have identified several distinguishing metabolites associated with T2DM in Western ancestry population. However, few previous metabolomics studies were conducted in Chinese populations.

Methods: We performed untargeted serum metabolic profiling between 30 T2DM patients and 30 healthy controls based on GC/TOFMS. Multivariate data analyses were applied to identify the distinguishing metabolites.

Results: Excellent separation was obtained between the two cases. And overall 54 distinguishing metabolites were identified with VIP $>$ 1 and $P < 0.05$, which were involved in metabolic pathways of amino acid, carbohydrate, lipids, membrane transport and nucleotides. To further analyze the correlation between the identified metabolites and T2DM, 17 metabolites were selected with FC $>$ 2.0, including gentisic acid, citraconic acid, succinic acid, 2-hydroxybutanoic acid and 3-hydroxy-L-proline, the corresponding FC were respectively 5.44, 2.21, 2.10, 2.21 and − 2.04.

Conclusion: Our results demonstrated that untargeted GC/TOFMS-based metabolic approach processed well performance to identify serum distinguishing metabolites of T2DM in Chinese adults, which may be as potential biomarkers in diagnose and treatment of diabetes. And the results also provided new insight into understand the underlying molecular mechanism.

1. Introduction

Diabetes mellitus (DM) is a constellation of chronic metabolic disorder in the endocrine system, which is characterized by hyperglycemia and impaired insulin sensitivity [1]. It is regarded as one of the most important threat to public health and civilization. According to epidemiological analysis, the number of people with DM was 387 million worldwide in 2014, and the number is expected to increase to 592 million by 2035, 90% of which are due to type 2 diabetes mellitus (T2DM) [2]. In China, the prevalence of T2DM has increased speedily because of environmental influences and diet and lifestyle modifications during recent decades [3,4]. The rate of the Chinese population with T2DM from < 1% in 1980 increased to 11.6% in 2010 [5], and the prevalence of T2DM in China is expected to continue accelerating [6]. The complications of T2DM are very complex, such as neuropathy, retinopathy and nephropathy may be already present before diagnosis. As conventional biomarkers, Random glucose concentrations, body mass index (BMI), fasting glucose (FPG), 2 h post prandial plasma glucose (2 h PG) or haemoglobin A1c (HbA1c) are widely used in clinic, but they are still imperfect and most of them fail to provide the whole insight of disease pathophysiology. Therefore, understanding the biological mechanism underlying T2DM and identifying novel biomarkers are particularly important to develop the effective strategy to prevent, diagnose and treat T2DM.

Abbreviations: T2DM, type 2 diabetes mellitus; GC/TOFMS, gas chromatograph system coupled with a Pegasus HT time-of-flight mass spectrometer; PCA, principal component analysis; PLS-DA, partial least squares discriminant analysis; OPLS, orthogonal projections to latent structures-discriminant analysis; VIP, variable importance in the projection; FC, fold change; FPG, fasting glucose; 2 h PG, 2 h post prandial plasma glucose

* Corresponding authors.
E-mail addresses: ailingji@163.com (A. Ji), 10190115@vip.henu.edu.cn (Y. Li).
[1] These authors contributed equally to this work and should be considered co-first authors.

https://doi.org/10.1016/j.cca.2017.11.036
Received 6 September 2017; Received in revised form 27 November 2017; Accepted 28 November 2017
Available online 02 December 2017
0009-8981/

Metabolomics is an emerging analytical technology and has already become a crucial strategy to measure the endogenous metabolites for mechanistic studies of metabolic diseases in cells, biofluids or tissues [7]. As the end point of the "-omics" cascades, metabolomics provides the most integrated profile of biological status at a specified time. There is a mass of evidences demonstrate that metabolomics can find out novel biomarkers for disease diagnose and provide significantly insight into the pathogenic mechanism underlying various diseases [8–11]. Metabolomics is a high-sensitivity and high-throughput approach which has been increasingly used to discover biomarkers of T2DM [6]. The susceptibility and pathophysiology of T2DM would be affected by racial/ethnic differences [3]. To date, most studies have been predominantly conducted in Western ancestry population. To our knowledge, a few studies are conducted in the global metabolomics profiling of T2DM in China [6,12].

In the present study, we conducted the serum metabolic profiling between T2DM patients and healthy controls within Henan province located in the central of China using gas chromatograph system coupled with a Pegasus HT time-of-flight mass spectrometer (GC/TOFMS), with the wish of identifying novel biomarkers of T2DM diagnose and understanding of the molecular mechanism involved in disease progression.

2. Materials and methods

2.1. Study subjects and serum collection

The study was approved by the ethical committee of Huaihe Hospital of Henan University. Informed consent was obtained from each participant before blood collection. A total of 30 T2DM subjects and 30 healthy controls were enrolled in this study from 2014 to 2015. T2DM was characterized based on symptoms and at least one of the following criteria: random glucose concentrations ≥ 11.1 mmol/L, fasting glucose (FPG) ≥ 7.0 mmol/L or 2 h post prandial plasma glucose (2 h PG) ≥ 11.1 mmol/L or haemoglobin A1c (HbA1c) ≥ 6.5%; The volunteer joined this study also provided physical examination information.

All venous blood samples (10 mL) were taken from all participants in the morning after overnight fasting for at least 8 h. After collection, the part of the samples was used for clinical chemistry tests. The remaining samples were immediately kept on ice and processed within 6 h to get serum, which then were stored at − 80 °C for long-term storage until metabolomics analysis.

2.2. Chemicals

L-2-Chlorophenylalanine, as an internal standard, was purchased from Shanghai Hengbai Biological Technology Co. Ltd. (China). Bis-(trimethylsilyl)-trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS) was obtained from REGIS Technologies. Inc. (USA). Methoxylamine was supplied by Sigma-Aldrich (USA). Methanol (high performance liquid chromatography grade, HPLC) was obtained from Merck KGaA (Germany). All other chemicals were supplied by Shanghai Ampu Company (China). Double-distilled water was supplied by Milli-Q Reagent Water System (USA).

2.3. Serum sample preparation and pretreatment

Serum samples in our study were stored at − 80 °C until GC/TOFMS analysis. After thawing, 100 μL sample was put into 1.5 mL EP tube. To extract the serum metabolites, 350 μL methanol (HPLC grade) was added. As an internal standard 30 μL L-2-chlorophenylalanine (0.1 mg/mL stock in ddH_2O) were also added into the tube. Then the mixture was vortexed vigorously for 10 s. The sample was centrifuged for 10 min at 12000 rpm at 4 °C. The 400 μL supernatant was carefully transferred into a fresh 2 mL GC/MS glass vial to dry in a vacuum concentrator without heating. Subsequently, 60 μL of methoxyamination reagent (20 mg/mL in pyridine) was added to dissolve the sample. After shaking, the solution was incubated for 2 h at 37 °C. And then 80 μL BSTFA (1% TMCS, v/v) was added promptly. The mixture was vortexed gently and incubated at 70 °C for 1 h. 10 μL FAMEs (standard mixture of fatty acid methyl esters, C8–C16: 1 mg/mL; C18–C30: 0.5 mg/mL in chloroform) was added into the sample when the temperatures get down to the room temperature. After vortexed gently, the 1 μL extraction was supplied for GC/TOFMS analysis.

2.4. GC/TOFMS analysis

According to previous studies [13,14], the serum metabolic profiling was analyzed by an Agilent 7890A gas chromatograph system coupled with a Pegasus HT time-of-flight mass spectrometer, which was equipped with a DB-5MS capillary column coated with 5% diphenyl cross-linked with 95% dimethyl polysiloxane (30 m × 250 μm diameter, 0.25 μm thickness) (J&W Scientific, USA). A 1 μL derivative was injected in splitless mode. Helium (purity > 99.999%) was used as a carrier gas with the front inlet purge flow was 3 mL/min. And the gas flow rate through the column was 1 mL/min. The initial column temperature was kept at 50 °C for 1 min, then increased to 330 °C at a rate of 10 °C/min, and maintained at 330 °C for 6 min. The injection temperature was 280 °C. The transfer line temperature was also 280 °C and the temperature of ion source was 220 °C. The energy was − 70 eV in electron impact mode. The MS data were acquired in full-scan mode with the range of 85–600 *m/z* at a rate of 20 spectra per second after a solvent delay of 366 s.

2.5. Multivariate data analysis

Chroma TOF 4.3 × software (LECO Corporation) and LECO-Fiehn Rtx5 database were used to preprocess the raw GC/MS data, which included peaks extracting, the data baselines filtering and calibrating, peak alignment, deconvolution analysis, peak identification and integration of the peak area [15]. The RI (retention time index) method was used in the peak identification, and the RI tolerance was 5000. The raw 625 peak between type 2 diabetes mellitus and healthy controls were preprocessed. Firstly, missing values of raw data were filled up by half of the minimum value. In addition, the filtered data was normalized using internal standard normalization method. Finally, 329 peaks were screened out after pretreatment and the normalized peak were fed to SIMCA-P 13.0 software package (Umetrics, Umea, Sweden) for principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA) and orthogonal projections to latent structures-discriminant analysis (OPLS). PCA model showed the separation of the origin MS data. PLS-DA were validated for a better separation. In order to guard against overfitting, the default 7-round cross-validation procedure was applied with exclusion of 1/7th samples in each round. In order to further validate the model, permutation test (200 times) was proceeded. The low Q2 values indicated the robustness of the models, which showed a low risk of over fitting and reliable. To select the important peak between two groups, OPLS model was employed [11,16]. A loading plot was constructed, which showed the contribution of variables to difference between two groups. To refine this analysis, the metabolites with the VIP (variable importance in the projection) values exceeding 1.0 were considered as changed ones. At the same time, Student's *t*-test (*t*-test), was also used to assess the significance of the remaining metabolites. $P < 0.05$, variables were retained between two comparison groups. Moreover, q-value for multiple testing by controlling for FDR (false discovery rate) was also supplied. Only both meeting VIP > 1 and $P < 0.05$, the metabolites were identified as distinguishing ones between T2DM and healthy subjects. The LECO/Fiehn Metabolomics Library was used to identify the compounds. It will give a similarity value for the compound identification accuracy. If several peaks had the same name, only the peak with the highest similarity was