Systems diagnosis of chronic diseases, explored by metabolomics and ultra-weak photon emission

He, M.

Chapter 6

Traditional Chinese medicine-based subtyping of early-stage type 2 diabetes using plasma metabolomics combined with ultra-weak photon emission

Min He, Eduard P. A. van Wijk, Mei Wang, Slavik Koval, Mengmeng Sun, Thomas Hankemeier and Jan van der Greef

Submitted for publication.

Abstract

Ethnopharmacological relevance: The prevalence of type 2 diabetes mellitus (T2DM) is increasing rapidly worldwide. Because of the limited success of generic interventions, focus has shifted toward personalized strategies, particularly in early stages of the disease. Traditional Chinese medicine (TCM) is based on a systems view combined with personalized strategies and has improved our knowledge with respect to personalized diagnostics. From a systems biology perspective, this understanding can be improved in order to yield a biochemical basis for such strategies, for example using metabolomics combined with other system-based diagnostic methods such as ultra-weak photon emission (UPE). In this respect, UPE has been used successfully to support TCM-based subtyping. Combining these technologies will further support TCM-based subtyping of diseases such as T2DM.

Aim of the study: The aim of this study was to investigate the feasibility of using plasma metabolomics to stratify the following TCM-based subtypes: Qi-Yin deficiency, Qi-Yin deficiency with dampness, and Qi-Yin deficiency with stagnation. Furthermore, we studied the relationship between plasma metabolomics and UPE with respect to TCM-based subtyping in order to obtain biochemical information for further interpreting disease subtypes.

Materials and methods: Plasma samples obtained from 44 subjects were extracted and analyzed using both liquid chromatography/tandem mass spectrometry and gas chromatography/tandem mass spectrometry. We then profiled various classes of metabolites, including amine metabolites, organic acids, sugars, and lysophosphatidic acid–derived metabolites, as well as lipids, including sphingomyelin phosphatidylcholine, phosphoethanolamine, lyso-

phosphatidylcholine, lyso-phosphoethanolamine, , cholesterol esters and triglycerides. Multivariate analysis (principal component analysis and orthogonal projections to latent structures discriminant analysis) was used to analyze the metabolomics profiles and to study TCM-based stratification. Finally, Spearman's rank correlation-based networks were used to correlate the metabolites with the UPE parameters.

Results and discussion: Principal component analysis of plasma metabolites revealed differences among the TCM-based pre-T2DM subtypes. Relatively high levels of lipids (e.g., triglycerides and cholesterol esters) were important discriminators of two of the three subtypes and may be associated with a higher risk of cardiovascular disease. Correlation networks revealed that plasma metabolomics and UPE yielded similar TCM-based subtypes. Finally, plasma metabolomics data indicate that the lipid profile is an essential component captured by UPE with respect to stratifying subtypes of T2DM.

Conclusions: Metabolic differences exist among different TCM-based subtypes of pre-T2DM, and profiling plasma metabolites can be used to discriminate among these subtypes. Plasma metabolomics provides biochemical insights into system-based UPE measurements.

Key words: Type 2 diabetes mellitus, plasma metabolites, disease subtypes, ultra-weak photon emission, correlation networks

1. Introduction

Type 2 diabetes mellitus (T2DM) is a chronic, devastating complex disease. T2DM is characterized by increased fasting plasma glucose levels, impaired postprandial insulin secretion, decreased insulin sensitivity, and impaired pancreatic beta-cell function [1]. In addition, patients with T2DM have increased levels of inflammatory factors such as TNFα, IL-6, IL-8, and reactive active species [2], [3], altered levels of hormones, peptides, proteins, and enzyme activity, as well as other metabolic perturbations [4]. Striking, nearly all of these metabolic changes are often present years before the patient presents with clinical symptoms leading to a diagnosis of T2DM [5], [6].

Based on epidemiology studies, an estimated 285 million individuals are affected by diabetes worldwide, and this number continues to increase [7]. Furthermore, this number is likely an underestimate, as many individuals are not diagnosed in an early stage due to insufficient knowledge regarding the multi-symptom relationships at a systems level [8], [9]. Receiving a diagnosis only in a later stage of diabetes—together with the severe complications associated with disease progression—can lead to high costs and can reduce the efficacy of treatment [10]. For example, long-term dysglycemia increases the risk of severe complications such as hypertension, blindness, renal failure, and cardiovascular disease [11], [12]. These complications reduce quality of life and are a major cause of morbidity, hospitalization, and mortality among patients with diabetes. Current diagnostic tests are based primarily on a single screening tool such as the oral glucose tolerance test or measuring fasting plasma glucose. Understanding the symptoms that develop in an early stage of the disease and developing indicators of disease progression would likely contribute to improving both prevention and treatment strategies, including strategies based on changes in lifestyle. Moreover, treatments based on generic observations—which have led to the notion of one drug-one target-one disease (or one-size-fits-all)—are extremely limited,

particularly in early stages of the disease. Therefore, system-based approaches are needed in order to achieve personalized approaches.

Integrative holistic forms of medicine such as traditional Chinese medicine (TCM) provide descriptions of disease syndromes and subtypes at a systems level, including descriptions that can be used to diagnose early syndromes of chronic diseases. Such descriptions can be used as a guide or reference in order to achieve personalized medicine. In this respect, TCM has provided descriptions of pre-T2DM syndromes, indicating its potential for helping develop personalized medicine [13], [14]. To bridge TCM with Western medicine, evidence-based scientific data is needed at the biochemical level. Thus, modern systems biology research—including metabolomics—is a promising approach for exploring the biochemistry underlying TCM subtyping.

Metabolic disorders are often present for years before the appearance of clinical disease, and metabolomics is a widely used technique for predicting and diagnosing disease [15]. Metabolomics provides a comprehensive profile of small molecular metabolites in biological systems and can be used as a readout of the organism's physiological status [16]. In principle, this approach is well suited to studying complex TCM-based diagnostics. Metabolomics is generally performed on fluids such as blood, urine, and cerebrospinal fluid. Urine is commonly used for metabolomics, as it easily obtained, contains information regarding the excretion of products, and can reflect how metabolic processes change during the disease process. Several studies have used urine metabolomics to explore TCM-based diagnostics and T2DM syndrome subtypes [17], [18]. In addition to urine, blood also contains information regarding the body's regulatory status and dynamics. Thus, performing metabolomics on different fluids can provide complementary information, thereby improving our understanding of T2DM. An explorative study at TNO (https://clinicaltrials.gov/ct2/show /NCT00469287) was designed in which 44 pre-T2DM subjects received a diagnosis by a panel of three TCM-trained

physicians [17], and we explored these TCM-based subtypes using plasma metabolomics.

Recently, a sensitive, non-invasive technique has been proposed for supporting TCM-based diagnostics [19]. This technique, called ultra-weak photon emission (UPE), is used to measure spontaneous photon emissions from the skin's surface [20]. Because UPE reflects the body's physiological and pathological status, it represents a promising tool for use in clinical diagnostics at a systems level [21], [22]. The underlying biochemistry of UPE is related to metabolism and is correlated with reactive oxygen species in oxidative metabolic processes [23]–[26]. Although the use of UPE properties for characterizing TCM-based diagnostics has been summarized previously [19], [20], [27], further understanding of the molecular basis of UPE is needed. Therefore, combining metabolomics with TCM-based diagnostics can be used to investigate the biological meaning of UPE and to explore the added value of each technology. Importantly, UPE was used previously to subtype the same cohort of 44 subjects with pre-T2DM [27], thereby enabling us to study the correlation between UPE and plasma metabolomics.

2. Materials and Methods

2.1 Inclusion criteria for the selection of pre-diabetic subjects and the diagnosis of syndrome subtypes based on TCM

The recruitment of subjects and the diagnosis of pre-T2DM subtypes by TCM-trained physicians were described previously [17]. In brief, clinical parameters were obtained from 44 male Dutch subjects who met the following inclusion criteria: 30-70 years of age, body mass index of 26-35 kg/m^2, and a fasting glucose level of 6.1-6.9 mmol/L. No other clinical abnormalities or evidence of diabetic complications were detected. The subjects were then diagnosed separately in a blinded study by three TCM-certified physicians with at least five years of training in TCM and at least ten years of clinical experience. Three categories were based on TCM-based diagnostic terms, and 85% consensus was reached among the three CM physicians with respect to diagnosing the subjects. These three categories are defined as follows: QYD (Qi-Yin deficiency, n=15 subjects), QYD_Damp (Qi-Yin deficiency with dampness, n=20 subjects), and QYD_Stag (Qi-Yin deficiency with stagnation, n=9 subjects). Blood samples were collected after overnight fasting and used for the metabolomics study. In addition, UPE was measured from the palmar and dorsal surfaces of both hands.

2.2 Ethics statement

This explorative study was designed and conducted by TNO (Zeist, the Netherlands; https://clinicaltrials.gov/ct2/show/NCT00469287) and was approved by the Medical Ethics Committee of Tilburg (METOPP).

2.3 Data acquisition

2.3.1 Plasma metabolomics profiling

Metabolic profiles were measured by the Netherlands Organization for Applied Scientific Research (TNO, Zeist, the Netherlands). Heparinized blood samples were collected, and plasma was obtained by centrifugation (2000×*g* at 4ºC for 15 min). The plasma samples were aliquoted and stored at -20ºC prior to metabolite extraction and mass spectrometry.

Using a gas chromatography/mass spectrometry (GC-MS) platform, a large variety of metabolic classes were measured, including amine metabolites, organic acids, sugars, and lysophosphatidic acid (LPA)-derived metabolites. The details of the extraction and the GC-MS analysis protocol have been published previously [28]. In brief, 100-µl aliquots of plasma were spiked with a mixture of internal standards (ISTDs) and deproteinized with methanol. After centrifugation, the supernatant was transferred to a new sample vial for evaporation and two-step derivatization. The derivatized extracts were then analyzed using an Agilent 6890 gas chromatograph on a DB5-MS capillary column (30 m × 250 µm i.d., 0.25-µm film thickness; J&W Scientific, Folsom, CA) coupled to an Agilent 5973 mass selective detector; helium was used as the carrier gas at a flow rate of 1.7 ml/min for temperature-programmed gradient chromatographic separation. The raw data were pre-processed and exported using ChemStation G1701CA software (version D.01.02, Agilent), providing response ratios to the appropriate internal ISTD for each metabolite; these ratios were used for further statistical analysis.

For liquid chromatography/tandem mass spectrometry (LC-MS) lipid measurements, seven classes of lipids, including both polar lipids—such as phosphatidylcholine, phosphoethanolamine, lyso-phosphatidylcholine, lyso-phosphoethanolamine, and sphingomyelin —and non-polar lipids—such as cholesterol esters and triglycerides—were investigated using targeted analysis as reported previously by van Wietmarschen et al. [29] and Draisma et al. [30]. In brief, 10-µl aliquots of plasma were deproteinized by the addition of isopropanol containing a mixture of ISTDs. The lipids were separated and analyzed using a TSQ Quantum Discovery Triple Quad mass spectrometer coupled to a Surveyor

MS HPLC system on an Alltech Prosphere C4 300Å column (150 x 3.2 mm, particle size of 5 μm; Alltech, Lexington, KY) in combination with a Symmetry 300 C4 guard column (2.1 × 10 mm, particle size of 3.5 μm; Waters, Milford, MA) in positive ionization mode. The peak areas of the target lipids were integrated, and raw data were exported using LCQuan software (version 2; Thermo Fisher Scientific, Waltham, MA), yielding response ratios to the appropriate internal ISTD for each metabolite; these ratios were used for further statistical analysis.

During the GC-MS and LC-MS experiments, quality control (QC) samples were prepared by pooling equal amounts of plasma from each sample, then dividing the pooled samples into aliquots; these QC samples were used to check the performance of the LC-MS platform as well as to identify temporal trends in the acquired data. The relative standard deviation (RSD) of each target peak in the QC samples was used to confirm the quality of the data acquired from each analytical platform.

2.3.2 UPE measurements

UPE signals were measured from the same cohort of 44 subjects. A photomultiplier system (provided by Meluna Research B.V., Geldermalsen, the Netherland) with two detecting heads located at the top of a dark chamber was used to measure UPE. Each detecting head contains a 9558QB photomultiplier tube within a spectral sensitivity range of 190-650 nm (Electron Tubes Enterprises Ltd., Ruislip, UK) and an electronically controlled shutter. The dark chamber was maintained at 20±1.0°C. The settings used to measure UPE have been described previously [31], [32]. All measurements were controlled automatically via computer-driven software. UPE signals were measured at the following four hand surfaces: left dorsal (LD), right dorsal (RD), left palm (LP), and right palm (RP).

2.4 Data preprocessing and statistical analysis

2.4.1 Metabolomics data processing and analysis

Before performing a statistical analysis on the metabolomics data, the log-transformed dataset was processed using various scaling options (i.e., autoscaling, range scaling, and pareto-scaling) using the online software package MetaboAnalyst 3.0 (http://www.metaboanalyst.ca/) [33]. The pareto-scaling approach (mean-centered and scaling by the square root of the standard deviation of each variable) was chosen because it provided the best grouping performance, consistently explaining the largest variabilities when considering the same number of principal components (both 2D and 3D) [34]–[36]. Preliminary selection of variables prior to multivariate analysis is needed in order to: *i*) limit the dataset of variables for reliably separating the sample groups; *ii*) remove irrelevant and/or confounding variables; and *iii*) decide which variables to retain for the multivariate analysis; however, this selection is not needed in order to identify potential biomarkers, which has been applied in metabolic profiling studies [37], using *p*-values obtained from a one-way analysis of variance (ANOVA) ($p<0.1$) in GC-MS and LC-MS. Multiple comparisons, including principal component analysis (PCA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA), were conducted using MetaboAnalyst 3.0, which provides standard validation information, including cross-validation and a permutation test to prevent over-fit of the models to the data [33].

2.4.2 Acquisition of UPE data and derived parameters

From a 50-ms bin, the following ten UPE properties were calculated from all four hand surfaces: strength, FF0, FF1, FF2, alpha, gamma, theta, phi, SSI, and SSR [31], [32], [38]. Thus, a total of 40 UPE parameters were obtained from each subject.

2.4.3 Correlation analysis

The statistics software package R (version 3.0.3) was used to calculate Spearman's rank correlation coefficient in order to examine the relationship between the

metabolites and UPE parameters. A graphical overview of the correlation networks was created using CytoScape version 3.3.0 (http://www.cytoscape.org) with the MetScape plugin [39], [40]. Positive and negative correlations are indicated by positive and negative values of r, respectively.

3. Results and Discussion

3.1 Subtyping based on plasma metabolomics

TCM-based diagnostics is based on several standard diagnostic steps, including inspection, listening and smelling, inquiry and question, and palpation. The outcomes from these steps are combined to create an individual profile, which is used to establish a diagnosis. In this study, 26 variables were determined using TCM-based diagnostics [17]. From this exploratory study, plasma samples were used to obtain evidence-based information that was used to help subtype the pre-T2DM subjects.

We used two validated metabolomics methods based on GC-MS and LC-MS. GC-MS yielded 147 untargeted metabolites, and LC-MS yielded 110 targeted metabolites; all of these metabolites were included in the total metabolomics profile. The metabolites detected by GC-MS included various metabolic classes, but primarily included amine metabolites, organic acids, sugars, and fatty acids such as LPA and LPA-derived metabolites. The metabolites detected by LC-MS included seven classes of lipids, including both polar lipids such as phosphatidylcholine, phosphoethanolamine, lyso-phosphatidylcholine, lyso-phosphoethanolamine, and sphingomyelin and non-polar lipids such as cholesterol esters (ChEs) and triglycerides (TGs). Given the relatively small number of subjects (44) compared to the large number of total variables (257), a first step in selecting variables was required before proceeding with a multivariate analysis; this step allowed us to optimize the variable/object ratio for discriminant type approaches, and it allowed us to remove potential irrelevant and/or confounding variables [37]. A total of 32 preliminary variables were selected based on an ANOVA analysis ($p<0.1$); these variables included 14 plasma metabolites identified by GC-MS and 17 plasma lipids identified by LC-MS. These variables were then used for subsequent multivariate analyses, including PCA, Partial least squares discriminant analysis (PLSDA), and OPLS-DA (see S-table 1 and S-fig. 1).

The first step in our analysis focused on investigating whether plasma metabolomics could be used to discriminate between the three TCM-based syndrome subtypes of pre-T2DM (i.e., QYD vs. QYD_Damp, QYD vs. QYD_Stag, and QYD_Damp vs. QYD_Stag). A 3D PCA plot was used to visualize the natural distribution of the three groups in 3-dimensional space [37], [41]. The first three principal components analyzed described 66.5% of the total variance in the plasma metabolome (Fig. 1). We found no large distance between the three subtypes reflected by PCA, which is not surprising given that their TCM-based diagnostic patterns are all-linked (interrelated) and TCM-based syndromes subtypes are not independent but with dynamic changes towards different direction [13], [14]. However, we did observe tendency of clusters within the subtypes, with minor overlap in the PCA analysis.

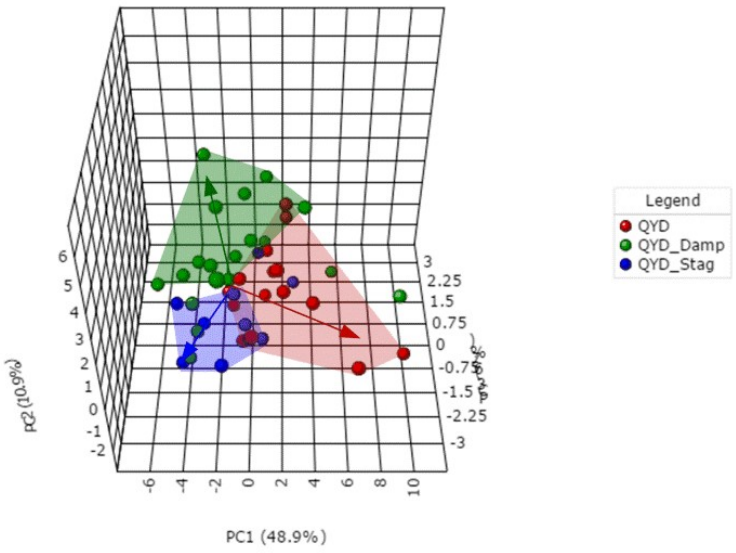


Fig. 1: 3D PCA score plot based on plasma metabolite profiling, acquired and integrated from GC-MS and LC-MS, for visualizing clusters of the three pre-T2DM subtypes (QYD, QYD_Damp and QYD_Stag).

Next, we used supervised models, including LDA, PLSDA, and OPLS-DA, in order to identify relevant plasma metabolites (S-fig. 2). The OPLS-DA model provided the highest R^2 and Q^2 values and was therefore used to identify the most relevant variables based on score plots [42][43]. Furthermore, permutation tests with 1000 iterations (p<0.05) showed a good performance of the model. Fig. 2 shows the OPLS-DA score plots for the first two principal components between each pair of subtypes (see also S-fig. 3).

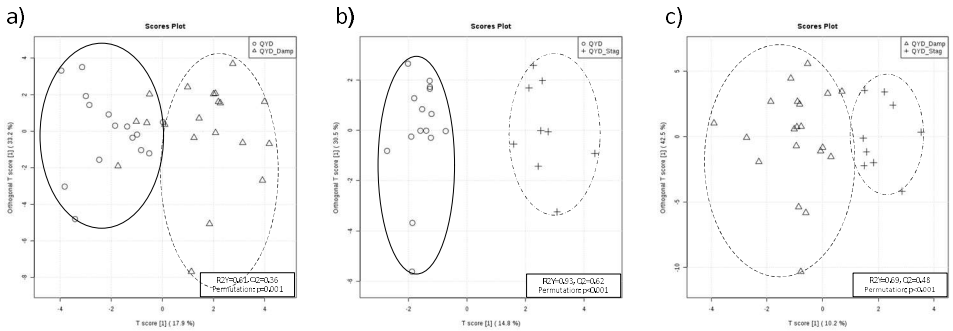


Fig. 2: OPLS-DA score plots of plasma metabolite profilling (integrated from LC-MS and GC-MS) for comparing differences between each pair of subtypes. a) QYD vs. QYD_Damp; b) QYD vs. QYD_Stag; and c) QYD_Damp vs. QYD_Stag.

Table 1 summarizes the relevant metabolites (defined as the combination of covariance |p[1]|>0.7 and correlation coefficient |p(corr)|>0.3[43]) for each pair of groups, together with their contribution between each pair of subtypes (QYD vs. QYD_Damp, QYD vs. QYD_Stag, and QYD_Stag vs. QYD_Damp). As shown in Table 1, 15 of the 18 metabolites that contributed to the differentiation between QYD and QYD_Damp are long-chain non-polar lipids (11 TGs and 4 ChEs); these metabolites were higher in the QYD_Damp group than in the QYD group. Fourteen of these same metabolites (10 TGs and 4 ChEs) were also higher in the QYD_Stag group than in the QYD group. Thus, we conclude that an increase in long-chain non-polar lipids is associated with the QYD_Damp and QYD_Stag groups.

The physiological mechanisms that underlie the early phases of T2DM have been linked to lifestyle issues such as the consumption of a diet high in fat and calories [44]–[47], which is similar to chronic fatigue syndrome and/or mild inflammatory status [17]. Triglycerides are the precursors of phospholipids, which are the building blocks of cell membranes and play an important role in energy homeostasis. Cholesterol esters are a stored form of cholesterol that is normally exported as a high-density lipoprotein (HDL) and returned to the liver. High levels of cholesterol and triglycerides (hypercholesterolemia and hypertriglyceridemia, respectively) are associated with fat accumulation, atherosclerosis, and cardiovascular disease [48], [49]. Therefore, patients in the pre-T2DM subgroups QYD_Damp and QYD_Stag may have an increased risk of developing atherosclerosis and/or cardiovascular disease in a later disease stage.

Table 1: List of relevant metabolites identified by OPLS-DA

QYD_Damp. vs. QYD		**QYD_Stag. vs. QYD**		**QYD_Stag. vs. QYD_Damp.**	
Metabolite	Change	Metabolite	Change	Metabolite	Change
C52_5_TG	↑	C22_5_ChE	↑	Beta-Alanine	↓
C54_6_TG	↑	C54_7_TG	↑	6926ukx10*	↓
C54_5_TG	↑	C54_6_TG	↑	1-Methylhistidine % 10227\01.03 uk x 45*	↓
C54_7_TG	↑	C58_10_TG	↑	31944uk05*	↓
C56_8_TG	↑	C52_6_TG	↑		
C56_7_TG	↑	C56_8_TG	↑		
C56_9_TG	↑	C18_3_ChE	↑		
C58_8_TG	↑	C52_5_TG	↑		
1-Methylhistidine % 10227\01.03 uk x 45*	↑	C22_6_ChE	↑		
C58_9_TG	↑	C56_7_TG	↑		
C52_6_TG	↑	C56_9_TG	↑		
C18_3_ChE	↑	C20_3_ChE	↑		
C16_0_ChE	↑	C54_5_TG	↑		
C22_6_ChE	↑	C58_9_TG	↑		
C20_3_ChE	↑	31944uk05	↓		
Creatinine	↑				
1-Palmitoyl-L-alpha-lysophosphatidic acid	↓				
1-Stearoyl-sn-glycero-3-phosphocholine	↓				

↑, increase; ↓, decrease.
*, Structural unidentified metabolites in GC-MS untargeted measurement.

Although TGs and ChEs were increased in both the QYD_Damp and QYD_Stag groups relative to the QYD group, these two groups had several metabolic differences (Table 1). The relatively lower levels of amine metabolites in the QYD_Stag group (and/or the relatively higher levels in the QYD_Damp group) may suggest that the difference between the QYD_Stag and QYD_Damp subtypes is based primarily in differences in the TCA cycle and/or muscle catabolism processes [50]. In summary, 23 metabolites contribute to the stratification of pre-T2DM subtypes. Thus, different subtypes of pre-T2DM may be discriminated based on differences in plasma metabolomics, including plasma lipids and amine metabolites.

Previously, Wei, et al. reported that urine metabolomics can be used to reflect changes in carbohydrate metabolism and renal function in patients with QYD_Stag syndrome; specifically, two of the three TCM-based subtypes could be stratified [17]. In contrast, plasma metabolomics provides stratification among the three subgroups, which is likely due to the use of a lipidomics platform, which measures a class of compounds that cannot be measured using urine metabolomics. This finding suggests that measuring lipid metabolomics is important for accurately subtyping pre-T2DM.

3.2 Correlation between metabolomics and UPE

We also measured UPE in our cohort of subjects with pre-T2DM. Stratification of the three TCM-based syndrome subtypes using 16 UPE parameters has been studied previously [27]. Given that both plasma metabolomics and UPE can stratify subjects into pre-T2DM subgroups, plasma metabolomics data may be used to obtain biochemical insight into UPE [51]–[53]. To explore the relationship between these two approaches, we used Spearman's rank correlation coefficient to establish a correlation-based metabolite-to-UPE network. Such a correlation

network may provide additional information that may further stratify disease subtypes and may provide a biochemical interpretation of UPE parameters.

We generated correlation networks between the 23 metabolites and 16 UPE parameters that contributed to the stratification of subtypes in order to visualize the most relevant correlations related to the three subtypes (Figure 3). These networks revealed clearly distinct distributions of UPE-to-metabolite correlations between the three subtypes. Specifically, the QYD_Damp subtype contained relatively few correlations, whereas the QYD and QYD_Stag subtypes contained relatively more positive and negative correlations, respectively. Moreover, although clear links are visible between UPE parameters and specific classes of metabolites (e.g., TGs and ChEs), the correlations differ among the subtypes. The differences between the three networks provide a clear distinction between the subgroups and might serve as an additional diagnostic tool.

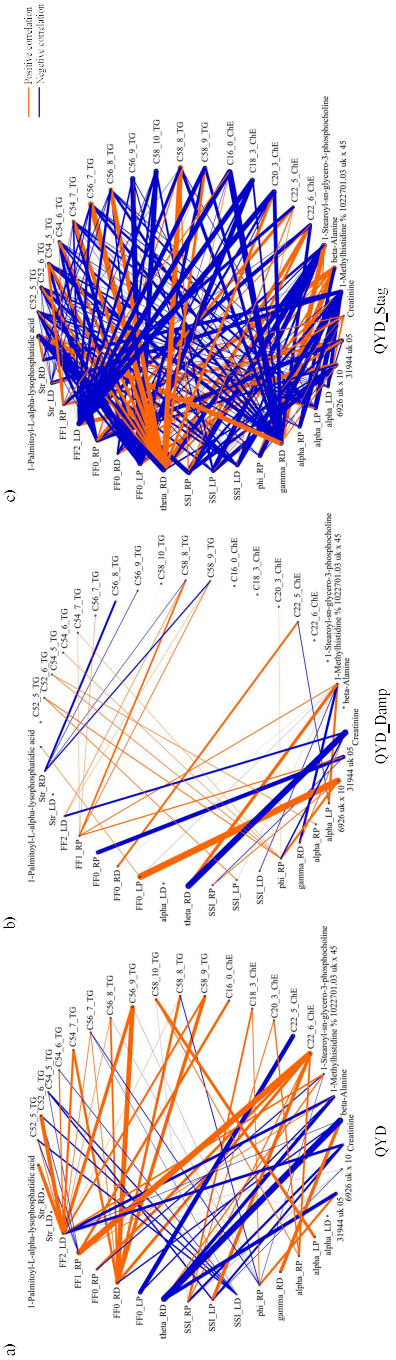


Fig. 3: Correlation networks between metabolic variables and UPE parameters for the indicated pre-T2DM subtypes. Only the correlations between metabolites and UPE parameters with $|r|>0.3$ are plotted。

4. Conclusions and perspectives

Here, we report that plasma metabolomics can be used to stratify the three TCM-based subtypes of early-stage type 2 diabetes, providing better stratification than urine metabolomics. Specifically, increased levels of plasma lipids such as TGs and ChEs may indicate a relatively higher risk of developing cardiovascular disease among patients with specific subtypes. In addition, we used UPE as a non-invasive method for subtyping pre-T2DM, and the UPE parameters were correlated with specific plasma metabolites—primarily lipid metabolites—and these correlations differed among the three subtypes. Thus, combining UPE and plasma metabolomics provides additional insight into the diagnosis of disease and the underlying biochemistry of UPE from a systems biology perspective.

The ability to identify the pre-T2DM syndrome subtype based on TCM is essential for achieving a personalized treatment plan, thereby significantly improving patient care. These results provide a window of opportunity for combining metabolomics with UPE in order to achieve personalized medicine and improve the early diagnosis of disease. Nevertheless, metabolomics platforms do not necessarily cover the entire metabolome, and choices must be made based on the metabolomics platforms that are currently available. Given the difficulties associated with obtaining comprehensive information regarding the dynamic changes reflected by measuring metabolomics, linking metabolomics to UPE under the guidance of TCM-based diagnostics is particularly attractive, promoting the early diagnosis of T2DM. Additional research is needed in order to expand the correlation networks between metabolites and UPE parameters. In addition, current approaches for stratifying T2DM are based on various criteria, which must be consistent for further clinical diagnosis. Therefore, additional research is needed in order to understand TCM-based concepts such as disease syndromes and subtypes.

5. Acknowledgement

Min He is supported by the Chinese Scholarship Council (CSC) during her PhD study at Leiden University in Netherlands (Scholarship File no. 20108220166). Therefore the authors would like to give thanks for the support program from CSC. The authors thank Herman van Wietmarschen for providing support information of this study.

6. References

[1] M. Virally, J.-F. Blicklé, J. Girard, S. Halimi, D. Simon, and P.-J. Guillausseau, "Type 2 diabetes mellitus: epidemiology, pathophysiology, unmet needs and therapeutical perspectives," *Diabetes Metab.*, vol. 33, no. 4, pp. 231–244, 2007.

[2] E. Wright, J. L. Scism-Bacon, and L. C. Glass, "Oxidative stress in type 2 diabetes: the role of fasting and postprandial glycaemia.," *Int. J. Clin. Pract.*, vol. 60, no. 3, pp. 308–314, 2006.

[3] S. I. Ahmad, Ed., *Diabetes: An Old Disease, a New Insight*. Nottingham: Springer Science & Business Media, 2013.

[4] C. D. Munhoz, J. O. Martins, G. A. Cerchiaro, C. Scavone, R. Curi, P. Sannomiya, D. Imunologia, and F. De Medicina, "Neutrophil function and metabolism in individuals with diabetes mellitus," vol. 40, pp. 1037–1044, 2007.

[5] A. Gonzalez-franquesa, A. M. Burkart, and E. Isganaitis, "What Have Metabolomics Approaches Taught Us About Type 2 Diabetes ?," 2016.

[6] A. G. Tabak, M. Jokela, T. Akbaraly, E. J. Brunner, M. Kivimäki, and D. R. Witte, "Trajectories of glycemia, insulin sensitivity and insulin secretion preceding the diagnosis of type 2 diabetes: the Whitehall II study," *Lancet*, vol. 373, no. 9682, pp. 2215–2221, 2009.

[7] J. E. Shaw, R. A. Sicree, and P. Z. Zimmet, "Global estimates of the prevalence of diabetes for 2010 and 2030," *Diabetes Res. Clin. Pract.*, vol. 87, no. 1, pp. 4–14, 2010.

[8] L. Chen, D. J. Magliano, and P. Z. Zimmet, "The worldwide epidemiology of type 2 diabetes mellitus—present and future perspectives," *Nat. Rev. Endocrinol.*, vol. 8, no. 4, pp. 228–236, Nov. 2011.

[9] Centers for Disease Control and Prevention, "National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States," 2014.

[10] Wenya Yang, T. M. Dall, P. Halder, P. Gallo, S. L. Kowal, and P. F. Hogan, "Economic Costs of Diabetes in the U.S. in 2012," *Diabetes Care*, vol. 36, no. 4, pp. 1033–1046, Apr. 2013.

[11] M. I. Mccarthy, "Genomics, type 2 diabetes, and obesity.," *N Engl J Med.*, vol. 363, no. 24, pp. 2339–2350, 2010.

[12] M. I. Harris and R. C. Eastman, "Early detection of undiagnosed diabetes mellitus : a US perspective," *Diabetes Metab Res Rev.*, vol. 7560, no. July, pp. 230–236, 2000.

[13] M. Jiang, C. Lu, C. Zhang, J. Yang, Y. Tan, A. Lu, and K. Chan, "Syndrome differentiation in modern research of traditional Chinese medicine," *J. Ethnopharmacol.*, vol. 140, no. 3, pp. 634–642, Apr. 2012.

[14] J. Guo, H. Chen, J. Song, J. Wang, L. Zhao, and X. Tong, "Syndrome Differentiation of Diabetes by the Traditional Chinese Medicine according to Evidence-Based Medicine and Expert Consensus Opinion.," *Evid Based Complement Altern. Med.*, vol. 2014, p. 492193, 2014.

[15] A. H. Zhang, H. Sun, S. Qiu, and X. J. Wang, "Recent highlights of metabolomics in chinese medicine syndrome research," *Evidence-based Complement. Altern. Med.*, vol. 2013, 2013.

[16] R. Ramautar, R. Berger, J. van der Greef, and T. Hankemeier, "Human metabolomics: Strategies to understand biology," *Curr. Opin. Chem. Biol.*, vol. 17, no. 5, pp. 841–846, 2013.

[17] H. Wei, W. Pasman, C. Rubingh, S. Wopereis, M. Tienstra, J. Schroen, M. Wang, E. Verheij,

and J. van der Greef, “Urine metabolomics combined with the personalized diagnosis guided by Chinese medicine reveals subtypes of pre-diabetes,” *Mol. Biosyst.*, vol. 8, no. 5, pp. 1482–91, Apr. 2012.

[18] T. Wu, M. Yang, H. F. Wei, S. H. He, S. C. Wang, and G. Ji, “Application of metabolomics in traditional chinese medicine differentiation of deficiency and excess syndromes in patients with diabetes mellitus,” *Evidence-based Complement. Altern. Med.*, vol. 2012, 2012.

[19] M. He, M. Sun, E. van Wijk, H. van Wietmarschen, R. van Wijk, Z. Wang, M. Wang, T. Hankemeier, and J. van der Greef, “A Chinese literature overview on ultra-weak photon emission as promising technology for studying system-based diagnostics,” *Complement. Ther. Med.*, vol. 25, pp. 20–26, 2016.

[20] R. van Wijk, J. van der Greef, and E. van Wijk, “Human Ultraweak Photon Emission and the Yin Yang Concept of Chinese Medicine,” *J. Acupunct. Meridian Stud.*, vol. 3, no. 4, pp. 221–231, 2010.

[21] J. A. Ives, E. van Wijk, N. Bat, C. Crawford, A. Walter, W. B. Jonas, R. van Wijk, and J. van der Greef, “Ultraweak Photon Emission as a Non-Invasive Health Assessment: A Systematic Review,” *PLoS One*, vol. 9, no. 2, p. e87401, Feb. 2014.

[22] R. Van Wijk, E. Van Wijk, H. van Wietmarschen, and J. Van der Greef, “Towards whole-body ultra-weak photon counting and imaging with a focus on human beings: A review,” *J. Photochem. Photobiol. B Biol.*, vol. 139, pp. 39–46, Oct. 2014.

[23] R. Van Wijk, E. Van Wijk, F. Wiegant, and J. Ives, “Free radicals and low-level photon emission in human pathogenesis: state of the art.,” *Indian J. Exp. Biol.*, vol. 46, no. 5, pp. 273–309, May 2008.

[24] M. Cifra and P. Pospíšil, “Ultra-weak photon emission from biological samples: Definition, mechanisms, properties, detection and applications,” *J. Photochem. Photobiol. B Biol.*, vol. 139, pp. 2–10, Oct. 2014.

[25] A. Prasad and P. Pospíšil, “Linoleic acid-induced ultra-weak photon emission from Chlamydomonas reinhardtii as a tool for monitoring of lipid peroxidation in the cell membranes.,” *PloS one*, vol. 6, no. 7. p. e22345, Jan-2011.

[26] M. Kobayashi, M. Takeda, T. Sato, Y. Yamazaki, K. Kaneko, K. Ito, H. Kato, and H. Inaba, “In vivo imaging of spontaneous ultraweak photon emission from a rat’s brain correlated with cerebral energy metabolism and oxidative stress.,” *Neurosci. Res.*, vol. 34, no. 2, pp. 103–13, Jul. 1999.

[27] M. Sun, E. Van Wijk, S. Koval, R. Van Wijk, M. He, M. Wang, T. Hankemeier, and J. van der Greef, “Measuring ultra-weak photon emission as a non-invasive diagnostic tool for detecting early-stage type 2 diabetes: a step toward personalized medicine.” J. Photochem. Photobiol. B, vol. 166 , 86–93, Jan. 2017.

[28] M. M. Koek, B. Muilwijk, M. J. Van Der Werf, and T. Hankemeier, “Microbial metabolomics with gas chromatography/mass spectrometry,” *Anal. Chem.*, vol. 78, no. 4, pp. 1272–1281, 2006.

[29] H. A. Van Wietmarschen, J. Van Der Greef, Y. Schroën, and M. Wang, “Evaluation of symptom, clinical chemistry and metabolomics profiles during Rehmannia six formula (R6) treatment: An integrated and personalized data analysis approach,” *J. Ethnopharmacol.*, vol. 150, no. 3, pp. 851–859, 2013.

[30] H. H. M. Draisma, T. H. Reijmers, I. Bobeldijk-Pastorova, J. J. Meulman, G. F. Estourgie-Van Burk, M. Bartels, R. Ramaker, J. van der Greef, D. I. Boomsma, and T. Hankemeier,

"Similarities and differences in lipidomics profiles among healthy monozygotic twin pairs," *OMICS*, vol. 12, no. 1, pp. 17–31, 2008.

[31] E. Van Wijk, R. Van Wijk, R. P. Bajpai, and J. van der Greef, "Statistical analysis of the spontaneously emitted photon signals from palm and dorsal sides of both hands in human subjects.," *J. Photochem. Photobiol. B.*, vol. 99, no. 3, pp. 133–43, Jun. 2010.

[32] R. P. Bajpai, E. P. A. Van Wijk, R. Van Wijk, and J. Van Der Greef, "Attributes characterizing spontaneous ultra-weak photon signals of human subjects," *J. Photochem. Photobiol. B Biol.*, vol. 129, pp. 6–16, 2013.

[33] J. Xia, I. V. Sinelnikov, B. Han, and D. S. Wishart, "MetaboAnalyst 3.0--making metabolomics more meaningful," *Nucleic Acids Res.*, pp. 1–7, 2015.

[34] B. Worley and R. Powers, "Multivariate Analysis in Metabolomics," *Curr. Metabolomics*, vol. 1, no. 1, pp. 92–107, 2013.

[35] R. A. van den Berg, H. C. J. Hoefsloot, J. A. Westerhuis, A. K. Smilde, and M. J. van der Werf, "Centering, scaling, and transformations: improving the biological information content of metabolomics data.," *BMC Genomics*, vol. 7, no. 1, p. 142, Jan. 2006.

[36] K. Benkali, P. Marquet, J. Rérolle, Y. Le Meur, and L. Gastinel, "A new strategy for faster urinary biomarkers identification by Nano-LC-MALDI-TOF/TOF mass spectrometry.," *BMC Genomics*, vol. 9, p. 541, 2008.

[37] I. L. Petersen, G. Tomasi, H. S??rensen, E. S. Boll, H. C. B. Hansen, and J. H. Christensen, "The use of environmental metabolomics to determine glyphosate level of exposure in rapeseed (Brassica napus L.) seedlings," *Environ. Pollut.*, vol. 159, no. 10, pp. 3071–3077, 2011.

[38] R. Van Wijk, E. P. a Van Wijk, and R. P. Bajpai, "Photocount distribution of photons emitted from three sites of a human body.," *J. Photochem. Photobiol. B.*, vol. 84, no. 1, pp. 46–55, Jul. 2006.

[39] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, "Cytoscape: a software environment for integrated models of biomolecular interaction networks.," *Genome Res.*, vol. 13, no. 11, pp. 2498–504, Nov. 2003.

[40] J. Gao, V. G. Tarcea, A. Karnovsky, B. R. Mirel, T. E. Weymouth, C. W. Beecher, J. D. Cavalcoli, B. D. Athey, G. S. Omenn, C. F. Burant, and H. V. Jagadish, "Metscape: A Cytoscape plug-in for visualizing and interpreting metabolomic data in the context of human metabolic networks," *Bioinformatics*, vol. 26, no. 7, pp. 971–973, 2010.

[41] P. Yin, A. Peter, H. Franken, X. Zhao, S. S. Neukamm, L. Rosenbaum, M. Lucio, A. Zell, H. U. H??ring, G. Xu, and R. Lehmann, "Preanalytical aspects and sample quality assessment in metabolomics studies of human blood," *Clin. Chem.*, vol. 59, no. 5, pp. 833–845, 2013.

[42] M. Rantalainen, O. Cloarec, J. K. Nicholson, E. Holmes, and J. Trygg, "OPLS discriminant analysis : combining the strengths of PLS-DA and SIMCA classification y," no. February, pp. 341–351, 2007.

[43] S. Wiklund, E. Johansson, L. Sjöström, E. J. Mellerowicz, U. Edlund, J. P. Shockcor, J. Gottfries, T. Moritz, and J. Trygg, "Visualization of GC/TOF-MS-Based Metabolomics Data for Identification of Biochemically Interesting Compounds Using OPLS Class Models," *Anal. Chem.*, vol. 80, no. 1, pp. 115–122, Jan. 2008.

[44] N. Esser, S. Legrand-Poels, J. Piette, A. J. Scheen, and N. Paquot, "Inflammation as a link between obesity, metabolic syndrome and type 2 diabetes," *Diabetes Res. Clin. Pract.*, vol. 105, no. 2, pp. 141–150, 2014.

[45] T. J. Wang, M. G. Larson, R. S. Vasan, S. Cheng, E. P. Rhee, E. McCabe, G. D. Lewis, C. S. Fox, P. F. Jacques, C. Fernandez, C. J. O'Donnell, S. a Carr, V. K. Mootha, J. C. Florez, A. Souza, O. Melander, C. B. Clish, and R. E. Gerszten, "Metabolite profiles and the risk of developing diabetes.," *Nat. Med.*, vol. 17, no. 4, pp. 448–53, Apr. 2011.

[46] R. M. Salek, M. L. Maguire, E. Bentley, D. V Rubtsov, T. Hough, M. Cheeseman, D. Nunez, B. C. Sweatman, J. N. Haselden, R. D. Cox, S. C. Connor, and J. L. Griffin, "A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human.," *Physiol. Genomics*, vol. 29, no. 2, pp. 99–108, 2007.

[47] R. Wang-Sattler, Z. Yu, C. Herder, A. C. Messias, A. Floegel, Y. He, K. Heim, M. Campillos, C. Holzapfel, B. Thorand, H. Grallert, T. Xu, E. Bader, C. Huth, K. Mittelstrass, A. Döring, C. Meisinger, C. Gieger, C. Prehn, W. Roemisch-Margl, M. Carstensen, L. Xie, H. Yamanaka-Okumura, G. Xing, U. Ceglarek, J. Thiery, G. Giani, H. Lickert, X. Lin, Y. Li, H. Boeing, H.-G. Joost, M. H. de Angelis, W. Rathmann, K. Suhre, H. Prokisch, A. Peters, T. Meitinger, M. Roden, H.-E. Wichmann, T. Pischon, J. Adamski, and T. Illig, "Novel biomarkers for pre-diabetes identified by metabolomics.," *Mol. Syst. Biol.*, vol. 8, no. 615, p. 615, 2012.

[48] S. P. Bagdade JD, Ritter MC, "Accelerated cholesteryl ester transfer in patients with insulin-dependent diabetes mellitus.," *Eur J Clin Invest.*, vol. 21, no. 2, pp. 161–7, 1991.

[49] H. N. Ginsberg, Y.-L. Zhang, and A. Hernandez-Ono, "Regulation of plasma triglycerides in insulin resistance and diabetes.," *Arch. Med. Res.*, vol. 36, no. 3, pp. 232–40, 2005.

[50] J. Sjolin, G. Hjort, G. Friman, and L. Hambraeus, "Urinary excretion of 1-methylhistidine: A qualitative indicator of exogenous 3-methylhistidine and intake of meats from various sources," *Metabolism*, vol. 36, no. 12, pp. 1175–1184, 1987.

[51] H. Chuang, E. Lee, Y. Liu, D. Lee, and T. Ideker, "Network-based classification of breast cancer metastasis," no. 140, pp. 1–10, 2007.

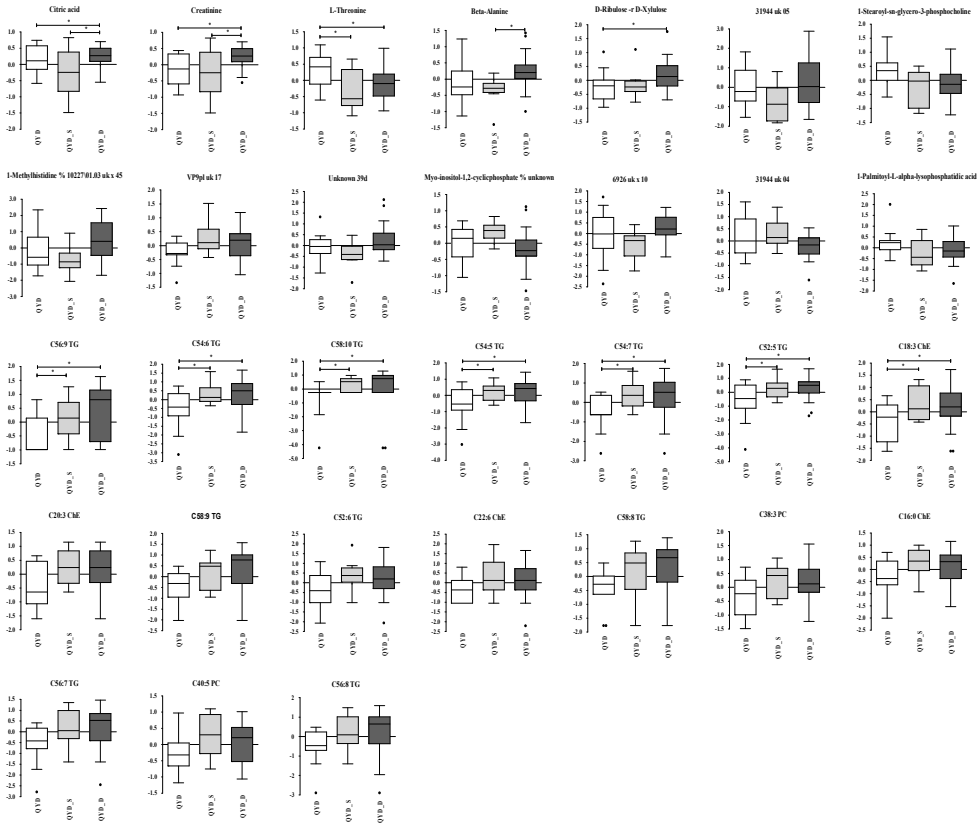
[52] A.-L. Barabasi, N. Gulbahce, and J. Loscalzo, "Network medicine: a network-based approach to human disease.," *Nat. Rev. Genet.*, vol. 12, no. 1, pp. 56–68, 2011.

[53] S. E. Calvano, W. Xiao, D. R. Richards, R. M. Felciano, H. V Baker, R. J. Cho, R. O. Chen, B. H. Brownstein, J. P. Cobb, S. K. Tschoeke, C. Miller-Graziano, L. L. Moldawer, M. N. Mindrinos, R. W. Davis, R. G. Tompkins, and S. F. Lowry, "A network-based analysis of systemic inflammation in humans.," *Nature*, vol. 437, no. 7061, pp. 1032–7, Oct. 2005.

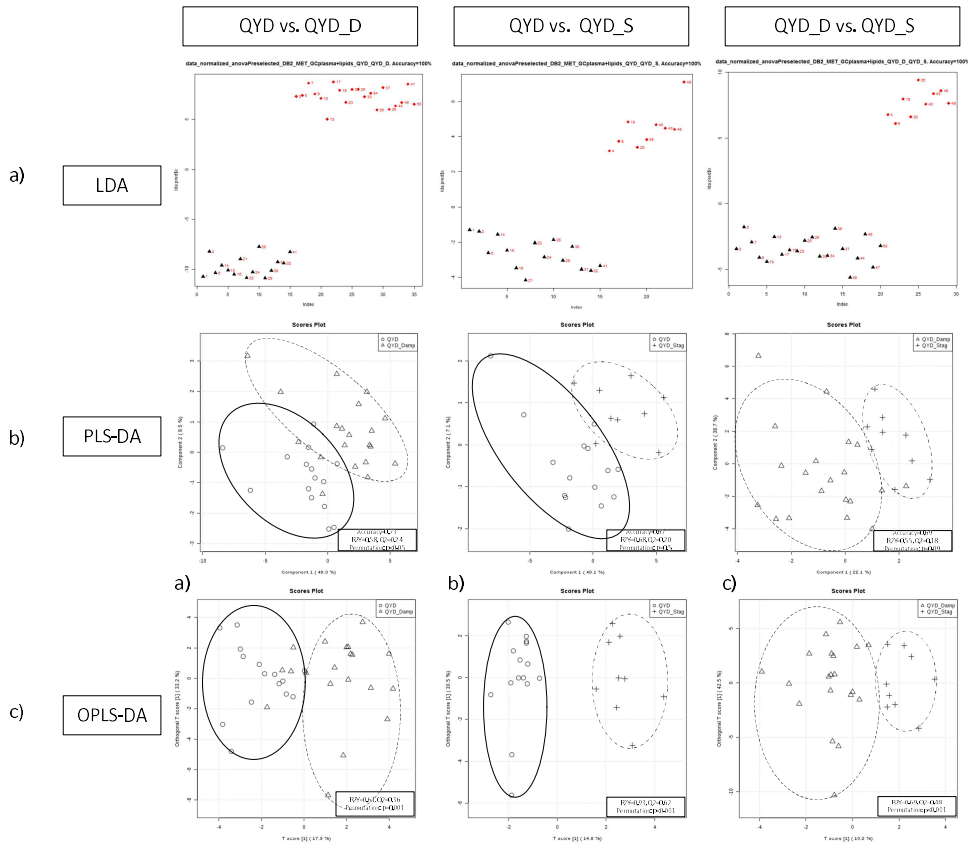
Supplementary information

S-table 1. Preliminary variables in the MS data identified by ANOVA (p<0.1)

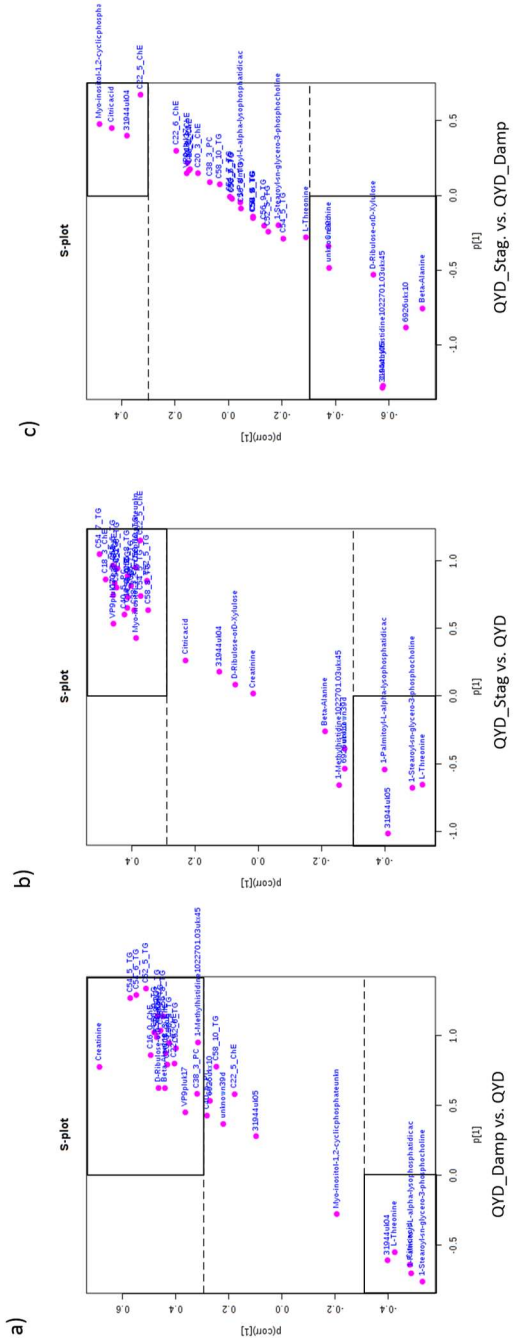
Compounds	*p*-value	Post-hoc test(Fisher's LSD) between groups
Metabolites in GC-MS		
Citric acid	0.018	QYD - QYD_Damp; QYD_Damp - QYD_Stag
Creatinine	0.019	QYD – QYD_Damp; QYD_Damp - QYD_Stag
L-Threonine	0.028	QYD - QYD_Damp; QYD - QYD_Stag
Beta-Alanine	0.039	QYD_Damp - QYD_Stag
D-Ribulose or D-Xylulose	0.049	QYD – QYD_Damp
31944 uk 05	0.052	QYD - QYD_Stag; QYD_Damp - QYD_Stag
1-Stearoyl-sn-glycero-3-phosphocholine	0.054	QYD - QYD_Damp; QYD - QYD_Stag
1-Methylhistidine % 10227\01.03 uk x 45	0.054	QYD_Damp - QYD_Stag
VP9pl uk17	0.054	QYD – QYD_Damp; QYD – QYD_Stag
unknown 39d	0.06	QYD_Damp - QYD_Stag
Myo-inositol-1,2-cyclicphosphate % unknown	0.065	QYD_Damp- QYD_Stag
6926 uk x 10	0.068	QYD_Damp - QYD_Stag
31944 uk 04	0.073	QYD_Damp- QYD_Stag
1-Palmitoyl-L-alpha-lysophosphatidic acid	0.091	QYD - QYD_Damp; QYD - QYD_Stag
Metabolites in LC-MS		
C56_9_TG	0.005	QYD - QYD_Damp; QYD - QYD_Stag
C54_6_TG	0.024	QYD - QYD_Damp; QYD - QYD_Stag
C58_10_TG	0.027	QYD - QYD_Damp; QYD - QYD_Stag
C54_5_TG	0.031	QYD - QYD_Damp; QYD - QYD_Stag
C54_7_TG	0.035	QYD - QYD_Damp; QYD - QYD_Stag
C52_5_TG	0.044	QYD - QYD_Damp; QYD - QYD_Stag
C18_3_ChE	0.050	QYD - QYD_Damp; QYD - QYD_Stag
C20_3_ChE	0.056	QYD - QYD_Damp; QYD - QYD_Stag
C58_9_TG	0.056	QYD - QYD_Damp
C52_6_TG	0.061	QYD - QYD_Damp; QYD - QYD_Stag
C22_6_ChE	0.070	QYD - QYD_Damp; QYD - QYD_Stag
C58_8_TG	0.081	QYD - QYD_Damp
C38_3_PC	0.085	QYD - QYD_Damp; QYD - QYD_Stag
C16_0_ChE	0.085	QYD - QYD_Damp; QYD - QYD_Stag
C56_7_TG	0.086	QYD - QYD_Damp; QYD - QYD_Stag
C40_5_PC	0.091	QYD - QYD_Stag
C56_8_TG	0.092	QYD - QYD_Damp



S-fig. 1: Box plots summarizing the 32 preliminary variables (plasma metabolites detected by LC-MS and GC-MS, identified by ANOVA (p<0.1)) in pre-T2DM subjects. Individual metabolite (peak area ratio between target metabolites and relevant internal standard) for the three groups are illustrated using boxplots after logarithmic transformation and pareto-scaling for data normal distribution. The metabolites which differed significantly based on ANOVA (p<0.05) were then followed by a post-hoc analysis (Fisher's least significant difference method) to show between which two groups the differences are significant (*).



S-fig. 2: Performance comparison between three supervised multivariate analysis models (LDA, PLSDA, and OPLS-DA), based on metabolite profiling in plasma of pre-T2DM subjects detected and integrated by LC-MS and GC-MS. A Permutation test with 1000 iterations (p<0.05) as well as the R2 and Q2 showed that the OPLS-DA model performed best.



S-fig. 3: Loading plots from the OPLS-DA analysis based on metabolite profiling in plasma of pre-T2DM subjects detected and integrated by LC-MS and GC-MS. Magnitude score (covariance) is combined with the reliability (correlation) results, for indicating variable influence that contributed to separation between each pair of subtypes. a) The QYD_Damp subtype versus the QYD subtype; b) the QYD_Stag subtype versus the QYD subtype; c) the QYD_Stag subtype versus the QYD_Damp subtype.