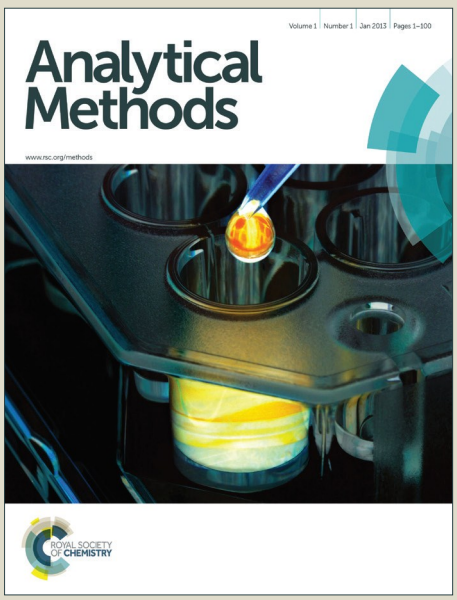
Analytical Methods

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

A Combination of GC-MS and Chemometrics Reveals Metabolic Differences between Serum and Plasma

Shuting Shi,[b] Lunzhao Yi,[a*] Yonghuan Yun,[b**] Xiaojuan Zhang,[b] Yizeng Liang[b]

[a] *Yunnan Food Safety Research Institute, Kunming University of Science and Technology, Kunming, Yunnan, 650500, P.R. China.*

[b] *College of Chemistry and Chemical Engineering, Central South University, Changsha, Hunan, 410083, P.R. China.*

Abstract: Blood serum and plasma are the most frequently used biofluids in metabolomics. The primary aim of this study was to ascertain the feasibility of the commonly used gas chromatography-mass spectrometer (GC-MS)-based methoximation followed by silylation with BSTFA+1%TMCS (MeOx-TMS) derivatization method, and to reveal the metabolic differences between serum and plasma. Individual variations were evaluated by different groups of sera and plasma samples collected from healthy volunteers in 2011 and 2013. The experimental results indicated that differences of metabolic levels among individuals were much higher than the variations of the experimental repeatability and precision. In addition, discriminant model between serum and plasma was established using partial least squares-discriminant analysis (PLS-DA). Six characteristic metabolites, phosphate, serine, 2,3,4-trihydroxybutyrate, citric acid, glucose and arachidonic acid, were screened out and considered to have the most important contribution to the discrimination. Results of this work will provide some valuable suggestions to researchers on the selection of suitable biofluids in metabolomic research.

Key words: metabolomics; MeOx-TMS derivatization method; GC-MS; PLS-DA; serum; plasma

1. Introduction

Metabolomics is defined as the quantitative measurement of the dynamic

*Corresponding author: Tel.: +86-871-65920302; E-mail address: ylz7910@hotmail.com.

**Corresponding author: Tel.: +86-731-88830831; E-mail address: yunyonghuan@foxmail.com.

multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification.[1] Compared to genomics and proteomics, in which the analytes of interest are biological macromolecules and the chemistries of the building blocks are relatively well defined, metabolomics is a rapidly growing discipline concerned with analyzing the entire measurable array of low-molecular-weight compounds, metabolites, in a given biological system. Typically, the endogenous metabolites of humans that are profiled include organic acids, amino acids, fatty acids, sugars, cholesterol and other substances that are intermediates in cellular metabolism.[2-4] The analytical technique in metabolomics is chosen according to the specific study being performed. GC-MS has been proven to be a potentially useful metabolic profiling platform due to its high sensitivity, peak resolution and reproducibility.[5] However, GC-MS makes it hard to directly get the information of metabolites which are difficult to volatilize. So, it is necessary to do chemical derivatization to reduce the polarity of the polar functional group, increase the thermal stability and volatility of the analytes. MeOx-TMS derivatization method is such a method which has been widely used in GC-MS based metabolomics studies.[6-8]

Chemometric analysis has, because of its ability to provide interpretable models for complex inter-correlated data, become an integrated part of the global metabolite analysis technique.[9] Partial least squares-discriminant analysis (PLS-DA) is a partial least squares regression of a set Y of binary variables describing the categories of a categorical variable on a set X of predictor variables. Currently, PLS-DA has been one of multivariate data analysis method widely employed to visualize metabolic disorder.[10-12]

In metabolomic research, blood serum and plasma are the most frequently studied samples for several reasons. Firstly, sample collection is minimally invasive compared with the collection of cerebrospinal fluid and tissues. Besides, as an integrative biofluid that incorporate the functions and phenotypes of many different parts of body in a single sample, blood is a ‘metabolic footprint’ of tissue metabolism.[13] The essential difference between serum and plasma is that whereas serum is collected after a process of clotting, plasma is collected without clotting.

However, as reviewed by Mannello,[14] using the wrong matrix (e.g. plasma in place of serum) can lead to improper diagnosis. Therefore, a deeper knowledge of differences between plasma and serum can contribute to effectively choosing samples. Judith R. Denery[15] has examined the proteomic difference between plasma and serum. In the newly emerging field of metabolomics,[1] there have been a few studies related to comparing differences between serum and plasma metabolite profiles.[16-18] For example, by analyzing 377 fasting individuals, better reproducibility in plasma and higher metabolite concentrations in serum were concluded by Zhonghao Yu et al.[18]

In this study, 151 samples were collected from 151 different healthy volunteers in 2011 (40 sera and 34 plasma samples) and 2013 (40 sera and 37 plasma samples). The two batches of samples were analyzed separately. It will help us to evaluate the influence of individual variation and to reveal the inherent differences in metabolic levels between serum and plasma. Detection of metabolites was performed using MeOx-TMS derivatization method followed by GC-MS. In addition, PLS-DA models were constructed to establish the significance of differences between these two fluids. We believe that results of this work will provided some valuable suggestions on the selection of suitable biofluids in metabolomics research.

2. Experimental

2.1. Chemicals and reagents

BSTFA+1%TMCS (N,O-Bis(trimethylsilyl)trifluoroacetamide with 1% trimethyl-chorosilane, for GC) (>99.0% purity), pyridine (>99.8% purity), methoxyamine hydrochloride (>98% purity), heptadecanoic acid (C17:0,>99.0% purity), sodium L-lactate (>99.0% purity), alanine (>99.5% purity), glycine (>99.5% purity), β-hydroxybutyrate (>99.5% purity), L-isoleucine (>99.0% purity), L-proline (>99.0% purity), L-serine (>99.0% purity), L-threonine (>99.0% purity), pyroglutamic acid (>99.0% purity), glutamic acid (>99.0% purity), L-phenylalanine (>99.0% purity), citric acid (>99.5% purity), 1,5-anhydro-D-sorbitol (>99.0% purity), fructose (>99.8% purity), galactose (>99.0% purity), glucose (>99.8% purity), hexadecanoic acid (C16:0, >99.0% purity), oleic acid (C18:1, >99.0% purity), linoleic acid

(C18:2ω-6, >99.0% purity), stearic acid (C18:0, >99.0% purity), arachidonic acid (C20:4ω-6, >99.0% purity) and cholesterol (>99.0% purity) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Methanol is analytical grade and purchased from Changsha Fufan Trade Ltd. (Changsha, China). The methoxyamine hydrochloride was dissolved in the pyridine at a concentration of 20 mg/mL. Heptadecanoic acid (Internal standard, IS) was dissolved in methanol at a concentration of 1 mg/mL.

2.2. Sample collection

In order to avoid the interferences from post-prandial phase, 80 fasting plasma samples and 71 sera samples were collected from 151 healthy volunteers. All volunteers were fasting at least eight hours. 40 sera samples (age range was 42-73, 21 males and 19 females) and 34 plasma samples (age range was 40-70, 17 males and 17 females) were obtained in June 2011and the other 40 sera samples (age range was 45-70, 20 males and 20 females) and 37 plasma samples (age range was 45-69, 19 males and 18 females) were acquired in September 2013 following informed consent from Xiangya Hospital of Central South University, Hunan, China, according to institutional regulations. All volunteers were tested in the Physical Examination Center of Xiangya Hospital. The results of the physical examination guaranteed that they did not have any diseases. Aliquots of plasma and serum samples were stored at -80℃ until required for the experiment.

2.3. Sample preparation

Each 100 μL serum or plasma sample was mixed with 350μL methanol to precipitate the proteins, and 50 μL of heptadecanoic acid solution (1 mg/mL in methanol) was added as the internal standard. Then, the mixture was vigorously vortexed for 1 min and centrifuged for 10 min at 16000 rpm (17800×g) at 4℃. The supernatant (400μL) was transferred into a 5 mL glass centrifugation tube, and then evaporated to dryness by N_2 gas. Next, 50μL of methoxyamine hydrochloride (20 mg/mL in pyridine) was added to the dry tube, and the resultant mixture was vigorously vortex-mixed for 1 min and heated in a water bath at 70℃ for 1 h with a glass plug. Finally, 100μL of BSTFA with 1% TMCS as catalyst was added to the solution, vortex-mixed vigorously for 1 min, and heated in a water bath at 70℃ for 1 h with a glass plug before GC-MS analysis.

2.4. Instruments and chromatographic conditions

Analyses were performed on a Shimadzu GC2010A (Kyoto, Japan) gas chromatography instrument, coupled with a QP2010 mass spectrometer (Compaq-Pro Linear data system, class 5K software). A sample of 1.0μL was injected into a DB-5 ms capillary column (0.25 mm × 30 m × 0.25 μm, Agilent, USA) at a split ratio of 1:10. The column temperature was initially maintained at 70 ℃ for 4 min, and then increased at a rate of 8 ℃/ min from 70 to 300 ℃ and held for 3 min. The helium carrier gas flow rate was 1.0 mL/min. The injector temperature was 280℃, the septum purge flow rate was 3 mL /min, and the purge was turned on all the time. The interface temperature was 250 ℃ and ion source temperature was 200 ℃. Ionization was achieved by a 70 eV electron beam. Masses were acquired in a full scan mode, over the range from m/z 35 to 800, with a scan speed of 0.2/sec when the 0.9kV of detector voltage was turned on after a solvent delay of 6 min.

2.5. Data analysis

The identification of structures of peaks-of-interest was based on the similarity search of the NIST/EPA/NIH Mass Spectra Library (NIST 05). 32 peaks were considered to be endogenous metabolites, in which 22 metabolites were identified by their chemical standard substances. As for the quantification, peak integration was employed. To normalize the blood data, the variables were expressed as the ratio of peak area of corresponding metabolites to that of the internal standard on the same chromatogram. A data matrix was generated for statistical analysis using PLS-DA and each row and column of the matrix represent a sample and a variable, respectively. The data matrix was auto-scaled, then analyzed by PLS-DA. For the purpose of cross-validation, ten-fold cross-validation was applied. Class membership was predicted using a discriminant line between two classes obtained by linear discriminant analysis (LDA). The statistical analysis was performed using the in-house software written in MATLAB (version 6.5, The Math Works, Natick, MA, USA).

3. Results and discussion

3.1. Validation of the feasibility of MeOx-TMS derivatization method

In this study, we analyzed the repeatability and precision of MeOx-TMS derivatization method to confirm its range of applications. Repeatability was obtained by analyzing 5 parallel plasma samples and precision was checked by 5 continuous injections of the same sample. Both repeatability and precision were expressed as the relative standard deviation (RSD, %), shown in Table 1. A large number of metabolites displayed good repeatability and precision. However, RSDs of 5 amino acids, including alanine, isoleucine, serine, threonine and phenylalanine, were found to be higher than 20% for precision. While for repeatability, 10 metabolites are higher than 20%, including 7 amino acid (alanine, glycine, isoleucine, serine, threonine, glutamic acid and phenylalanine), glycerate, 2,3,4-trihydroxybutyrate and creatinine enol. Experimental results demonstrated that some metabolites, especially amino acid, are unstable under the analytical conditions of the MeOx-TMS derivatization method. If very accurate quantitative analysis is indispensable for a study, some other methods should be employed to analyze these unstable metabolites.

In addition, RSDs of metabolites' concentrations in serum or plasma samples was calculated to reflect the individual differences. It turns out that the RSD value are much higher than the values of repeatability and precision (Table 1), indicating that differences between individuals were far greater than variations from the experimental factors. On this basis, MeOx-TMS derivatization method was suitable for our further metabolomics analysis of serum and plasma.

Insert Table 1

3.2. Metabolic profiling of serum and plasma

The GC-MS total ion chromatograms (TICs) of serum and plasma are presented in Figure 1(a) and (b), respectively. Visual inspection of the TICs revealed that some differences existed between serum and plasma. The variations of concentrations were demonstrated to be similar between samples collected from 2011 and 2013 for the majority of the metabolites (23/32), shown in Table 1.The results of t-test are the same for the 23 metabolites. That is, no significant difference existed in metabolic levels of β-hydroxy butyrate and mannose (t-test $p>0.05$ with a signed t value of "0"), 21 metabolites showed significant differences between serum and plasma (t-test $p< 0.05$

with a signed t value of "1"). 15 of them, including lactate, alanine, sarcosine, glycine, urea, phosphate, isoleucine, glycerate, serine, threonine, hexadecanoic acid, myo-inositol, oleic acid, stearic acid and arachidonic acid, presented a higher concentration in serum. On the contrary, the other 6 variables, including oxalic acid, proline, 2,3,4-trihydroxybutyrate, citric acid, 1,5-anhydro-sorbitol and fructose, displayed a higher concentration in plasma. These results demonstrated that metabolic phenotype of serum was markedly different from plasma and once again proved that higher levels were existed in blood serum for most metabolites. As for another 9 metabolites, the t-test results were opposite between 2011 and 2013. It may mainly due to the individual variations of samples collected from various sources. It seems that improvement of samples' representativeness is very important in metabolomics research.

Insert Figure 1

The higher concentrations of metabolites may lead to a higher sensitivity in biomarker detection. [18] Table 2 shows an overview of diseases linked to each metabolite which showed a significant difference between serum and plasma. Refer to Table 2, most metabolites were reported to be associated with various diseases. Based on the results obtained in our study, appropriate fluid was suggested for further metabolomic investigation of these diseases. In addition, one disease may be linked with a variety of metabolites. For an instance, Alzheimer's disease was found to be associated with glycine, isoleucine, serine, threonine, myo-inositol, proline and 1,5-anhydro-sorbitol. As shown in table 1, while glycine, isoleucine, serine, threonine and myo-inositol showed a higher concentration in serum, proline and 1,5-anhydro-sorbitol displayed an opposite result. It reminded us that fluid for the analysis should be chosen according to a certain research target.

Insert Table 2

3.3. Metabolomics-based discrimination model

It is relatively challenging to visualize the metabolic profiles between serum and plasma just based on GC-MS data because of the inter-subject variations in the data matrix composition, the complexity of TICs and differences of metabolites'

concentrations. Therefore, PLS-DA was applied to construct a visible model to discriminate serum from plasma.

In this study, the peak areas of 32 metabolites were used as the input data of PLS-DA to establish a visible model for discrimination of serum and plasma. The final optimized 2-dimensional (2D) PLS-DA models of sera and plasma collected from 2011 and 2013 were presented in Figure 2a and c, respectively. Sera and plasma were apart and distinctly grouped with PLS-DA. For the samples collected from 2011, the correct rates of ten-fold cross validation for sera, plasma were both 100%. The AUC value was 97.06%. For the samples from 2013, the correct rate was 100% and the AUC value was 97.30%.

Insert Figure 2

Following PLS-DA model construction, the absolute values of the coefficients (β) of the 32 metabolites were applied to identify which metabolites conspicuously contributed to the metabolomic differences between sera and plasma. The higher the absolute value of β is, the more the influence of corresponding variables is. As shown in Figure 2b, 6 metabolites were classified with higher coefficients in the data matrix obtained in 2011, including phosphate, isoluecine, proline, 2,3,4-trihydroxybutyrate, citric acid and stearic acid. While for the data matrix aquired in 2013 (Figure 2d), the 6 variables were alanine, 2,3,4-trihydroxybutyrate, citric acid, glucose, myo-inositol and stearic acid.

To guarantee the validity of conclusions, the two batches of data were pooled together for PLS-DA analysis. Serum and plasma were separated clearly by the discriminant line with the correct rates of 100% (shown in Figure 3a). The AUC value was 98.59%. AUC value increased by more than 1% in contrast to the models of 2011 and 2013, indicating that the integrated model is more reliable to discriminate serum from plasma. While as shown in Figure 3b, the 6 metabolites, which were classified with higher coefficients in the whole data matrix, were phosphate, serine, 2,3,4-trihydroxybutyrate, citric acid, glucose and arachidonic acid. Five of them are in accordance with the results of 2011 or 2013 except for arachidonic acid. It indicated that different batches of samples may lead to various results. However, most screened

metabolites were similar. More attention should be paid to these common metabolites in choosing the suitable fluids for metabolic research of diseases. For example, serum was recommended for diseases such as Alzheimer's disease, hypophosphatemia which are associated with serine or phosphate. Similarly, plasma was proposed for diseases such as childhood obesity and type 2 diabetes mellitus which are linked with 2,3,4-trihydroxybutyrate and citric acid, respectively.

Insert Figure 3

4. Conclusions

A protocol of MeOx-TMS derivatization method followed by GC-MS is very popular in metabolomics. Although the evaluation of feasibility of this method is very important for its correct applications, few publications have reported the repeatability and precision of the entire detected metabolites. In this study, experimental and individual variations were evaluated and compared. The results indicated that differences between individuals were far greater than variations from any of the experimental factors. For most cases, MeOx-TMS derivatization method was suitable for the metabolomic research. On this basis, metabolic levels of serum and plasma were analyzed and compared comprehensively. A discriminant model between serum and plasma was established by PLS-DA, phosphate, serine, 2,3,4-trihydroxybutyrate, citric acid and glucose were screened out as the key metabolites. This study will be very helpful for the right application of MeOx-TMS derivatization method and will give some reasonable suggestions on the biofluids selection in metabolomics.

Acknowledgements

This work was supported financially by National Nature Foundation Committee of P.R. China (No.21105129, No.21465016), the Fundermental Research Funds for the Central Universities (2011QNZT053).

References

1 J. K. Nicholson, J. C. Lindon and E. Holmes, *Xenobiotica; the fate of foreign compounds in biological systems*, 1999, **29**, 1181-1189.

2 S. J. Mihalik, S. F. Michaliszyn, J. de las Heras, F. Bacha, S. Lee, D. H. Chace, V. R. DeJesus, J. Vockley and S. A. Arslanian, *Diabetes care*, 2012, **35**, 605-611.

3 M. Trupp, H. Zhu, W. R. Wikoff, R. A. Baillie, Z.-B. Zeng, P. D. Karp, O. Fiehn, R. M. Krauss and R. Kaddurah-Daouk, *PLoS ONE*, 2012, **7**, e38386.

4 L. Z. Yi, C. C. Song, Z. Y. Hu, L. F. Yang, L. B. Xiao, B. Yi, W. Z. Jiang, Y. Cao and L. Q. Sun, *Metabolomics*, 2013, DOI: 10.1007/s11306-013-0606-x, 1-12.

5 K. K. Pasikanti, P. Ho and E. Chan, *J. Chromatogr. B*, 2008, **871**, 202-211.

6 L. Z. Yi, S. T. Shi, Z. B. Yi, R. H. He, H. M. Lu and Y. Z. Liang, *Anal. Methods*, 2014, **6**, 4380-4387.

7 J.-H. Huang, R.-H. He, L.-Z. Yi, H.-L. Xie, D.-s. Cao and Y.-Z. Liang, *Talanta*, 2013, **110**, 1-7.

8 W. Zhang, L. X. Zhang, H. D. Li, Y. Z. Liang, R. Hu, N. N. Liang, W. Fan, D. S. Cao, L. Z. Yi and J. D. Xia, *Chromatographia*, 2012, **75**, 799-808.

9 R. Madsen, T. Lundstedt and J. Trygg, *Anal. Chim. Acta*, 2010, **659**, 23-33.

10 E. Szymańska, E. Saccenti, A. K. Smilde and J. A. Westerhuis, *Metabolomics*, 2012, **8**, 3-16.

11 D. MacIntyre, B. Jimenez, E. J. Lewintre, C. R. Martín, H. Schäfer, C. G. Ballesteros, J. R. Mayans, M. Spraul, J. García-Conde and A. Pineda-Lucena, *Leukemia*, 2010, **24**, 788-797.

12 M. R. Pears, J. D. Cooper, H. M. Mitchison, R. J. Mortishire-Smith, D. A. Pearce and J. L. Griffin, *J. Biol. Chem.*, 2005, **280**, 42508-42514.

13 D. B. Kell, M. Brown, H. M. Davey, W. B. Dunn, I. Spasic and S. G. Oliver, *Nat. Rev. Microbiol.*, 2005, **3**, 557-565.

14 F. Mannello, *Arterioscl. Throm. Vas.*, 2008, **28**, 611-614.

15 J. R. Denery, A. A. Nunes and T. J. Dickerson, *Anal. Chem.*, 2010, **83**, 1040-1047.

16 L.-S. Liu, J.-Y. Aa, G.-J. Wang, B. Yan, Y. Zhang, X.-W. Wang, C.-Y. Zhao, B. Cao, J. Shi and M.-J. Li, *Anal. biochem.*, 2010, **406**, 105-112.

17 D. C. Wedge, J. W. Allwood, W. Dunn, A. A. Vaughan, K. Simpson, M. Brown, L. Priest, F. H. Blackhall, A. D. Whetton and C. Dive, *Anal. Chem.*, 2011, **83**, 6689-6697.

18 Z. H. Yu, G. Kastenmüller, Y. He, P. Belcredi, G. Möller, C. Prehn, J. Mendes, S. Wahl, W. Roemisch-Margl and U. Ceglarek, *PLoS One*, 2011, **6**, e21230.

19 G. Blomqvist, M. Alvarsson, V. Grill, G. Von Heijne, M. Ingvar, J. O. Thorell, S. Stone-Elander, L. Widen and K. Ekberg, *Am. J. Physiol. Endocrinol. Metab.*, 2002, **283**, E20-28.

20 S. A. Khan, I. J. Cox, G. Hamilton, H. C. Thomas and S. D. Taylor-Robinson, *Liver Int.*, 2005, **25**, 273-281.

21 N. Redjems-Bennani, C. Jeandel, E. Lefebvre, H. Blain, M. Vidailhet and J. L. Gueant, *Gerontology*, 1998, **44**, 300-304.

22 H. L. Levy, J. T. Coulombe and R. Benjamin, *Pediatrics*, 1984, **74**, 509-513.

23 A. N. Fonteh, R. J. Harrington, A. Tsai, P. Liao and M. G. Harrington, *Amino acids*, 2007, **32**, 213-224.

24 K. Q. Do, C. J. Lauer, W. Schreiber, M. Zollinger, U. Gutteck-Amsler, M. Cuenod and F. Holsboer, *J. Neurochem.*, 1995, **65**, 2652-2662.

25 S. Rainesalo, T. Keranen, J. Palmio, J. Peltola, S. S. Oja and P. Saransaari, *Neurochem. Res.*, 2004, **29**, 319-324.

26 C. T. Peng, K. H. Wu, S. J. Lan, J. J. Tsai, F. J. Tsai and C. H. Tsai, *Eur. J. Cancer*, 2005, **41**, 1158-1163.

27 H. Norrelund, H. Wiggers, M. Halbirk, J. Frystyk, A. Flyvbjerg, H. E. Botker, O. Schmitz, J. O. Jorgensen, J. S. Christiansen and N. Moller, *J. Intern. Med.*, 2006, **260**, 11-21.

28 B. Marescau, P. P. De Deyn, J. Holvoet, I. Possemiers, G. Nagels, V. Saxena and C. Mahler, *Metabolism: clinical and experimental*, 1995, **44**, 584-588.

29 A. Subramanian, A. Gupta, S. Saxena, A. Gupta, R. Kumar, A. Nigam, R. Kumar, S. K. Mandal and R. Roy, *NMR Biomed.*, 2005, **18**, 213-225.

30 J. Amanzadeh and R. F. Reilly, *Nat. Clin. Pract. Nephrol.*, 2006, **2**, 136-148.

31 D. J. Dietzen, T. R. Wilhite, D. N. Kenagy, D. S. Milliner, C. H. Smith and M. Landt, *Clin. Chem.*, 1997, **43**, 1315-1320.

32 J. L. Sepulveda, Y. C. Tanhehco, M. Frey, L. D. Guo, L. J. Cropcho, K. M. Gibson and H. C. Blair, *Arch. Pathol. Lab. Med.*, 2010, **134**, 73-80.

33 H. U. Shetty, H. W. Holloway and M. B. Schapiro, *Clin. Chem.*, 1996, **42**, 298-302.

34 D. S. Wishart, C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, D. D. Hau, N. Psychogios, E. Dong, S. Bouatra, R. Mandal, I. Sinelnikov, J. Xia, L. Jia, J. A. Cruz, E. Lim, C. A. Sobsey, S. Shrivastava, P. Huang, P. Liu, L. Fang, J. Peng, R. Fradette, D. Cheng, D. Tzur, M. Clements, A. Lewis, A. De Souza, A. Zuniga, M. Dawe, Y. Xiong, D. Clive, R. Greiner, A. Nazyrova, R. Shaykhutdinov, L. Li, H. J. Vogel and I. Forsythe, *Nucleic Acids Res.*, 2009, **37**, D603-610.

35 Y. Min, K. Ghebremeskel, C. Lowy, B. Thomas and M. A. Crawford, *Diabetologia*, 2004, **47**, 75-81.

36 S. Wang, A. Ma, S. Song, Q. Quan, X. Zhao and X. Zheng, *Hypertens. Res.*, 2008, **31**, 623-632.

37 Y. Ogawa, N. Machida, M. Jahana, M. Gakiya, Y. Chinen, M. Oda, M. Morozumi and K. Sugaya, *Front. Biosci.*, 2004, **9**, 2901-2908.

38 D. S. Raj, M. Ouwendyk, R. Francoeur and A. Pierratos, *Blood Purif.*, 2000, **18**, 97-102.

39 M. M. Zeng, Y. Z. Liang, H. D. Li, M. Wang, B. Wang, X. Chen, N. Zhou, D. S. Cao and J. Wu, *J. Pharmaceut. Biomed.*, 2010, **52**, 265-272.

40 R. A. Wevers, U. Engelke, U. Wendel, J. G. de Jong, F. J. Gabreels and A. Heerschap, *Clin. Chem.*, 1995, **41**, 744-751.

41 E. Bairaktari, K. Katopodis, K. C. Siamopoulos and O. Tsolas, *Clin. Chem.*, 1998, **44**, 1256-1261.

42 E. Bairaktari, K. Seferiadis, G. Liamis, N. Psihogios, O. Tsolas and M. Elisaf, *Clin. Chem.*, 2002, **48**, 1106-1109.

43 P. Schrauwen and M. K. C. Hesselink, *Diabetologia*, 2008, **51**, 1694-1697.

44 K. M. Dungan, J. B. Buse, J. Largay, M. M. Kelly, E. A. Button, S. Kato and S. Wittlin, *Diabetes care*, 2006, **29**, 1214-1219.

45 T. Kawasaki, H. Akanuma and T. Yamanouchi, *Diabetes care*, 2002, **25**, 353-357.

Table1. Methodological repeatability and precision comparing to individual differences reflected in serum and plasma

No.	Metabolites	Repeatability (RSD(%), n=5)	Precision (RSD(%), n=5)	2011					2013				
				Serum(n=40)		Plasma(n=34)		T	Serum(n=40)		Plasma(n=37)		T
				Concentrations (mg/mL)	RSD (%)	Concentrations (mg/mL)	RSD (%)		Concentrations (mg/mL)	RSD (%)	Concentrations (mg/mL)	RSD (%)	
1	Oxalic acid	12.27	11.06	0.025±0.003	12.66	0.028±0.004	15.20	1	0.025±0.007	26.93	0.027±0.007	25.90	1
2	Lactate*	15.10	9.73	1.615±0.602	37.24	0.994±0.338	34.01	1	1.083±0.327	30.14	0.869±0.290	33.38	1
3	Tartronic acid	12.18	4.31	0.012±0.001	12.35	0.011±0.004	35.71	0	0.007±0.002	22.33	0.013±0.003	22.63	1
4	Alanine*	27.01	22.07	0.153±0.062	40.77	0.113±0.036	32.13	1	0.124±0.034	27.46	0.030±0.013	41.33	1
5	Sarcosine	12.12	11.26	0.078±0.026	33.66	0.050±0.021	41.22	1	0.157±0.064	40.69	0.118±0.047	39.69	1
6	Glycine*	25.72	13.94	0.108±0.034	31.69	0.070±0.031	44.26	1	0.063±0.029	45.91	0.025±0.012	41.14	1
7	β-Hydroxy butyrate*	12.82	13.76	0.036±0.030	84.14	0.03±0.028	93.87	0	0.031±0.033	58.77	0.025±0.007	91.76	0
8	Urea	18.81	13.99	1.481±0.295	19.93	0.960±0.256	26.68	1	1.378±0.410	27.08	0.671±0.300	44.69	1
9	phosphate	12.41	8.12	0.076±0.014	18.52	0.594±0.146	24.50	1	0.483±0.116	15.85	0.354±0.113	32.01	1
10	Isoleucine*	31.57	21.35	0.146±0.047	32.58	0.030±0.007	32.40	1	0.046±0.027	41.72	0.025±0.010	40.79	1
11	Proline*	13.39	8.82	0.038±0.006	17.31	0.070±0.028	39.55	1	0.021±0.007	34.46	0.050±0.034	67.99	1
12	glycerate	25.90	13.65	0.012±0.003	27.54	0.009±0.003	39.59	1	0.011±0.004	39.18	0.009±0.004	51.82	1
13	Serine*	28.68	35.92	0.085±0.016	38.83	0.046±0.012	46.76	1	0.059±0.02	34.03	0.033±0.023	72.49	1
14	Threonine*	27.52	23.47	0.071±0.013	18.36	0.052±0.013	24.62	1	0.056±0.021	36.81	0.046±0.025	57.26	1
15	Pyroglutamate*	16.06	11.49	0.221±0.029	13.24	0.161±0.040	24.96	1	0.185±0.045	24.19	0.160±0.042	16.09	0
16	2,3,4-trihydroxybutyrate	21.29	14.28	0.005±0.002	31.22	0.010±0.004	36.85	1	0.004±0.003	33.07	0.015±0.006	41.28	1
17	Creatinine enol	21.07	13.88	0.014±0.004	25.80	0.009±0.004	48.07	1	0.013±0.005	34.99	0.015±0.007	40.35	0
18	glutamic acid*	21.19	12.10	0.036±0.018	48.66	0.012±0.005	38.95	1	0.014±0.007	52.12	0.016±0.011	38.30	0
19	Phenylalanine*	24.08	20.22	0.031±0.007	32.57	0.021±0.005	33.79	1	0.023±0.016	26.66	0.021±0.006	30.36	1
20	Citric acid*	10.30	7.25	0.032±0.009	28.50	0.058±0.015	25.18	1	0.020±0.009	41.75	0.054±0.010	18.85	1

1,5-Anhydro-sorbitol*	6.21	6.88	0.157±0.046	29.35	0.118±0.043	36.08	1	0.139±0.028	27.81	0.097±0.038	39.09	1
Fructose*	10.15	9.85	0.092±0.024	26.17	0.116±0.031	27.02	1	0.027±0.012	28.49	0.039±0.010	33.57	1
Galactose*	7.80	7.67	0.050±0.009	17.88	0.045±0.013	29.00	1	0.111±0.522	22.25	0.037±0.009	25.90	0
Glucose*	7.64	6.36	3.485±0.390	11.20	3.278±0.720	21.97	0	4.152±0.433	10.42	2.932±0.191	10.59	1
Mannose	9.66	5.80	0.027±0.015	54.61	0.073±0.064	88.20	0	0.038±0.026	69.25	0.094±0.092	98.43	0
Hexadecanoic acid (C16:0)*	7.61	9.94	0.214±0.034	15.77	0.124±0.028	22.34	1	0.208±0.032	15.41	0.163±0.035	21.43	1
Myo-inositol	9.30	9.11	0.019±0.008	41.07	0.012±0.003	28.14	1	0.050±0.017	34.60	0.018±0.008	47.54	1
Oleic acid (C18:1)*	14.49	13.36	0.119±0.035	29.29	0.091±0.029	31.69	1	0.133±0.029	21.68	0.086±0.015	26.65	1
Linoleic acid (C18:2ω-6)*	17.07	13.57	0.180±0.052	29.22	0.122±0.046	37.38	1	0.192±0.066	29.30	0.170±0.030	30.92	0
Stearic acid (C18:0)*	11.06	10.61	0.097±0.016	16.63	0.045±0.008	17.87	1	0.115±0.023	15.98	0.070±0.023	16.17	1
Arachidonic acid (C20:4ω-6)*	9.21	6.34	0.023±0.010	41.47	0.010±0.004	42.18	1	0.031±0.010	31.65	0.010±0.004	32.56	1
Cholesterol*	8.34	8.53	0.463±0.050	10.70	0.437±0.108	24.61	0	0.413±0.046	11.21	0.349±0.048	13.89	1

*identified by standard substances. Concentrations are presented as mean±SD by the ratio of its peaks area to that of the internal standard on the same chromatogram.A p value of <0.05 is considered statistically significant and signed T value is “1”, otherwise “0”.

Table 2. Asoociated diseases of each metabolite and its suggesting fluids

Compounds	HMDB Hyperlink	Associated diseases	Suggesting Fluids
lactate	HMDB00190	Type 1diabetes mellitus [19]	Serum
		Hepatic and biliary malignancies[20]	
		Dementia[21]	
sarcosine	HMDB00271	Sarcosinemia[22]	
Alanine Glycine Isoleucine Serine Threonine	HMDB00161 HMDB00123 HMDB00172 HMDB00187 HMDB00167	Alzheimer's disease[23]	
		Schizophrenia[24]	
		Epilepsy[25]	
		Leukemia[26]	
		Heart failure[27]	
urea	HMDB00294	Cirrhosis[28]	
		Meningitis[29]	
phosphate	HMDB01429	Hypophosphatemia[30]	
glycerate	HMDB00139	Primary hyperoxaluria[31]	
hexadecanoic acid	HMDB00220	Cardiovascular diseases[32]	
myo-inositol	HMDB00211	Alzheimer's disease[33]	
		Cachexia[34]	
oleic acid	HMDB00207	Gestational diabetes[35]	
stearic acid	HMDB00827	Nasopharyngeal carcinoma[4]	
arachidonic acid	HMDB01043	Hypertension[36]	
		Gestational diabetes[35]	
oxalic acid	HMDB02329	Hemodialysis[37]	Plasma
proline	HMDB00162	Alzheimer's disease[23]	
		Hemodialysis[38]	
2,3,4-trihydroxybutyrate	HMDB00943	Childhood obesity[39]	
citric acid	HMDB00094	Canavan disease[40]	
		Paraquat poisoning[41]	
		Rhabdomyolysis[42]	
		Type 2 diabetes mellitus[43]	
1,5-anhydro- sorbitol	HMDB02712	Alzheimer's disease[33]	
		Type 2 diabetes mellitus[44]	
fructose	HMDB00660	Type 2 diabetes mellitus[45]	

Figure caption

Figure 1. GC-MS TICs of serum (a) and plasma (b).

Figure 2. 2D-Projection plots of PLS-DA for the discrimination between serum and plasma of 2011 (a) and 2013 (c) and absolute value of coefficients (β) of each metabolite for PLS discrimination between serum and plasma of 2011(b) and 2013 (d). a, c, scores of the first two latent variables (PLS-1 and PLS-2) were ploted agaist each other. Each point indicated either a serum (2011, blue triangle; 2013, blue circle) or plasma (2011, red suare; 2013 red star) sample.b, d, the higher the absolute value of β is, the more the influence of correspounding variables is. In turn, these compounds corresponding to these variables might be key variables to discriminate sera and plasma.

Figure 3. 2D-Projection plots of PLS-DA for the discrimination between all sera and plasma (a) samples collected from 2011 and 2013, and absolute value of coefficients (β) of each metabolite for PLS discrimination (b). a, each point indicated either a serum (2011, blue triangle; 2013, blue circle) or plasma (2011, red suare; 2013 red star) sample.

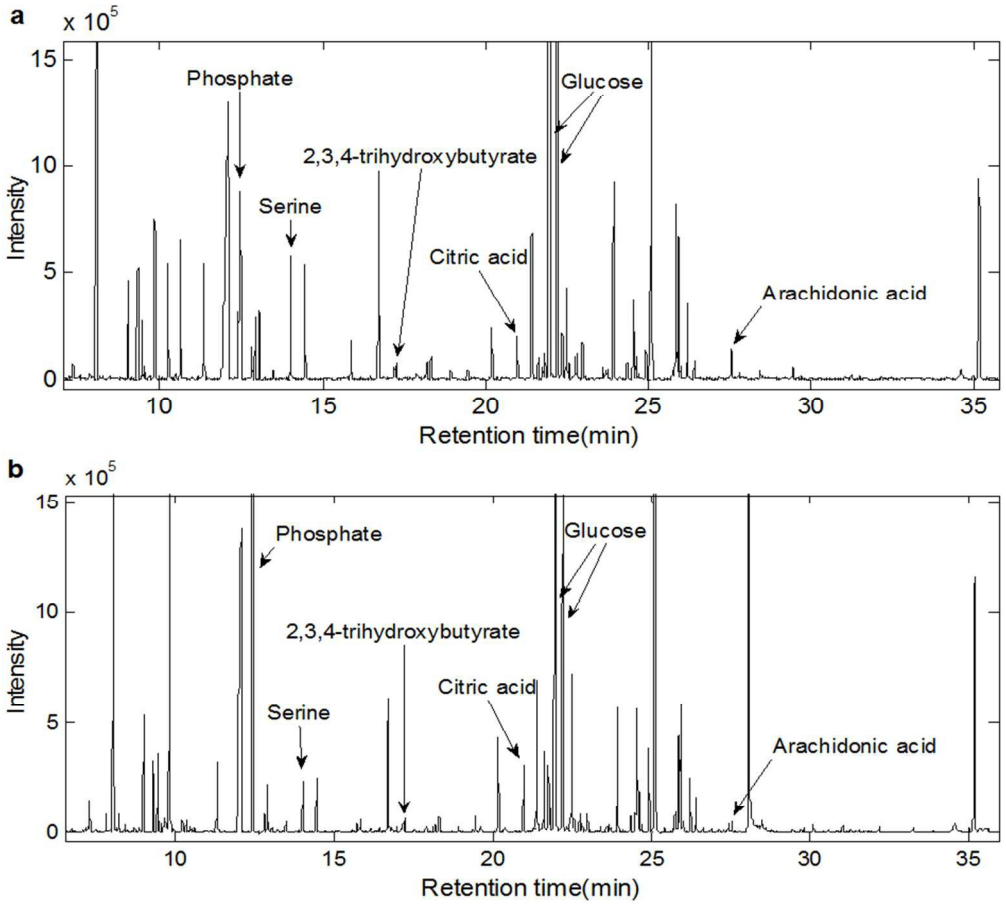


Figure 1. GC-MS TICs of serum (a) and plasma (b).
79x72mm (300 x 300 DPI)

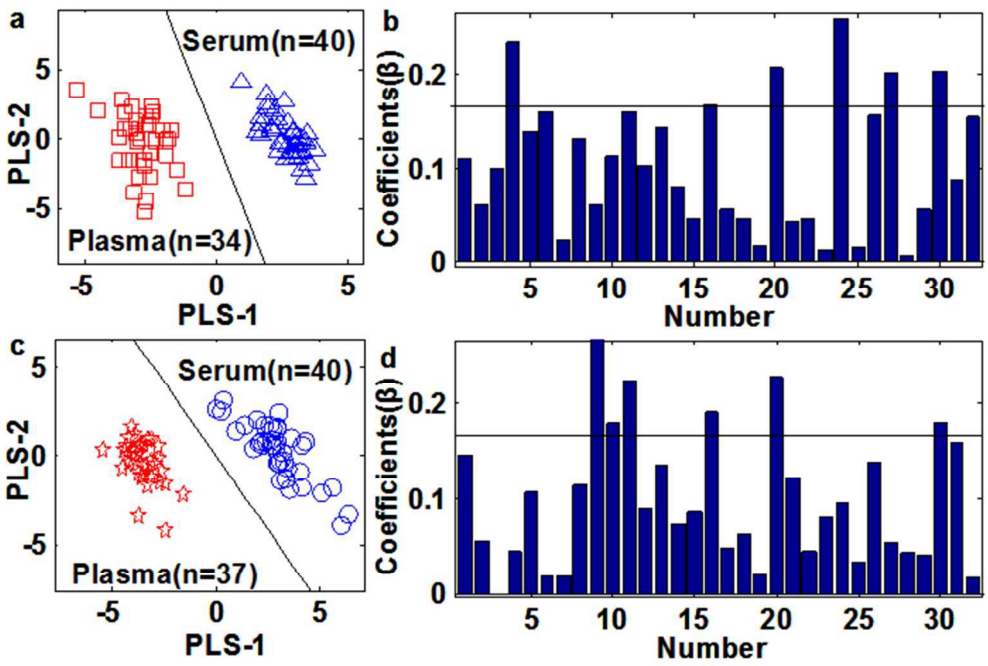


Figure 2. 2D-Projection plots of PLS-DA for the discrimination between serum and plasma of 2011 (a) and 2013 (c) and absolute value of coefficients (β) of each metabolite for PLS discrimination between serum and plasma of 2011(b) and 2013 (d). a, c, scores of the first two latent variables (PLS-1 and PLS-2) were ploted againt each other. Each point indicated either a serum (2011, blue triangle; 2013, blue circle) or plasma (2011, red suare; 2013 red star) sample.b, d, the higher the absolute value of β is, the more the influence of correspounding variables is. In turn, these compounds corresponding to these variables might be key variables to discriminate sera and plasma.
60x39mm (300 x 300 DPI)

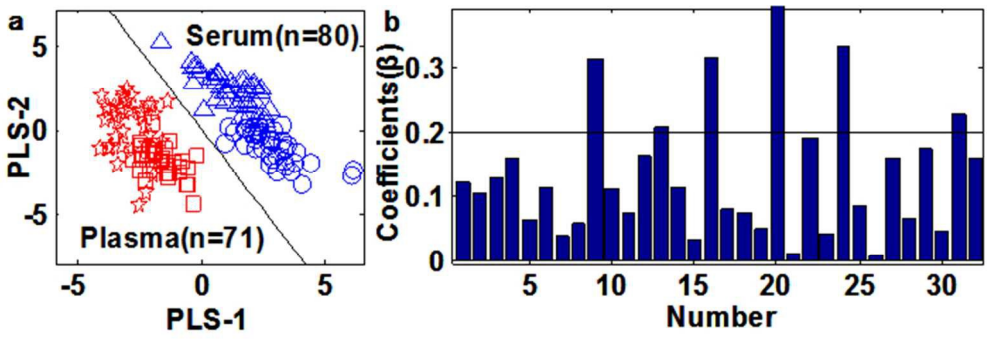


Figure 3. 2D-Projection plots of PLS-DA for the discrimination between all sera and plasma (a) samples collected from 2011 and 2013, and absolute value of coefficients (β) of each metabolite for PLS discrimination (b). a, each point indicated either a serum (2011, blue triangle; 2013, blue circle) or plasma (2011, red suare; 2013 red star) sample.
60x19mm (300 x 300 DPI)