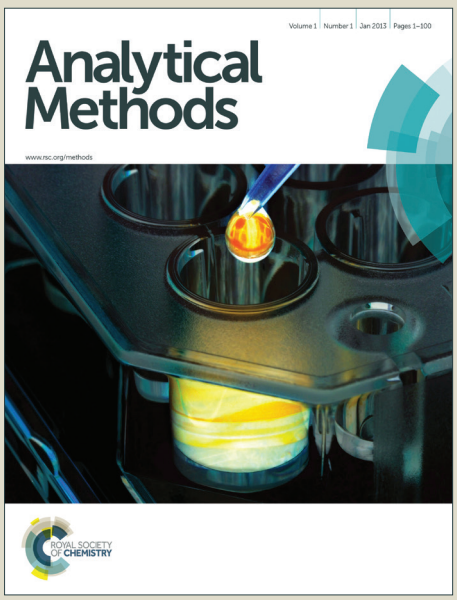
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

MeOx-TMS derivation for GC/MS metabolic profiling of urine and the application to the discrimination between normal C57BL/6J and type 2 diabetic KK-Ay mice

Lunzhao Yi[a*,b], Shuting Shi[b], Zhibiao Yi[c], Ruihua He[b], Hongmei Lu[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*

[c]*Dongguan Mathematical and Engineering Academy of Chinese Medicine, GuangZhou University of Chinese Medicine, Dongguan, 523808, PR China*

Abstract

Derivation of metabolites is inevitable for GC/MS based global metabolic profiling. The article reports a GC/MS-based protocol using methoximation followed by silylation with BSTFA+1%TMCS for analysis of urine metabolites, which is thoroughly developed and optimized from derivatization to detection. The obtained chromatograms were much cleaner due to the absence of multi-peaks of sugars, such as glucose. Validation was performed with chemical standards and urine samples and proved the methodology to be efficient, rapid and reliable with linear responses, low detection limits and good precision and recovery. The method was successfully applied to characterize the metabolic phenotype of type 2 diabetic KK-Ay mice. Partial least squares-discriminant analysis (PLS-DA) and t-test analysis illustrated that there were seven metabolites (glyceric acid, hippuric acid, glucose, sorbitol,

* Corresponding author: Tel.: +86-731-88830824; Fax: +86-731-88830831.
E-mail address: yilunzhao@mail.csu.edu.cn (Lunzhao Yi).

galactonic acid, myo-inositol, turanose) having distinct differences between normal C57BL/6J and type 2 diabetic KK-Ay mice.

Keywords: MeOx-TMS derivation; GC/MS; Metabolic profiling; Type 2 diabetes mellitus

1. Introduction

Metabolomics is defined as "the quantitative measurement of the dynamic multi-parametric responses of a living system to pathophysiological stimuli or genetic modification"[1], which depend on the ability to describe the changes of low molecular weight metabolites in various biofluids, such as plasma, urine and cerebrospinal fluid, etc. [2-4]. Because NMR and MS based platforms could produce comprehensive profiles of metabolites from the biological samples, these analytical methods had been widely used in metabolomics [5-7].

Among these analytical techniques, gas chromatography/mass spectrometry (GC/MS) is a relatively low cost alternative that provides high separation efficiency to resolve the complex biological mixtures, and gained increased implementation recently in performing the global metabolic profiles[8-13]. Furthermore, GC/MS is a mature technology applicable to a large number of samples. It is relatively easy to perform peak identification and prediction compared with other technical platforms such as liquid chromatography- and capillary electrophoresis-MS [14]. However, because many metabolites are polar and heat-labile, the pretreatment of derivation before GC/MS analysis is inevitable. There are many derivation methods to resolve these problems,

such as trimethylsilylation (TMS) derivation [13], propyl chloroformate (PCF) derivation [15], methoximation reaction (MeOx-derivation) [16], etc.. TMS-derivation is the most commonly used derivation method for the chemical components with active H, such as –OH, -COOH, -NH_2, -SH. This pretreatment method has being widely employed in metabolomics research nowadays[17-19]. In addition, the combination of methoximation followed by TMS (MeOx-TMS method) was developed and applied in metabolomics [11].

Type 2 diabetes mellitus (T2DM) is a complex and heterogeneous metabolic disorder disease, which is characterized by both impaired insulin secretion and insulin resistance [20-22] . There are predictions that T2DM could increase worldwide to more than 250 million individuals within the next decade or so. Despite this, we still are unclear as to its causes and optimal treatment. The research of metabolic phenotype of T2DM will help us to understand its pathogenesis to some extent. Because the collection of urine is non-invasive and convenient, urinary metabolomics of type 2 diabetic KK-Ay mice is employed to represent the metabolic characteristics of T2DM. After optimization of MeOx-TMS reaction conditions, we described a simple and stable method to derivate several kinds of metabolites in urine step by step. A gas chromatography–mass spectrometry (GC–MS) profiling protocol was used to isolate and identify the detected metabolites and establish stable metabolic profiles of C57BL/6J and type 2 diabetic KK-Ay mice. The endogenous metabolites of mice urine were identified by the standard components and the NIST library. Furthermore, partial least squares-discriminant analysis (PLS-DA) was applied to discriminate the

C57BL/6J and type 2 diabetic KK-Ay mice, and to screen the key metabolites (potential biomarkers).

2. Experimental

2.1. Chemicals and reagents

N,O-Bis(trimethylsilyl) trifluoroacetamide with 1% trimethylchorosilane (BSTFA+1%TMCS), pyridine, methoxyamine hydrochloride, ribitol (internal standard, IS), urease, lactic acid, L-threonine, succinic acid, cis-aconitic acid, isocitric acid, fructose, glucose, sorbitol, palmitic acid and lactose were purchased from Sigma-Aldrich (St. Louis, MO, USA). Methanol was analytical grade and purchased from Changsha Fufan Trade Ltd. (Changsha, China). Ribitol was used as internal standard, urease was used to decompose and remove the excess urea. The methoxyamine hydrochloride was dissolved in the pyridine at a concentration of 15 mg/mL. The urease was dissolved in the ultra-pure water at a concentration of 5 mg/mL.

2.2. Experimental animals

C57BL/6J and the KK-Ay mice specimens were purchased from Beijing HuaFuKang biological technology Co., LTD. Until the urine samples were collected, the experimental mice were raised by institute of laboratory animal science, Chinese Academy of Medical Sciences (CAMS) (Beijing, China). The barrier environment conditions were as followed: temperature, 20–26°C; humidity, 40–70%; aeration frequency, 10~15 times/h; illuminance, 150~300 Lux; and a 12h light and dark cycle

(lighting from 07:00 to 19:00). The mice were fed a standard mouse diet and given water *ad libitum*. Urine samples were taken at the same time each morning and were stored at -80℃ prior to analysis. All experiments were performed in compliance with the relevant laws and institutional guidelines. The study was approved by institute of laboratory animal science, Chinese Academy of Medical Sciences (CAMS) (Beijing, China).

2.3. 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). In the gas chromatographic system, a DB-5MS fused-silica capillary column (0.25 mm × 30 m × 0.25 μm, Agilent, USA) was used. The helium carrier gas flow rate was 1.0 mL/min. The column initial temperature was kept at 70 ºC for 5 min. Then the temperature was ramped at a rate of 20ºC /min to 160 ºC, 4 ºC /min to 180 ºC and 10 ºC /min to 300 ºC, and held for 1.5min at 300ºC. 1μL of the metabolite derivative solution with the derivatization reagent was run through the gas chromatograph-mass spectrometer with a 10:1 split throughout. The injector temperature was 280ºC, the septum purge flow rate was 3mL /min, and the purge was turned on all the time. The total GC run time was 28min. The interface temperature was 250 ºC and ion source temperature was 200 ºC. 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 0.9 kV of detector voltage was turned on after a solvent delay of 5 min.

2.4. Preparation of stock and working solutions

Stock solutions of lactic acid, succinic acid, isocitric acid, glucose, sorbitol and lactose were prepared in methanol with concentrations of 2 mg/mL, respectively. Then the six standard samples were mixed together to obtain the mixed standard with concentration of 2 mg/mL, and further diluted into 0.005-2 mg/mL as working solutions. The concentration of ribitol-methanol solution was 2 mg/mL.

2.5. Preparation of standard and quality control (QC) samples

The calibration standard samples were prepared by diluting the mixed standard into the concentration of 2.0, 1.0, 0.5, 0.1, 0.05 0.025, 0.010, 0.005mg/mL, and processed as described in the sample preparation. Quality control (QC) samples (for accuracy and precision) were prepared at a concentration of 0.025 mg/mL (low), 0.1 mg/mL (medium) and 1.0 mg/mL (high) for lactic acid, succinic acid, isocitric acid, glucose, sorbitol and lactose. All standard stock solutions were stored at -20℃.

2.6. Sample preparation

An aliquot (300 μL) of mice urine was centrifuged for 5 min at 16000 rpm at 4 °C. 150 μL supernatant was transferred to a 2 mL plastic centrifugation tube and 150 μL, 5 mg/mL urease solution was added to remove and decompose excess urea. After the mixture was incubated at room temperature for 30 min, internal standard (I.S.) working solution (2 mg/mL ribitol-methanol, 50 μL) and protein precipitant (methanol, 800 μL) were added and mixed by vigorously vortexing for 1 min, following room temperature reaction lasting 10 min. Then the mixture was centrifuged at 4 °C, 16000 rpm for 5 min, 150 μL supernatant was taken out and

evaporated to dryness in a vacuum oven. Methoximation was carried out at 70 ºC for 30 min using methoxyamine hydrochloride in pyridine (40 μL, 15 mg/mL in pyridine), then the samples were trimethylsilylated by 60 μL BSTFA, containing 1% TMCS as a catalyst, at 70 ºC for another 30 min before GC-MS analysis.

2.7. Validation of the method

Calibration curves were established from peak area ratios (analyte/IS) versus nominal concentrations using linear least-squares regression model ($1/X^2$ weighting). Intra- and inter-day precisions were determined by assessing measured results of QC samples at low, medium and high concentrations. Precisions were expressed by the relative standard deviation (R.S.D, %). Extraction recoveries were determined by comparing the ratio of the analytes' peak areas of the extracted QC samples with those of un-extracted standard solutions at the same nominal concentrations. Stability was checked by comparing measured results with those of freshly prepared samples. The short- and long-term stabilities were evaluated by analyzing urine samples kept at room temperature (about 25 ºC) for 4 h and in the freezer (−80 ºC) for 3 months, respectively; the freeze–thaw stability was carried out by detecting urine samples undergoing three freeze (−80 ºC)–thaw (room temperature) cycles; the post-preparation stability was assessed by determining the extracted urine samples stored under 4 ºC (in refrigerator) for 24 h.

2.8. 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). 39 peaks were considered to

be endogenous metabolites. Lactic acid, L-threonine, succinic acid, cis-aconitic acid, isocitric acid, fructose, glucose, sorbitol, palmitic acid and lactose were identified by their chemical standard substances. As for the quantification, peak integration was employed. To normalize the urinary data, the variables were expressed as the ratio of peak area of corresponding metabolites to that of the internal standard on the same total ion chromatogram (TIC). 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 autoscaled, then, was analyzed by PLS-DA. The PLS-DA models were constructed to establish the significance of the difference between the C57BL/6J and KK-Ay mice. For the purpose of cross-validation and due to the small number of samples, ten-fold cross validation was applied. The 10 validation PLS-DA models were calculated excluding 10% of the mice samples in the validation model. Class membership was predicted using discriminant line between two classes obtained by linear discriminant analysis (LDA). In order to screen out the key metabolites contributed to the metabolic disorder of T2DM mice (potential biomarkers), the coefficients of PLS transformed equation were calculated. The discriminant equation expressed by latent variables obtained by PLS could be transformed to that expressed by original variables. The detailed deduction process was reported by Yi et al.[21]. Briefly, if the first two PLS latent variables (LVs) were extracted to establish the LDA model, the equation of the discrimination line could be expressed as follows:

$$\mathbf{t}_1*\alpha_1+\mathbf{t}_2*\alpha_2=c \qquad (1)$$

where $\mathbf{t}_1$ and $\mathbf{t}_2$ are the first two LVs obtained by PLS decomposition, α_1 and α_2 are the coefficients of discrimination line equation, c is a constant. Equ. 1 can be written as matrix form:

$$\mathbf{T}*\boldsymbol{\alpha}=c \qquad (2)$$

where $\mathbf{T}=[\mathbf{t}_1\ \mathbf{t}_2]$, $\alpha=[\alpha_1\ \alpha_2\ ]^t$, subscript "t" stand for transport.

As we know, each latent variables $\mathbf{t}_i$ is the linear combination of column vectors of **X**, that is

$$\mathbf{T}=\mathbf{X}*\mathbf{H} \qquad (3)$$

where **X** is the data matrix, **H** is the weight matrix obtained by PLS decomposition.

Inserting Equ. (3) into Equ. (2),

$$\mathbf{X}*\mathbf{H}*\boldsymbol{\alpha}=c \qquad (4)$$

Donated $\boldsymbol{\beta}=\mathbf{H}*\boldsymbol{\alpha}$, so

$$\mathbf{X}*\boldsymbol{\beta}=c \qquad (5)$$

Then, the line equation expressed by PLS LVs (scores) was transformed to that expressed by original variables. In the transformed equation (Equ.5), the absolute values of coefficients ($\boldsymbol{\beta}$) can render the influence of corresponding variables on separation between sample classes. In turn, these compounds corresponding to these variables might be likely candidates for biomarkers.

The statistical analysis was performed using the in-house software written in MATLAB (version 6.5, The MathWorks, Natick, MA, USA) .

3. Results and discussion

3.1. Selection of derivation method

Gas chromatography/mass spectrometry (GC/MS) is one of the most frequently used tools to analyze a variety of metabolites in metabolomics because of its high sensitivity, high resolution and good reproducibility. However, many metabolites are thermally unstable, such as sugars, organic acids, amino acids, fatty acids, alcohols, etc. They cannot be analyzed without derivation. Trimethylsilylation (TMS-derivation) is the most commonly used derivation method for the chemical components with active H. It is a mild and universal reaction to increase the volatility of non-volatile hydrophilic compounds by exchanging their acidic protons. However, our researches demonstrated that silylation is not enough for comprehensive analysis of metabolites. For example, there are more than twenty peaks for glucose after trimethylsilylation in a total ion chromatogram (TIC), which is shown in Fig.1 (A). The TMS derives of glucose are very complicated, including different isomers with or without ring, five or six carbons. In order to reduce the number of derivatized products, another deriviation method, methoximation reaction (MeOx-derivation) was employed before TMS-derivation[11]. Methoximation is performed to inhibit ring formation of reducing sugars by protecting aldehyde and ketone groups. It is usually needed for opening ring sugars and results in two stereoisomers, which will later be separated in GC-MS. The result of MeOx-TMS derivation of glucose is shown in Fig.1 (B). The reaction equations of methoximation (1) and trimethylsilylation (2) of glucose are shown in Fig.2. There are only two stereoisomers of glucose derivatives after MeOx-TMS derivation. Most metabolites in urine are organic compounds. Many of them not only have active H, but also carbonyl. For that reason, the MeOx-TMS derivation method

was applied to the analysis of urine metabolites. Furthermore, several kinds of metabolites, lactic acid, succinic acid, isocitric acid, glucose, sorbitol and lactose, were employed to evaluate the efficiency of MeOx-TMS derivation.

Insert Figure 1

Insert Figure 2

3.2. Chromatographic conditions and selectivity

First of all, one should find a suitable GC/MS condition to establish a stable quantitative metabolite profile with good separation, since there are abundant metabolites in rat urine. Those metabolic mixtures containing multi-classes components, even positional and geometrical isomers, cause the most problems in isolation and identification, as well as in GC methodology. Column selection is one of the focuses in the process of optimizing chromatographic conditions. The separation ability of DB-23, HP-1 and DB-5MS were compared and DB-5MS was selected. Furthermore, other chromatographic conditions were modified, such as temperature process, split ratio, injection temperature. Typical chromatograms of rat urine samples are shown in Fig. 3. 39 metabolites in rat urine were cleanly isolated with lower baselines within 28 min under those final chromatographic conditions.

Insert Figure 3

3.3. Linearity and detection limit

Linearity of calibration was tested and assayed in consecutive 5 days. Calibration curves in the concentration range of 0.005–2.0 mg/mL for lactic acid, succinic acid, isocitric acid, glucose, lactose and 0.005-0.5 mg/mL for sorbitol were constructed by

plotting the peak area ratios of analyte/I.S. to the spiked concentrations. The linearity was determined from linear regression analysis on the calibration curves. The detection limits of lactic acid, succinic acid, isocitric acid and lactose were 0.0025mg/mL. The detection limits of glucose and sorbitol were 0.001mg/mL.

3.4. Precision and recovery

Intra-day accuracy and precision (each, n=5) were evaluated by analysis of QC samples at different times during the same day; the precisions obtained were 3.63-6.24% (RSD). Inter-day accuracy and precision were determined by repeated analysis of QC samples over 5 consecutive days (n = 1 series per day), obtained the precision were 6.51-8.28% (RSD). The results suggested that the accuracy and precision in the present assay are acceptable for the analysis.

The extraction and the MeOx-TMS derivation efficiency for lactic acid, succinic acid, isocitric acid, glucose, sorbitol, lactose and ribitol (IS) in rat urine with methanol as precipitant in sample preparation were consistent, precise and reproducible. The yields of the mixed standards were 80.1-107.1% (low, n=5), 86.7-105.3% (medium, n=5), 84.6-97.2% (high, n=5) with the RSD values being less than 7% at each QC level by comparing the peak areas of analytes and IS. Furthermore, the mixed standards were spiked to the rat urine samples to evaluate yields of the reaction process under matrix effect and obtained the reaction yields of 82.1-111.2% (low, n=3), 85.0-108.3% (medium, n=3), 83.2-112.2% (high, n=3), respectively, with the RSD values being less than 10%. Additionally, internal standard, ribitol, was used to evaluate the extraction and reaction yields with matrix effect, which the mean values

were 92.2% (low, n=3), 96.1% (medium, n=3), 84.6% (high, n=3) with the RSD values being less than 8%.

3.5. Stability

The stability results illustrated that lactic acid, succinic acid, isocitric acid, glucose, sorbitol, lactose are stable in rat urine for at least 3 months when stored at -80◦C for three freeze–thaw cycles and in the reconstituted solutions when stored under 4 ºC for 24 h.

3.6. The application to the discrimination between C57BL/6J and the KK-Ay mice

In this research, 39 main metabolites, involved in metabolism of sugars, organic acids, amino acids, fatty acids, *et al*., were identified to discriminate the metabolic phenotype of type 2 diabetic KK-Ay mice. Table 1 showed the quantitative metabolic profile of these 39 metabolites in urine of both C57BL/6J and the KK-Ay mice (relative quantity of each metabolite to internal standard). As shown in table 1, while 20 metabolites remained unchanged between C57BL/6J and the KK-Ay mice (t-test $p>0.05$ with a signed t value of "0"), there are 19 metabolites significantly changed in the urine from C57BL/6J and the KK-Ay mice (t-test $p<0.05$ with a signed t value of "1"), including ethylene glycol, acetic acid, β-aminoisobutyric acid, glyceric acid, 2,3-dihydroxybutanoic acid, 3,4-dihydroxybutanoic acid, N-(1-oxobutyl)-glycine, isovalero glycine, 2,3,4-trihydroxybutyrate, N-(1-oxohexyl)-glycine, cis-aconitic acid,

hippuric acid, N-phenyl glycine, glucose, sorbitol, galactonic acid, myo-inositol, turanose, lactose. These results indicated that the metabolic phenotype of KK-Ay mice was significantly different from that of C57BL/6J. Because of the inter-subject variations in urine matrix composition, the complexity of TICs and the differences of metabolites concentrations, it is hard to visualize these metabolic profiles just based on the GC-MS data. In this case, we employed chemometric methods of multivariate statistical analysis to construct a visible model to discriminate KK-Ay mice.

Peak areas of 39 metabolites were used as the input data of PLS-DA to establish a visible model for the discrimination of C57BL/6J and the KK-Ay mice. After comparison, the final optimized 2-dimentional PLS-DA model by the first two latent variables (PLS1 and PLS2) was obtained (Figure 4 (A)). KK-Ay mice (the hollow circles in Figure 4 (A)) and C57BL/6J (diamonds in Figure 4(A)) were homogeneously distributed. C57BL/6J and the KK-Ay mice were separated clearly by the discriminant line (Figure 4(A)) with a total recognition rate and predictive rate (10-fold cross validation) of 100% (Table 2). For type 2 diabetic KK-Ay mice, the contents of sugars were significantly higher than for C57BL/6J mice. To find other key metabolites besides sugars, a new PLS-DA model was established using 36 metabolites as input data (Figure 4(B)). Glucose, turanose and lactose were kicked out. The total recognition rate and predictive rate (10-fold cross validation) were 100% and 93.75% (Table 2). The result illustrated that the rest 36 metabolites also have good discrimination ability.

Insert Figure 4

Insert Table 2

In the next step, we applied the absolute values of coefficients (β) of the metabolites to identify which metabolites were conspicuously contributed to the metabolic differences between C57BL/6J and the KK-Ay mice. It is very interesting that the only difference of the screened metabolites between the two PLS-DA models (39 variables and 36 variables) is glucose and turanose. That is to say, the screened key metabolites were all the same except sugars for the two PLS-DA models. As shown in figure 5, seven metabolites were classified with higher coefficients, including glyceric acid, hippuric acid, glucose, sorbitol, galactonic acid, myo-inositol, turanose, which were concordant with the t-test results (t signed as ‘‘1’’, Table 1). Using the seven metabolites as input data to establish a PLS-DA model, the recognition rate is 100% (Table 2). These results indicated that these seven metabolites might be more correlated with the metabolic disturbances of type 2 diabetes and possibly be served as biomarker candidates of T2DM.

Insert Figure 5

4. Conclusions

Metabolomics is a growing research field where new protocols are rapidly developed and new applications discovered. However, the development of such protocols rarely includes a systematic optimization followed by validation with chemical standards and real samples. In this paper, we described a simple and stable MeOx-TMS method to derivate different kinds of metabolites from urine. A GC/MS profiling protocol was

used to isolate and identify the detected metabolites and establish stable metabolic profiles of C57BL/6J and type 2 diabetic KK-Ay mice. 39 endogenous metabolites of mice urine were identified by the standard components and the NIST library. Furthermore, discriminant model between C57BL/6J and type 2 diabetic KK-Ay mice was establish by PLS-DA, and glyceric acid, hippuric acid, glucose, sorbitol, galactonic acid, myo-inositol, turanose were screened out as the key metabolites (potential biomarkers).

Acknowledgment

This work was supported financially by National Nature Foundation Committee of China (21105129 and 21175157), Science and Technological Program for Dongguan′s Higher Education, Science and Research, and Health Care Institutions （2012108102032）, Special Foundation of China Postdoctoral Science (200902481) and the Fundamental Research Funds for the Central Universities (2011QNZT053).

References

1. J. K. Nicholson, J. C. Lindon and E. Holmes, *Xenobiotica*, 1999, **29**, 1181.
2. L. Akesson, J. Trygg, J. M. Fuller, R. Madsen, J. Gabrielsson, S. Bruce, H. Stenlund, T. Tupling, R. Pefley, T. Lundstedt, A. Lernmark and T. Moritz, *Metabolomics*, 2011, **7**, 593.
3. J. Yang, G. W. Xu, Y. F. Zheng, H. W. Kong, T. Pang, S. Lv and Q. Yang, *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 2004, **813**, 59.
4. M. Coen, Y. S. Hong, T. A. Clayton, C. M. Rohde, J. T. Pearce, M. D. Reily, D. G. Robertson, E. Holmes, J. C. Lindon and J. K. Nicholson, *Journal of Proteome Research*, 2007, **6**, 2711.
5. H. Wu, T. Liu, C. Ma, R. Xue, C. Deng, H. Zeng and X. Shen, *Analytical and Bioanalytical Chemistry*, 2011, **401**, 635.
6. M. Coen, E. Holmes, J. C. Lindon and J. K. Nicholson, *Chemical Research in Toxicology*, 2008, **21**, 9.
7. S. H. Lee, H. M. Woo, B. H. Jung, J. Lee, O. S. Kwon, H. S. Pyo, M. H. Choi and B. C. Chung, *Analytical Chemistry*, 2007, **79**, 6102.

8. K. Bando, R. Kawahara, T. Kunimatsu, J. Sakai, J. Kimura, H. Funabashi, T. Seki, T. Bamba and E. Fukusaki, *Journal of Bioscience and Bioengineering*, 2010, **110**, 491.
9. H. Tsugawa, T. Bamba, M. Shinohara, S. Nishiumi, M. Yoshida and E. Fukusaki, *Journal of Bioscience and Bioengineering*, 2011, **112**, 292.
10. Y. Qiu, M. Su, Y. Liu, M. Chen, J. Gu, J. Zhang and W. Jia, *Analytica Chimica Acta*, 2007, **583**, 277.
11. Q. Zhang, G. Wang, Y. Du, L. Zhu and A. Jiye, *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 2007, **854**, 20.
12. X. Huang, L. Shao, W. Gong, Y. Mao, C. Liu, H. Qu and Y. Cheng, *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences*, 2008, **870**, 178.
13. A. P. H. Danielsson, T. Moritz, H. Mulder and P. Spegel, *Metabolomics*, 2012, **8**, 50.
14. O. Fiehn, *Trac-Trends in Analytical Chemistry*, 2008, **27**, 261.
15. X. Zheng, Y. Qiu, W. Zhong, S. Baxter, M. Su, Q. Li, G. Xie, B. M. Ore, S. Qiao, M. D. Spencer, S. H. Zeisel, Z. Zhou, A. Zhao and W. Jia, *Metabolomics*, 2013, **9**, 818.
16. S. Abate, Y. G. Ahn, T. Kind, T. R. I. Cataldi and O. Fiehn, *Rapid Communications in Mass Spectrometry*, 2010, **24**, 1172.
17. M. Zeng, Y. Liang, H. Li, M. Wang, B. Wang, X. Chen, N. Zhou, D. Cao and J. Wu, *Journal of Pharmaceutical and Biomedical Analysis*, 2010, **52**, 265.
18. L. Coulier, B. Muilwijk, S. Bijlsma, M. Noga, M. Tienstra, A. Attali, H. van Aken, E. Suidgeest, T. Tuinstra, T. M. Luider, T. Hankemeier and I. Bobeldijk, *Metabolomics*, 2013, **9**, 78.
19. Z. Hong, Z. Lin, Y. Liu, G. Tan, Z. Lou, Z. Zhu, Y. Chai, G. Fan, J. Zhang and L. Zhang, *Journal of Chromatography A*, 2012, **1254**, 14.
20. B. S. Baliga and V. A. Fonseca, *American family physician*, 1997, **55**, 817.
21. L.-Z. Yi, J. He, Y.-Z. Liang, D.-L. Yuan and F.-T. Chau, *Febs Letters*, 2006, **580**, 6837.
22. S. I. Taylor, *Cell*, 1999, **97**, 9.

Figure Legend

Fig.1	GC/MS total ion chromatograms (TICs) of glucose after trimethylsilylation (Fig.1 (A)) and MeOx-TMS derivation (Fig.1 (B)).
Fig.2	The reaction equations of methoximation (1) and trimethylsilylation (2) of glucose.
Fig.3	Metabolic profile of mice urine is illustrated as the total ion chromatogram (TIC) in GC/MS analysis. The whole profile is shown in subfigure A and specific areas of it are magnified in subfigures B and C. Overloaded metabolites: 1. Ethylene glycol, 2. N, N-Diethylacetamide, 3. Lactic acid, 4. Acetic acid, 5. Phosphate, 6. L-Threonine, 7. β-Aminoisobutyric acid, 8. Succinic acid, 9. 1,2-Hydroquinone, 10. Glyceric acid, 11. 2,3-Dihydroxybutanoic acid, 12. 2,4-Dihyoxybutanoic acid, 13. 3,4-Dihydroxybutanoic acid, 14. N-(1-oxobutyl)-Glycine, 15. Isovalero glycine, 16. Threitol, 17. N-Crotonyl glycine, 18. 2,3,4-Trihydroxybutyrate, 19. N-(1-oxohexyl)-glycine, 20. (3-Hydroxyphenyl) acetic acid, 21. Xylose, 22. Ribose, 23. Arabitol, 24. 6-Deoxy-D-Galactose, 25. Mannonic acid, 26. cis-Aconitic acid, 27. Isocitric acid, 28. Hippuric acid, 29. Fructose, 30. N-Phenyl glycine, 31. Glucose, 32. Altronic acid, 33. Sorbitol, 34. Galactonic acid, 35. Palmitic acid, 36. Myo-Inositol, 37. Turanose, 38. Lactose monohydrate, 39. Lactose. An asterisk behind the peak number indicates this metabolite has been further identified by the commercial standards.

Fig.4	Projection plots of metabolites from PLS-DA of the first two latent variables (PLS-1 and PLS-2) for the eight C57BL/6J and eight KK-Ay mice. Figure 4(A) and (B) were using 39 and 36 variables (except glucose, turanose and lactose) as input data, respectively. The discrimination line between C57BL/6J and KK-Ay mice were obtained by linear discrimination analysis (LDA). Key: ◇= C57BL/6J; ○= KK-Ay.
Fig.5	Absolute value of **β** of each variable for discrimination of C57BL/6J and KK-Ay mice. The component with an asterisk is the screened key metabolite.

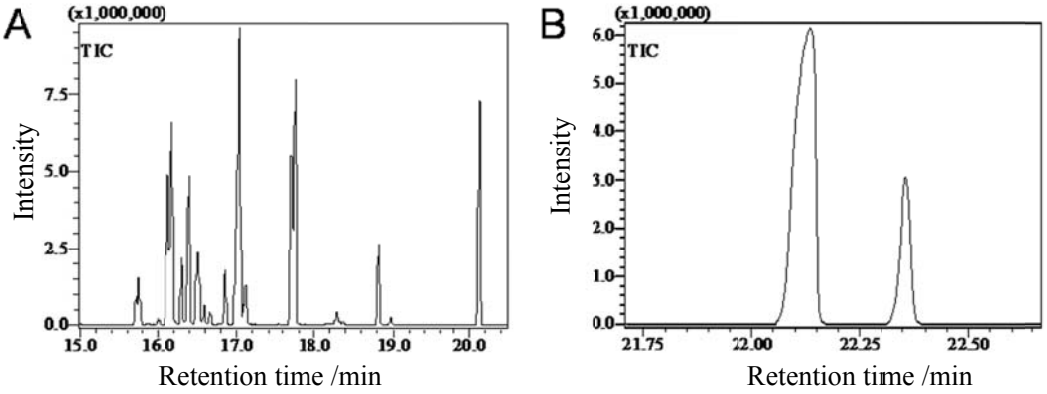


Fig.1

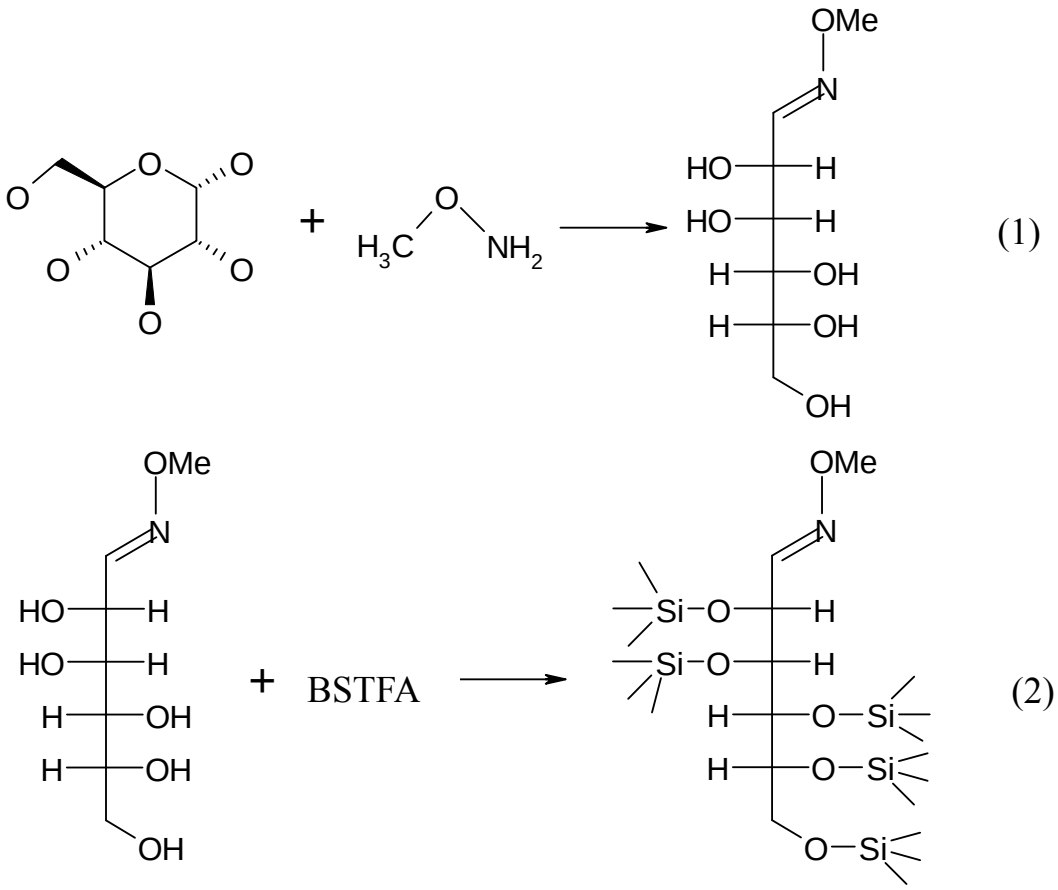


Fig.2

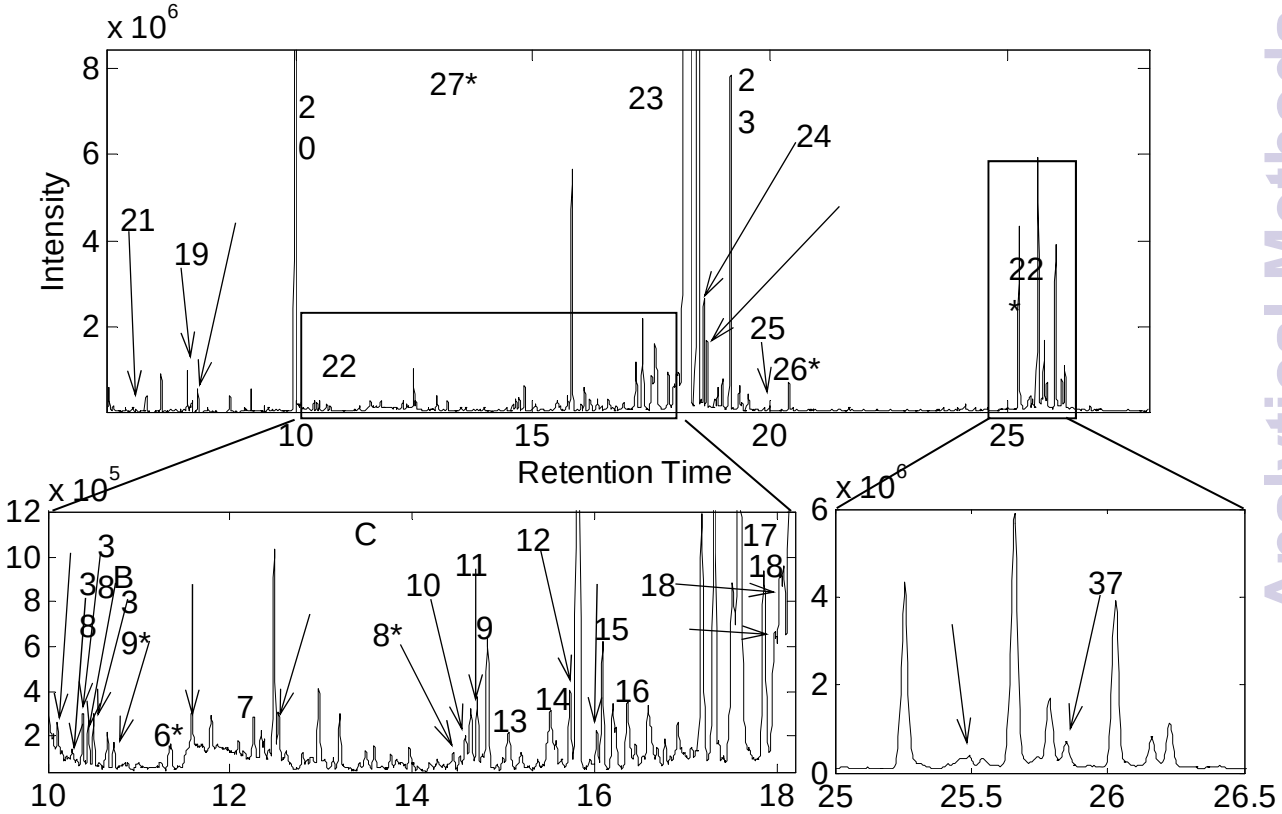


Fig.3

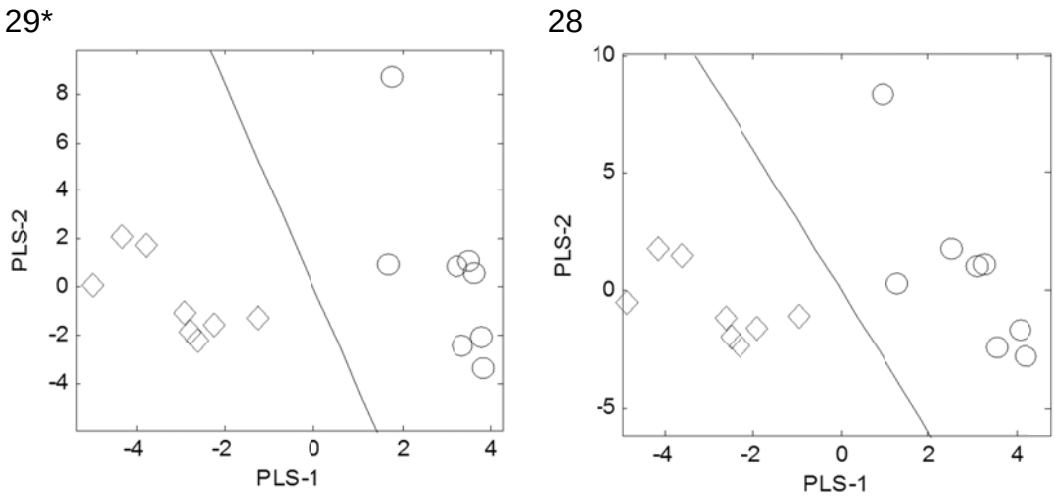


Fig.4

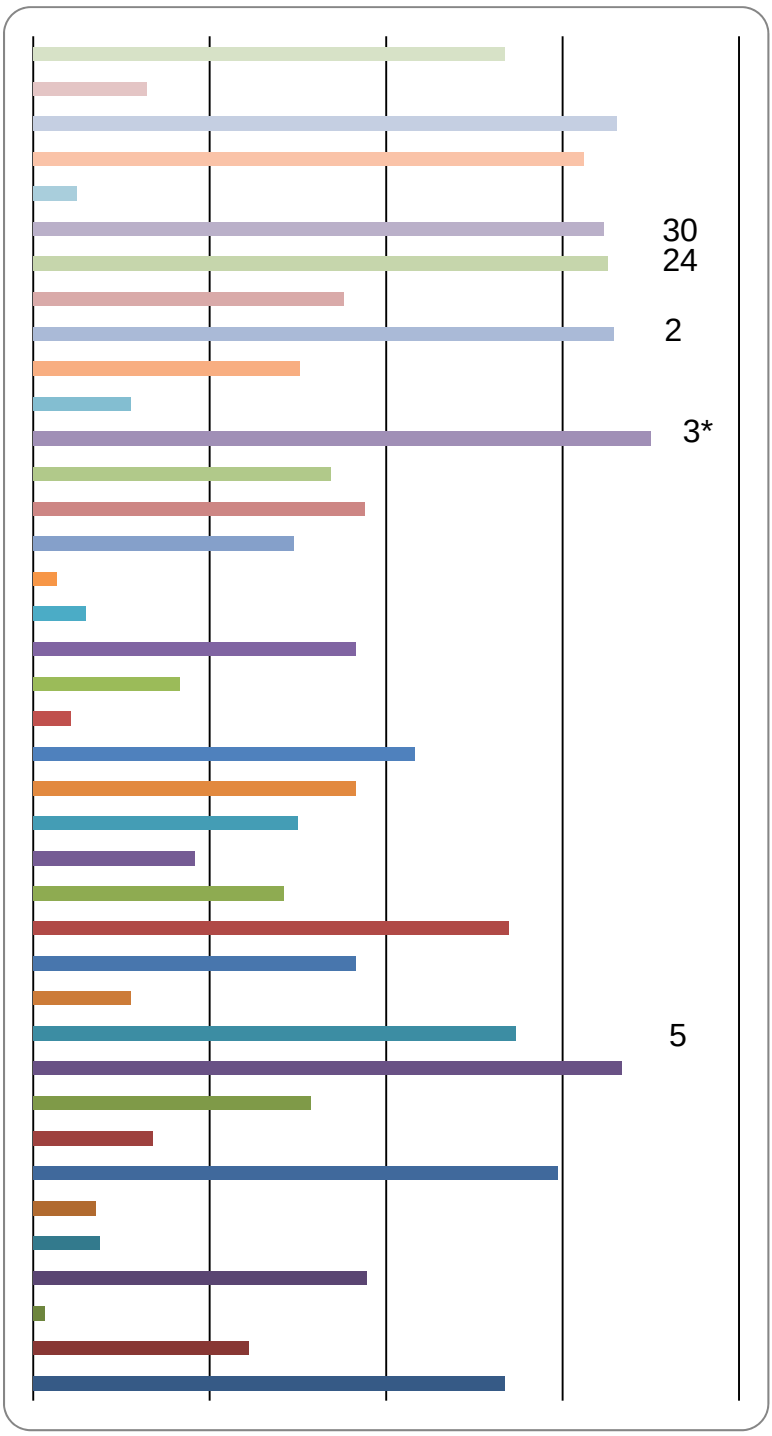


Fig.5

Table 1 Quantitative analysis of metabolic profiles of the C57BL/6J and KK-Ay mice (relative quantity of each metabolite to internal standard) and t-test results

No.	Metabolites	GC/MS m/z	Relative quantity		t	*P* value
			C57BL/6J (n=8)	KK-Ay (n=8)		
1	Ethylene glycol	73, 103, 147, 191	0.017±0.004	0.052±0.043	1	0.039
2	N, N-Diethylacetamide	43, 58, 72	0.066±0.009	0.058±0.010	0	>0.05
3	Lactic acid*	73, 117, 147	0.150±0.120	0.116±0.092	0	>0.05
4	Acetic acid	73, 160	0.086±0.033	0.049±0.023	1	0.023
5	Phosphate	73, 299, 301, 314	3.128±1.017	2.491±1.534	0	>0.05
6	L-Threonine*	73, 117, 130, 219	0.017±0.010	0.025±0.051	0	>0.05
7	β-Aminoisobutyric acid	73, 147	0.025±0.009	0.007±0.005	1	<0.001
8	Succinic acid*	73, 147, 247	0.081±0.043	0.048±0.056	0	>0.05
9	1,2-Hydroquinone	73, 112, 239, 254	0.012±0.007	0.035±0.069	0	>0.05
10	Glyceric acid	73, 147, 189	0.096±0.027	0.020±0.014	1	<0.001
11	2,3-Dihydroxybutanoic acid	73, 117, 147, 292	0.027±0.005	0.012±0.009	1	0.001
12	2,4-Dihyoxybutanoic acid	73, 103, 147, 219	0.018±0.011	0.014±0.008	0	>0.05
13	3,4-Dihydroxybutanoic acid	73, 147, 189, 233	0.030±0.010	0.016±0.011	1	0.015
14	N-(1-oxobutyl)-Glycine	73, 43, 158, 202	0.065±0.024	0.023±0.018	1	0.001
15	Isovalero glycine	73, 158, 216	0.046±0.013	0.024±0.023	1	0.039
16	Threitol	73, 103, 117, 147	0.091±0.027	0.095±0.098	0	>0.05
17	N-Crotonyl glycine	69, 73, 156, 200	0.061±0.014	0.038±0.029	0	>0.05
18	2,3,4-Trihydroxybutyrate	73, 147, 205, 220, 292	0.128±0.016	0.063±0.063	1	0.013
19	N-(1-oxohexyl)-glycine	73, 102	0.116±0.072	0.044±0.037	1	0.024
20	(3-Hydroxyphenyl) acetic acid	73, 147, 164, 281, 296	0.031±0.010	0.032±0.022	0	>0.05
21	Xylose	73, 103, 307	0.041±0.015	0.052±0.055	0	>0.05
22	Ribose	73, 103, 147, 189, 217	0.113±0.037	0.183±0.152	0	>0.05

23	Arabitol	73, 103, 147, 217, 307	0.079±0.046	0.079±0.054	0	>0.05
24	6-Deoxy-D-Galactose	73, 117	0.031±0.011	0.022±0.024	0	>0.05
25	Mannonic acid	73, 147, 217	0.045±0.022	0.068±0.067	0	>0.05
26	cis-Aconitic acid*	73, 147, 229, 285, 375	0.084±0.039	0.035±0.039	1	0.025
27	Isocitric acid*	73, 147, 273	0.465±0.412	0.18±0.099	0	>0.05
28	Hippuric acid	73, 105, 147, 206	0.067±0.043	0.297±0.092	1	<0.001
29	Fructose*	73, 103, 147, 217, 307	0.171±0.159	0.111±0.048	0	>0.05
30	N-Phenyl glycine	73, 91	0.080±0.021	0.049±0.031	1	0.035
31	Glucose*	73, 147, 160, 205, 319	0.631±0.765	53.676±39.557	1	0.002
32	Altronic acid	73, 147, 217	0.030±0.007	0.019±0.015	0	>0.05
33	Sorbitol*	73, 147, 205, 319	0.160±0.067	0.690±0.375	1	0.001
34	Galactonic acid	73, 147, 333	0.120±0.054	0.348±0.161	1	0.002
35	Palmitic acid*	73, 262, 337	0.014±0.015	0.012±0.006	0	>0.05
36	Myo-Inositol	73, 147, 217, 305	0.035±0.023	0.350±0.211	1	0.001
37	Turanose	73, 103,129, 147, 204, 217, 361	0.022±0.014	0.302±0.258	1	0.008
38	Lactose monohydrate	73, 204, 361	1.040±0.335	1.000±0.539	0	>0.05
39	Lactose*	73, 191, 204, 217	0.014±0.004	0.377±0.279	1	0.002

*Identified by chemical standards.

Table 2 Recognition ability and predictive ability [a] of the multivariate models

		Recognition ability	Predictive ability
The whole metabolites (n=39)	Sensitivity	100%	100%
	Specificity	100%	100%
	Recognition rate	100%	100%
36 metabolites (except glucose, turanose and lactose)	Sensitivity	100%	87.5%
	Specificity	100%	100%
	Recognition rate	100%	93.75%
The screened 7 metabolites	Sensitivity	100%	100%
	Specificity	100%	100%
	Recognition rate	100%	100%

[a] Recognition rate is the correct classification of the training set. Predictive ability of the multivariate models was evaluated by 10-fold cross validation. Sensitivity is the number of true positives classified as positive. Specificity is the number of true negative classified as negative.