Journal of Chromatography B, 878 (2010) 2817–2825

Contents lists available at ScienceDirect

Journal of Chromatography B

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

Free fatty acid metabolic profile and biomarkers of isolated post-challenge diabetes and type 2 diabetes mellitus based on GC–MS and multivariate statistical analysis

Liyan Liu, Ying Li, Chunmei Guan, Kang Li, Cheng Wang, Rennan Feng, Changhao Sun *

Department of Nutrition and Food Hygiene, Public Health College, Harbin Medical University, 157 Baojian Road, Nangang District, Harbin 150086, PR China

ARTICLE INFO

Article history:
Received 13 May 2010
Accepted 24 August 2010
Available online 16 September 2010

Keywords:
Free fatty acid
Biomarker
Isolated post-challenge diabetes
Type 2 diabetes mellitus
Gas chromatography–mass spectrometry

ABSTRACT

Isolated post-challenge diabetes (IPD, 2h-PG ≥11.1 mmol/L and FPG <7.0 mmol/L) is often ignored in screening for diabetes by fasting plasma glucose (FPG) levels. The aim of this study was to investigate the metabolic profiles of serum free fatty acids (FFAs) and to identify biomarkers that can be used to distinguish patients with IPD from those with type 2 diabetes mellitus (T2DM) or healthy control individuals. FFA profiles of the subjects were investigated using gas chromatography–mass spectrometry (GC–MS). Principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA) were used for classification and prediction among the three groups. The predictive correct rates were 92.86% for IPD and healthy control individuals and 90.70% for T2DM and healthy control individuals, indicating that PLS-DA could satisfactorily distinguish IPD individuals from healthy controls and those with T2DM. Finally, palmitic acid, stearic acid, oleic acid, linoleic acid and α-linolenic acid were identified as potential biomarkers for distinguishing IPD from healthy control and T2DM individuals. These potential biomarkers might be helpful for diagnosis and characterization of diabetes.

1. Introduction

Currently, more than 10% of the global elderly population suffers from diabetes; of these individuals, 90% have type 2 diabetes mellitus (T2DM). T2DM can be diagnosed by fasting plasma glucose (FPG) ≥7.0 mmol/L or 2-h postprandial plasma glucose (2h-PG) ≥11.1 mmol/L using a 75 g oral glucose tolerance test (OGTT) [1,2]. However, in large-scale screening of diabetes or in routine physical examination, fasting plasma is usually collected to determine FPG for diabetes screening, because FPG is relatively convenient and easily obtained. For this reason, many patients with isolated post-challenge diabetes (IPD, 2h-PG ≥11.1 mmol/L and FPG <7.0 mmol/L) are often ignored. In the DECODE (Diabetes Epidemiology: Collaborative Analysis of Diagnostic Criteria in Europe) study, 50% patients with 2h-PG ≥11.1 mmol/L had a FPG <7.0 mmol/L [3]. In the Third National Health and Nutrition Examination Survey, 41% of previously undiagnosed T2DM patients aged 40–74 years had FPG <7.0 mmol/Land 2h-PG ≥11.1 mmol/L [4]. In addition, increasing evidence shows that IPD is associated with increased risk of cardiovascular disease and diabetic complications [5,6]. Therefore, it is important to identify biomarkers that are present in the fasting serum and that can be used to identify IPD subjects.

It is well known that diabetes is closely associated with metabolic lipid disorders, especially those involving free fatty acids (FFAs). Increased blood FFA levels have a central role in the development of diabetes, leading to insulin resistance (IR), impaired insulin signal pathways and destruction of β-cells [7–9]. However, different kinds of FFA have different or even opposite effects on the progress of IR and T2DM. For example, saturated fatty acids (SFA) worsened insulin sensitivity and increased the risk of T2DM, but polyunsaturated fatty acids (PUFA), particularly n-3 fatty acids, improved IR [10] and are potentially protective against T2DM [11]. On the other hand, plasma FFA levels are likely to increase with increases in blood glucose level and hepatic glucose production in patients with diabetes [12]. Previous studies have suggested that types and levels of FFAs have differing roles in the development of diabetes and the investigation of the relationship between specific types of FFAs and diabetes is more important than that of the total FFA level. Therefore, an accurate method is necessary to study the profile of FFAs and identify biomarkers that can distinguish healthy control, IPD and T2DM individuals.

Metabolomics, defined as the quantitative measurement of all low-molecular-weight metabolites in an organism at a specified time under specific environmental conditions [13], has been successfully applied to many fields such as disease diagnosis [14,15], biomarker screening [16,17], and nutrition research [18]. Over the past few years, metabolomics has combined data-rich advanced analytical techniques such as nuclear magnetic resonance (NMR)

* Corresponding author. Tel.: +86 451 8750 2801; fax: +86 451 8750 2885.
E-mail address: Sun2002changhao@yahoo.com (C. Sun).

1570-0232/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2010.08.035

Table 1
Clinical characteristics and dietary intake of 105 subjects (mean ± SD).

Parameter	Health control (n = 50)	IPD (n = 52)	T2DM (n = 53)
Age (years)	55.7 ± 10.4	54.7 ± 10.6	53.5 ± 10.5
Sex (female/male)	30/20	31/21	31/22
Smoker/non-smoker	17/34	16/36	18/35
Alcohol consumption (%)	37.25	40.38	39.54
Protein (g/d)	82.29 ± 26.15	83.04 ± 25.76	83.04 ± 26.08
Fat (g/d)	90.87 ± 32.24	91.86 ± 36.70	91.86 ± 36.70
Carbohydrate (g/d)	342.28 ± 129.26	341.48 ± 128.92	339.98 ± 127.02
BMI (kg/m^2)	22.1 ± 1.6	26.2 ± 3.8	25.8 ± 3.1
SBP (mmHg)	74.3 ± 6.7	81.3 ± 11.9	83.3 ± 11.7
DBP (mmHg)	113.1 ± 6.3	142.8 ± 24.9	142.1 ± 2.6
FPG (mmol/L)	4.2 ± 0.5	5.0 ± 0.6	10.8 ± 5.1
2h-PG (mmol/L)	4.6 ± 1.0	14.9 ± 3.9	19.1 ± 5.1
TG (mmol/L)	0.9 ± 0.3	2.1 ± 1.1	2.9 ± 3.5
TC (mmol/L)	4.2 ± 0.5	5.1 ± 1.0	5.01 ± 1.2

SBP: systolic blood pressure; DBP: diastolic blood pressure; BMI: body mass index; TG: triglycerides; TC: total cholesterol. FBG: fasting plasma glucose; 2h-PG: 2h-postprandial plasma glucose.

spectroscopy and mass spectrometry with multivariate statistical analysis. Gas chromatography–mass spectrometry (GC–MS) is a relatively low-cost alternative that provides high separation efficiency to resolve the complex biological mixtures and is a common analytical tool in metabolomics. This technique has been used in profiling of the metabolites of some diseases as well as serum fatty acids [19,20].

In the present study, we used GC–MS technology to investigate the comprehensive metabolic profile of human serum FFAs. Multivariate statistical analyses, including to principal component analysis (PCA) and the partial least squares-discrimination analysis (PLS-DA), were used to map the distribution of all human serum samples and identify healthy control individuals, IPD patients and T2DM patients. The results indicated that the predictive power, sensitivity and specificity of the PLS-DA model were better than those of the PCA model. Fatty acid metabolic profile analysis combined with PLS-DA could be useful for identifying potential biomarkers and diagnosis of IPD and T2DM.

2. Materials and methods

2.1. Subjects

Fifty-three patients with newly diagnosed T2DM and 52 patients with IPD were selected according to FPG and 2h-PG levels from a population-survey study to investigate the prevalence and risk factors of diabetes in Harbin. All subjects were recruited from communities of Harbin, Heilongjiang Province, in the North of China. T2DM patients were diagnosed according to the 1997 American Diabetes Association (ADA) criteria (Report of the expert committee on the diagnosis and classification of diabetes mellitus 1997). The cutoff value for FPG was 7.0 mmol/L and for 2h-PG was 11.1 mmol/L. The 50 healthy adults in the healthy control group were from the same population-survey study of patients and were not related to the patients.

There were no significant differences among the three groups in age, sex, smoking, alcohol consumption and dietary intake (Table 1). In addition, body mass index (BMI), triglyceride (TG) level, total cholesterol (TC) level, systolic blood pressure (SBP) and diastolic blood pressure (DBP) were not significantly different between IPD and T2DM individuals (Table 1). The study was approved by the Ethics Committee of Harbin Medical University and informed consent was obtained from each participant.

2.2. Chemicals and regents

Fatty acid standards were purchased from Sigma (St Louis, MO, USA, ≥99% purity): myristic acid (C14:0), palmitic acid (C16:0), palmitoleic acid (C16:1n-7), heptadecanoic acid (C17:0), stearic acid (C18:0), oleic acid (C18:1n-9), linoleic acid (C18:2n-6), linolenic acid (C18:3n-3), γ-linolenic acid (C18:3n-6), *cis*-11,14-eicosadienoic acid (C20:2n-6), arachidonic acid (C20:4n-6), *cis*-5,8,11,14,17-eicosapentaenoic acid (C20:5n-3), *cis*-7,10,13,16,19-docosapentaenoic acid (C22:5n-6), *cis*-4,7,10,13,16,19-Docos-ahexaenoic acid (C22:6n-3), tetracosanoic acid (C24:0), and selacholeic acid (C24:1n-9). The

Table 2
The equation and the correlation coefficient of fatty acids.

Fatty acid	Calibration equation	Regression coefficient	Range of quality (μg/mL)	Limit of detection (μg/mL)
C14:0	$Y = 0.4414X - 0.2206$	0.9964	2.96–25.00	0.05
C16:0	$Y = 0.0289X + 0.418$	0.9957	249.78–750.5	0.03
C16:1	$Y = 0.0425X - 0.0822$	0.9996	25.10–100.10	0.08
C18:0	$Y = 0.0248X - 0.0479$	0.9992	61.00–253.00	0.03
C18:1	$Y = 0.0319X - 0.4984$	0.9991	150.00–750.00	0.04
C18:2	$Y = 0.0331X - 0.2653$	0.9999	250.00–1000.00	0.05
γ-C18:3	$Y = 0.2107X - 0.4323$	0.9995	5.30–25.60	0.08
C18:3	$Y = 0.0689X - 0.0033$	0.9981	1.10–50.00	0.08
C20:2	$Y = 0.0483X - 0.0267$	0.9988	5.30–31.00	0.08
C20:4	$Y = 0.0269X - 0.034$	0.9997	50.00–300.00	0.05
C20:5	$Y = 0.0242X + 0.0328$	0.9953	4.90–50.00	0.05
C22:5	$Y = 0.0069X + 0.0051$	0.9975	2.50–56.00	0.06
C22:6	$Y = 0.0241X - 0.1348$	0.9993	20.00–100.00	0.07
C24:0	$Y = 0.0201X + 0.0483$	0.9976	1.20–15.50	0.08
C24:1	$Y = 0.0117X + 0.013$	0.9967	1.45–5.80	0.05

*X: the quality of fatty acids; Y: relative peak area = peak area of fatty acid/peak area of internal standard.

Table 3
The repeatability of the calibration samples ($n = 6$).

Fatty acids	The true concentration	Measured concentration	CV%	Bias%
C14:0	5.00	4.64 ± 0.24	5.17	92.80
C16:0	350.00	352.24 ± 19.26	5.47	100.61
C16:1	50.00	48.17 ± 1.69	3.51	96.34
C18:0	150.00	142.38 ± 8.29	5.82	94.90
C18:1	300.00	304.26 ± 17.36	5.71	101.42
C18:2	450.00	443.37 ± 34.28	7.73	98.53
γ-C18:3	15.00	14.02 ± 0.91	6.49	93.47
C18:3	10.00	9.34 ± 0.56	6.00	93.40
C20:2	15.00	14.13 ± 0.42	2.97	94.20
C20:4	150.00	147.67 ± 8.67	5.87	98.45
C20:5	5.00	5.19 ± 0.37	7.13	103.80
C22:5	10.00	9.07 ± 0.26	2.87	90.70
C22:6	50.00	48.05 ± 3.07	6.39	96.10
C24:0	5.00	4.17 ± 0.25	6.00	83.40
C24:1	5.00	4.09 ± 0.14	3.42	81.80

Values shown are mean ± SD for fatty acid concentration (μg/mL)

solution 10% H_2SO_4/CH_3OH was freshly prepared by diluting H_2SO_4 (purity: ≥98.0%) in chromatographic grade methanol. *n*-Hexane (chromatographic grade), ethyl acetate (analytical reagent) and sodium chloride (analytical reagent) were purchased from Tianjin Guangfu Chemical Reagent Co (Tianjin, China).

2.3. Preparation of standard solutions and method validation

Stock solutions of the 15 fatty acids and an internal standards (heptadecanoic acid) were prepared at 1000.00 μg/mL in methanol. Working solutions were prepared with methanol at concentrations of 1.10–1000.00 μg/mL (see Table 2). All standard solutions were stored at −20 °C until required.

Calibration samples were prepared by spiking with 15 different concentrations of fatty acids standards. Limit of detection (LOD) were defined as lowest concentrations with signal-to-noise (S/N) ratios of 10. Repeatability of calibration samples was expressed as coefficients of variation (CV%) and percentage biases (bias%), respectively (see Table 3).

2.4. Sample preparation

Fasting blood samples were immediately centrifuged at 3,000 × *g* for 10 min at room temperature and then stored at −80 °C until analysis. Samples were randomly selected for FFA extraction and GC–MS acquisition. Briefly, aliquots (200 μL) of serum were spiked with internal standard (I.S.) working solution (200 μL heptadecanoic acid C17:0 200 μg/mL), and 1 mL 0.05% H_2SO_4 was then added to deposit protein. The FFA was extracted using 3 mL ethyl acetate and shaking with a vortex mixer for 60 s, then centrifuged at 4,000 × *g* for 10 min at room temperature. The ethyl acetate phase was evaporated to dryness under N_2. Following the addition of 2 mL 10% H_2SO_4–CH_3OH and incubation in a 62 °C water bath for 2 h, 2 mL saturated sodium chloride and 2 mL hexane were sequentially added and mixed for 60 s to obtain the fatty acid methyl esters. Samples were evaporated to dryness under N_2 gas, and 100 μL hexane was added to each tube prior to analysis.

2.5. Gas chromatography–mass spectrometry

GC–MS analysis was performed using a TRACE gas chromatograph with a Polaris Q mass spectrometer (Thermo Finnigan, Austin, TX, USA). Helium was used as the carrier gas. A split injector (the split ratio being 1:10) at 230 °C was used to add the sample (1.0 μL) onto a J&W DB-WAX (30 m × 0.25 mm I.D., 0.25 μm film thickness) capillary column. Fatty acid methyl esters were separated at constant flow with the following oven program: (*a*) initially 50 °C for 2 min; (*b*) temperature was increased at a rate of 10 °C/min up to 200 °C; (*c*) maintained at 200 °C for 10 min; (*d*) increased at a rate of 10 °C/min up to 220 °C; (*e*) maintained at 220 °C for 15 min. The transfer line was maintained at 230 °C. The ion trap mass spectrometer was operated under electron bomb ionization (EI) mode. Mass spectra of m/z 30–450 were collected by full scan mode with 0.58 s/scan velocity. Solvent delay time was 5 min. The source temperature was 230 °C with the electron energy at 70 eV.

2.6. Statistical analysis

2.6.1. One-way ANOVA analysis

An internal standard method was used for quantitative analysis in this study and all of the peaks exceeding a signal-to-noise (S/N) of 10 were selected. The results were presented as mean ± standard deviation (SD). Data that were not normally distributed were logarithmically transformed to obtain normal distribution before analysis. Continuous variables were analyzed by one-way ANOVA with Tukey's test. $P < 0.05$ was considered statistically significant.

2.6.2. Principal component analysis (PCA)

The FFA data were analyzed by PCA (SIMCA-P, Umea, Sweden) to establish any 'groupings' with respect to three groups including healthy control, IPD and T2DM. A PCA model was constructed with all samples. The score plot of PC1 versus PC2 was used to examine separation or clusters for the three groups.

2.6.3. Partial least square-discriminant analysis (PLS-DA)

The PLS is described as the regression extension of PCA. Instead of describing the maximum variation in the measured data (*X*) for PCA, PLS attempts to derive latent variables and imitates principal components (PCs); this maximizes the co-variation between the exploratory variable (*X*) and the response variable (*Y*). PLS-DA that discriminates the known classes in a calibration set is a special form of PLS modeling that aims to identify the variables and directions in multivariate space. In PLS-DA, an indicator *Y* matrix of category variables contains as many columns as there are known classes in the calibration set—each class has a column in *Y*. In the current work, PLS-DA was used to generate models for distinguishing the three groups. The metabolites with the greatest variable importance in projection (VIP) values [21] (VIP > 1.0) in the model were regarded as potential biomarkers.

3. Results and discussion

3.1. Method validation

Method validation was conducted by evaluating the linearity, limit of detection (LOD), repeatability of the present method using spiked calibration samples prepared from 15 free fatty acids. The method had excellent linearity (correlation efficient *r*, 0.9953–0.9999). LODs ranged from 0.03 to 0.08 μg/mL (Table 2). Repeatability was determined by analyzing calibration samples. CV% ranged from 2.87% to 7.73%, whereas bias% ranged from 81.80% to 103.80% (Table 3).

3.2. GC–MS profiles of serum samples

The quantitative results of FFA analysis of the three groups are presented in Table 2. Fifteen types of FFA were detected in the serum of all samples, including four saturated fatty acids (C14:0, C16:0, C18:0 and C24:0), three monounsaturated fatty acids (C16:1n-7, C18:1n-9 and C24:1n-9) and eight polyunsaturated fatty acids (C18:2n-6, C18:3n-6, C18:3n-3,C20:2n-6, C20:4n-6,

Table 4
Quantitative analysis of FFA of the three groups.

FFA (μg/mL)	Control (n = 50)	IPD (n = 52)	T2DM (n = 53)
C14:0	3.61 ± 0.52	3.93 ± 0.64*	4.15 ± 0.80**
C16:0	356.76 ± 79.39	510.28 ± 143.57**	570.83 ± 174.39**,#
C16:1	40.38 ± 8.79	52.05 ± 18.57**	57.40 ± 18.80**
C18:0	87.64 ± 28.38	117.52 ± 42.98**	161.98 ± 80.39**,##
C18:1	268.96 ± 41.67	430.38 ± 128.98**	510.73 ± 220.91**,##
C18:2	455.57 ± 71.06	870.44 ± 185.47**	898.44 ± 208.18**
γ-C18:3	11.14 ± 0.67	11.96 ± 1.05**	12.89 ± 2.91**,##
C18:3	5.50 ± 2.40	13.83 ± 4.84**	19.71 ± 13.64**,##
C20:2	14.37 ± 4.34	15.21 ± 5.40	14.46 ± 6.85
C20:4	160.48 ± 37.10	201.11 ± 48.18**	252.24 ± 53.63**
C20:5	2.54 ± 2.18	4.78 ± 3.69	6.00 ± 4.01*
C22:5	11.59 ± 7.44	21.01 ± 17.18**	23.39 ± 14.65**
C22:6	51.64 ± 14.09	58.14 ± 15.25	61.59 ± 15.49**
C24:0	2.84 ± 2.56	4.49 ± 1.06**	4.14 ± 1.24**
C24.:1	1.84 ± 0.98	3.00 ± 1.97**	3.26 ± 1.83**
Total	1474.85 ± 157.90	2318.14 ± 488.38**	2601.22 ± 643.25**

* P < 0.05 compared with healthy control.
** P < 0.001 compared with healthy control.
\# P < 0.05 compared with IPD.
\#\# P < 0.001 compared with IPD.

C20:5n-3, C22:5n-6 and C22:6n-3). Most FFA concentrations differed significantly (P < 0.05) between healthy control and IPD individuals and between healthy control and T2DM individuals. There were significant differences in five FFAs between IPD and T2DM individuals. For the total FFA (the sum of the fifteen detected FFAs), healthy controls had lower total FFA levels than did IPD and T2DM individuals, and there was a significant difference in total FFA levels between IPD and T2DM individuals (P < 0.001). As the data showed different FFA profiles (Table 4), PCA and PLS-DA were used for discrimination of the three groups and identification of potential biomarkers.

3.3. Discrimination of the three groups by PCA

FFA profiles were analyzed by PCA. The score plot of PCA represented the distribution of all samples. In Fig. 1a, data of three samples were discarded because they were outside of the 95% confidence interval (the ellipse). The first two components of the model (PC1 and PC2) explained 57.4% of the variance (R^2 = 0.574, and Q^2 = 0.386). As shown in Fig. 1a, many samples in different groups were scattered and could not be completely separated. This is because PCA finds a lower dimensional space capturing the maximum amount of variance in an input data matrix, X, without losing any useful information. However, the excessive number of irrelevant variables could result in confusion during sample discrimination [22]. PLS is a similar approach to PCA except it reduces the dimension of both input and output data matrices, X and Y, by capturing the maximum amount of covariance between X and Y to best predict Y. Therefore, the discrimination analysis based on PLS would be expected to more powerfully separate the three groups than would PCA.

3.4. Discrimination of the three groups by PLS-DA

PLS-DA is a multivariate classification method based on PLS, the regression of PCA. PLS-DA explains maximum separation between defined class samples in the data set [23]. After crude screening by PCA, a PLS-DA model was constructed for classification of the three groups 16 type of fatty acids used of latent variable. The score plot of PLS-DA is shown in Fig. 1b. The PLS-DA model explained 62.2% of the variance (R^2 = 0.622, and Q^2 = 0.526); this was more than that explained by the PCA model. The samples in healthy control and other groups were separated clearly. However, owing to the poor recognition rate (50.46%) for 'all group' discrimination, three sub-comparisons were performed to compare healthy control and IPD individuals, healthy control and T2DM individuals, and IPD and T2DM individuals. The living form of mapping a spatial FFA profile was shown in three-dimensional (3D) data matrices (Fig. 2a–c). The results of PLS-DA showed that the samples in healthy control versus IPD and T2DM individuals (Fig. 2a and b) were separated clearly; there was also acceptable separation in the comparison of IPD and T2DM individuals (Fig. 2c). Furthermore, all the parameters of PLS-DA models are listed in Table 4. The values of R^2 and Q^2 in different models were more than 0.5, which indicates that these models were suitable for these recognition analysis. The permutation testing (999 times) [24] (Fig. 3, Table 5) was operated to test

Table 5
Parameters of PLS-DA models based on the date from sub-comparisons.

Groups	Component	R^2X	R^2Y	Q^2Y	R^2-intercept	Q^2-intercept
Healthy control-IPD	3	0.567	0.859	0.838	0.336	−0.43
Healthy control-T2DM	3	0.582	0.826	0.793	0.326	−0.397
IPD-T2DM	3	0.541	0.513	0.645	0.311	−0.367

Component: number of significant component calculated by seven times cross-validation.

Table 6
Classification of IPD, T2DM and health control by PLS-DA method.

	Recognition rate	Prediction rate	Sensitivity	Specificity	Correct rate
Healthy control-IPD	91.67%(55/60)	92.86%(39/42)	96.00%(48/50)	88.46%(46/52)	90.16%(94/102)
Healthy control-T2DM	90.00%(54/60)	90.70%(39/43)	92.00%(46/50)	88.67%(47/53)	90.29%(93/103)
IPD-T2DM	84.61%(55/65)	77.50%(31/40)	78.85%(41/52)	84.91%(45/53)	81.90%(86/105)

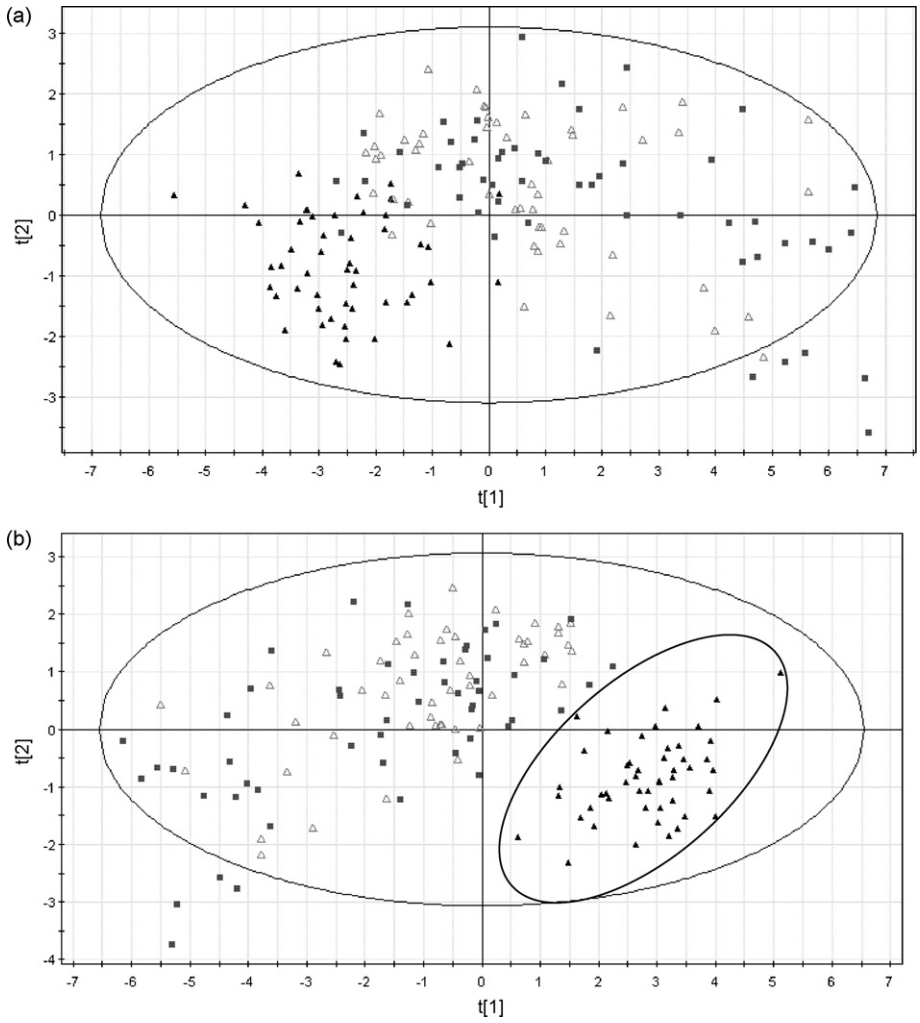

Fig. 1. Score plot from multivariate statistical analysis of the three groups: (a) PCA; (b) PLS-DA; health control (▲); IPD (△); T2DM (■).

the over-fitting of PLS-DA after modeling the data. The low value of intercepts, R^2 and Q^2, showed that the model is not over-fitted.

To determine the effect of different data preprocessing techniques on the predictive ability of the resulting modes, approximately two-thirds of the samples (the "training set") were selected randomly to construct a PLS-DA model that could then be used to predict the class membership of the remaining one-third of samples (the 'test set'). The classification results are shown in Table 6. The results showed relatively high predictive ability, with 92.86% and 90.70% prediction rates for healthy control versus IPD and T2DM individuals, respectively. The prediction rate for identification of IPD and T2DM individuals was 77.50%. All recognition and correct rates were greater than 81.90%. Furthermore, the specificity for prediction was more than 84.91% in all PLS-DA models. All these results indicate that discrimination models obtained by PLS-DA had good recognition and predictive abilities.

3.5. Potential biomarkers

In this study, a PLS-DA model was used for classification of healthy control, IPD and T2DM and the identification of potential biomarkers. The VIP [18] value was used to reflect the variable importance and identify potential biomarkers. The variables with VIP values more than 1.0 and with standard deviations less than the mean values were selected as possible biomarker candidates and are described in Table 7. It is interesting that we identified three types of FFA (C18:1, C18:2 and C18:3) as potential biomarkers for identification of healthy control versus IPD and T2DM individuals. Although total FFA is important for clinical diagnosis of diabetes, the VIP value of total FFA was lower (VIP value <0.9) than those of the three identified biomarkers. This indicates that the use of C18:1, C18:2 and C18:3 as index substances may be more accurate and sensitive than total FFA in the diagnosis of diabetes. Therefore, according with Table 4, the concentration ranges of FFAs biomarkers are proposed in this study (Table 8).

Table 7
Identification results of the potential biomarkers discovered by VIP values.

Groups	Biomarkers	VIP values	*P*-Values[a]
Healthy control-IPD	C18:2	1.47	$2.53E^{-22}$
	C18:3	1.39	$3.84E^{-18}$
	C18:1	1.10	$1.49E^{-8}$
Healthy control-T2DM	C18:2	1.29	$1.49E^{-7}$
	C18:3	1.21	$7.53E^{-13}$
	C18:1	1.08	$1.49E^{-7}$
IPD-T2DM	C18:2	2.03	0.181
	C18:0	1.39	0.0014
	C16:0	1.05	0.025

[a] *P*-Values were calculated by ANOVA analysis.

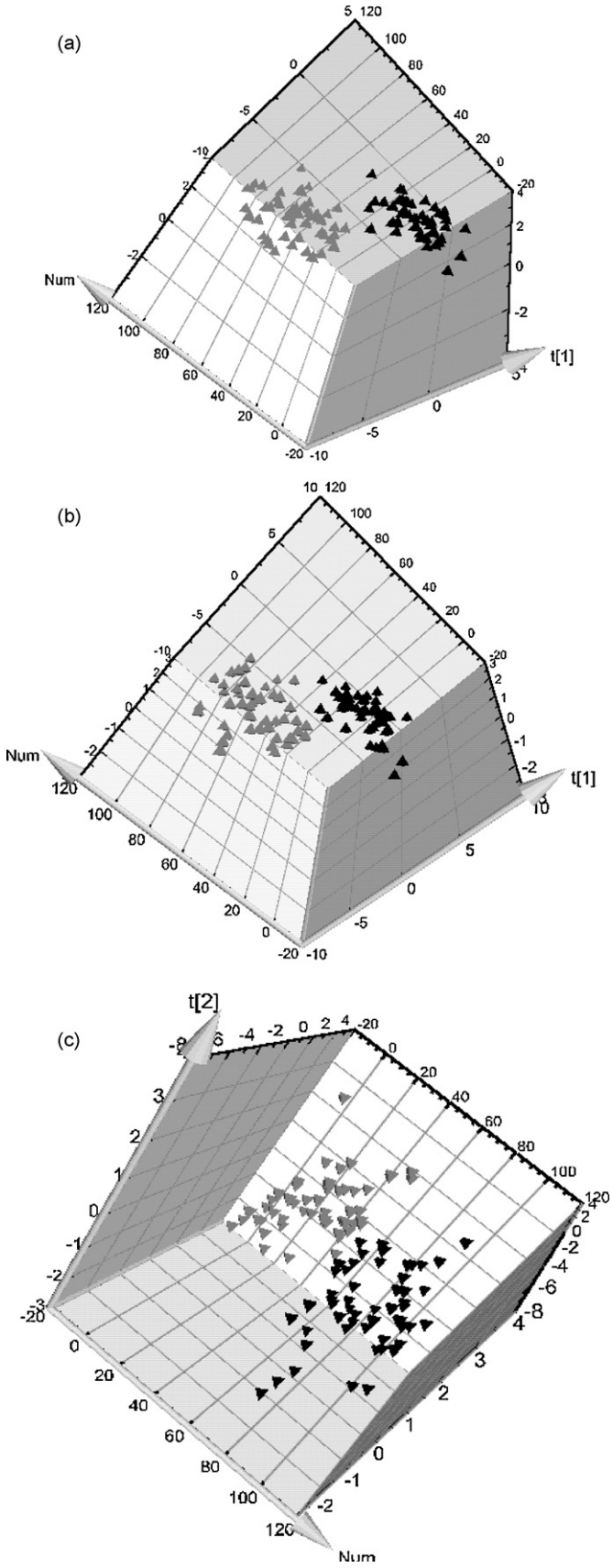

Fig. 2. 3D scores plots of free fatty acids from PLS-DA of models: (a and b) healthy control (black); IPD, T2DM (red); (c) IPD (black); T2DM (red).

Table 8
The concentration ranges of the FFA biomarkers in three groups.

Biomarkers	Healthy control	IPD	T2DM
C18:2	350.00–550.00	650.00–800.00	800.00–1250.00
C18:3	2.50–8.00	9.00–20.00	20.00–40.00
C18:1	200.00–300.00	400.00–550.00	550.00–800.00

For IPD versus T2DM, the VIP values of C18:2, C16:0 and C18:0 fatty acids were more than 1.0. However, there was no significant difference in the concentration of C18:2 between IPD and T2DM individuals, so C16:0 and C18:0 were selected as potential biomarkers. Moreover, the potential biomarkers can also be identified from the PLS-DA model loadings plot. The loadings plot (Fig. 4) identified the variables that strongly contribute to the separation of classes. In turn, these fatty acids might be biomarker candidates. From the loadings plot, we can also detect the potential biomarkers furthest from the origin. The potential biomarkers found in the loadings plot were in agreement with those identified by VIP values. Thus, patients who have higher levels of C18:2 are likely to be diagnosed with IPD or T2DM, and it is necessary to determinate C16:0 and C18:0 levels to help discriminate IPD and T2DM.

There have been increasing numbers of studies on the use of FFA metabolic profiles and biomarkers to distinguish T2DM patients from healthy control individuals. Lun-Zhao et al. suggested that FFAs might be suitable for classification of T2DM patients and healthy control individuals, and identified three FFAs as potential biomarkers for discrimination of T2DM patients and healthy control individuals [20]. Moreover, it has also been reported that the organic acids in urine, and glucose, phosphate and linoleic acid in plasma might be potential biomarkers to distinguish T2DM patients from healthy control individuals [25,26]. Although these studies have investigated the relationship between fatty acids and T2DM and identified some possible biomarkers that can be used to distinguish T2DM and healthy control individuals, none of them was involved in the differences of FFA profiles and distinguishing biomarkers of healthy control versus IPD patients. By contrast, in our study, the biomarkers identified will help us identify patients with diabetes from large populations, and be particularly useful for identification of IPD patients.

The potential biomarkers C16:0 (palmitic acid), C18:0 (stearic acid), C18:1 (oleic acid,n-9), C18:2 (linoleic acid, n-6) and C18:3 (α-linolenic acid, n-3) are all important bioactive molecules, and have effects on the function of pancreas and signal transporting in various cellular processes. For example, palmitic acid (C16:0), a biomarker that can distinguish IPD versus T2DM individuals, may inhibit the metabolic actions of insulin and attenuate insulin signal transduction [27]; whereas linoleic acid (C18:2,n-6), an important biomarker for distinguishing healthy control versus IPD and T2DM individuals, may preserve pancreatic β-cell function and improved peripheral glucose utilization [28]. These biomarkers (identified FFAs) may reflect the different features of a glucose metabolic disorder. Although the mechanisms by which FFAs affect pancreatic function are not yet clearly understood, it has recently been reported that long-chain FFAs amplify insulin secretion from pancreatic β-cells by activating GPR40, and that high levels of FFAs may affect β-cell function by over stimulating GPR40 [29,30]. This is one possible explanation for the effects of FFAs on diabetes, but it is not yet known whether the different ratios of various FFAs could influence GPR40 signaling. The effects of different ratios and interactions of these biomarkers should be further investigated *in vitro* and *in vivo*.

Taken together, these potential biomarkers reflect the deregulation of fatty acid metabolism in diabetic individuals, and might be beneficial in the diagnosis of or further pathogenesis research in diabetes.

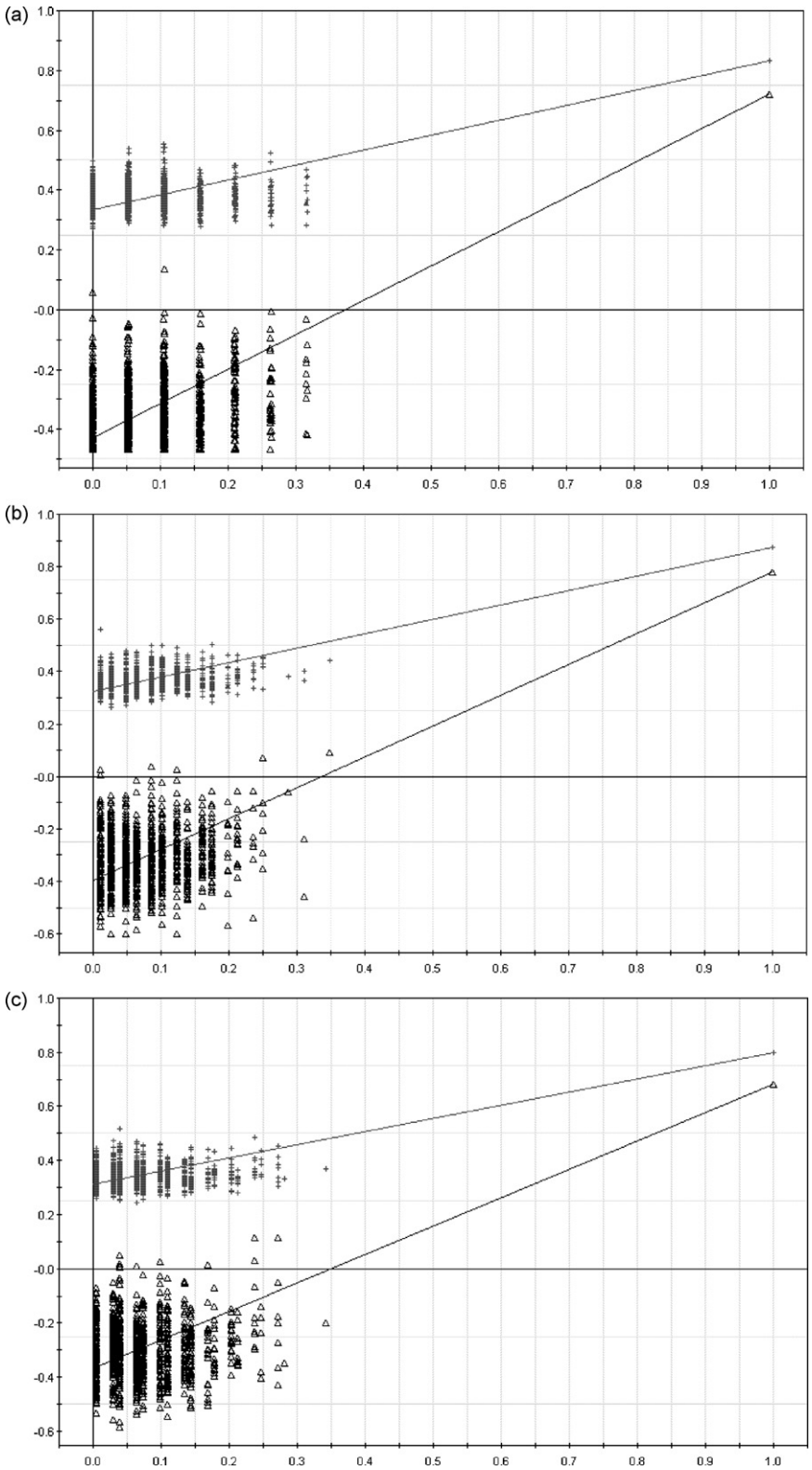

Fig. 3. Permutation testing results of the PLS-DA models, which was used to test the possibility of over-fitting: (+) R^2 and (△) Q^2; (a) healthy control and IPD; (b) healthy control and T2DM; (c) IPD and T2DM.

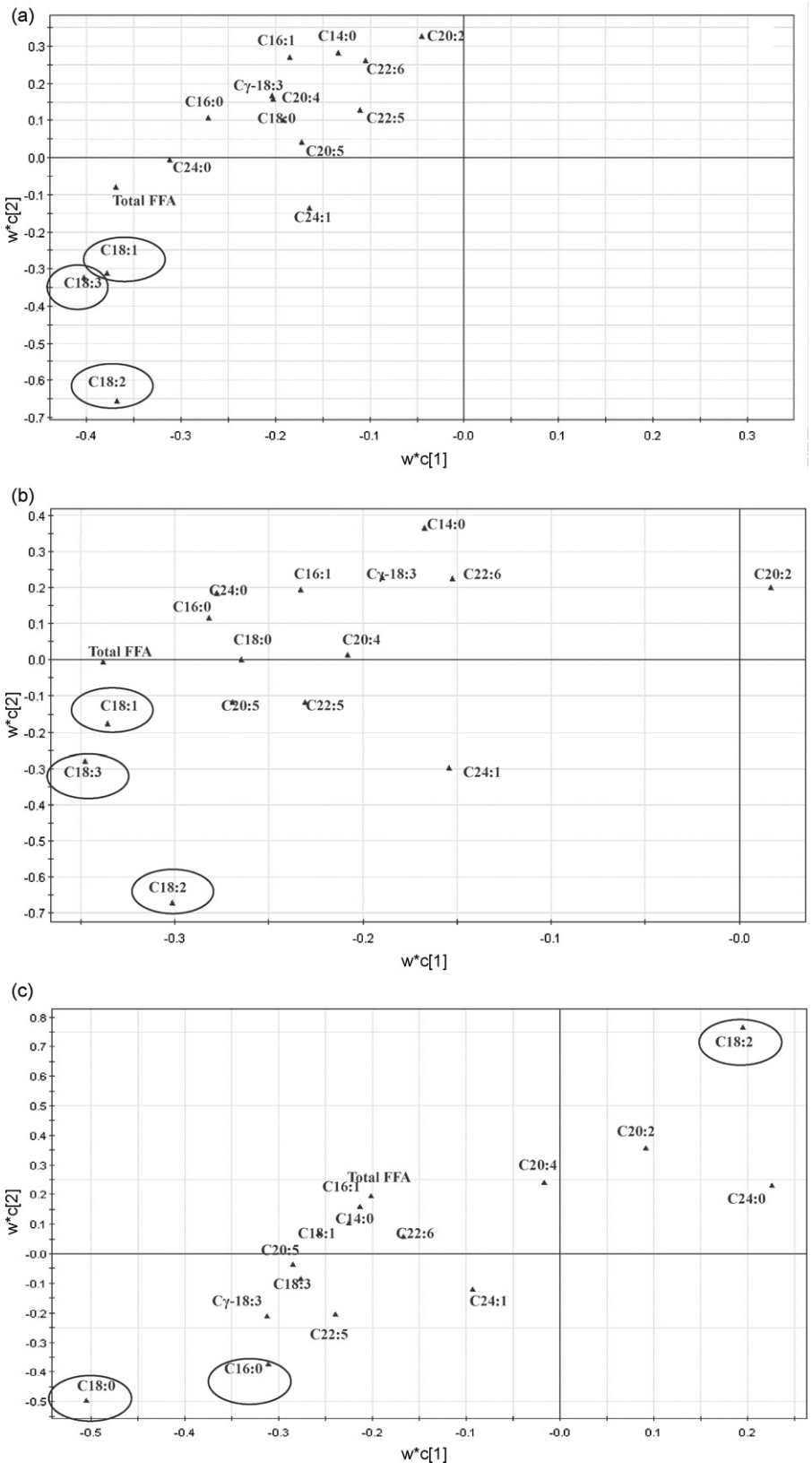

Fig. 4. Loadings plots of free fatty acids from PLS-DA models: (a) healthy control and IPD; (b) healthy control and T2DM; (c) IPD and T2DM. Corresponding loadings plot; possible biomarkers are marked with a circle.

4. Conclusions

Metabolomics is now recognized as an independently and widely used technique for identifying combination biomarkers for disease. IPD and T2DM are both associated with FFA metabolism disorders. In this study, metabolomics was used to investigate serum fatty acid metabolic profiles and biomarkers of healthy control, IPD and T2DM patients. We have demonstrated the application of GC–MS coupled with multivariable analysis in the classification of healthy control, IPD and T2DM patients. Moreover, we identified potential biomarkers that could be used to distinguish IPD patients from healthy controls, T2DM patients from healthy controls, and IPD from T2DM patients. Therefore, the detection of FFA metabolic profiles by an efficient multivariate statistical method might be useful in the diagnosis of IPD and T2DM patients, and identification of IPD individuals from large populations using fasting serum levels.

Acknowledgements

This work was supported by the Natural Science Foundation of China (No. 30671756) and China's 11th Five-Year Scientific and Technical Plan (No. 2006BAD27B05).

References

[1] K.G. Alberti, P.Z. Zimmet, Diabetes Med. 15 (1998) 539.
[2] R. Kahn, Diabetes Care 20 (1997) 1183.
[3] European Diabetes Epidemiology Group, Diabetes epidemiology: collaborative analysis of diagnostic criteria in Europe, Diabetologia 42 (1999) 647.
[4] H.E. Resnick, M.I. Harris, D.B. Brock, T.B. Harris, Diabetes Care 23 (2000) 176.
[5] E. Barrett-Connor, A. Ferrara, Diabetes Care 21 (1998) 1236.
[6] J.E. Shaw, A.M. Hodge, P. Chitson, P.Z. Zimmet, Diabetologia 42 (1999) 1050.
[7] R.N. Bergman, M. Ader, Trends Endocrin. Met. 11 (2000) 351.
[8] K.L. Wyne, Am. J. Med 115 (2003) 29.
[9] G. Boden, Proc. Assoc. Am. Phys. 3 (1999) 241.
[10] L.H. Storlien, A.B. Jenkins, D.J. Chishoim, Diabetes 40 (1991) 280.
[11] J. Tuomìlehto, J. Undstròm, J.G. Eriksson, N. Engl. J. Med. 344 (2001) 1343.
[12] G.M. Reaven, Y.D. Chen, Am. J. Med. 85 (1988) 106.
[13] A.A. Evans, G. Chen, E.A. Ross, F.M. Shen, W.Y. Lin, W.T. London, Cancer Epidemiol. Biomarkers Prev. 11 (2002) 369.
[14] W.M. Claudino, A. Quattrone, L. Biganzoli, M. Pestrin, I. Bertini, A. Di Leo, J. Clin. Oncol. 25 (2007) 2840.
[15] W.R. Wikoff, J.A. Gangoiti, B.A. Barshop, G. Siuzdak, Clin. Chem. 53 (2007) 2169.
[16] R. Xue, Z. Lin, C. Deng, L. Dong, T. Liu, J. Wang, X. Shen, Mass Spectrometr. 22 (2008) 3061.
[17] M. Bogdanov, W.R. Matson, L. Wang, T. Matson, R. Saunders-Pullman, S.S. Bressman, M. Flint Beal, Brain 131 (2008) 389.
[18] S.H. Zeisel, Am. J. Clin. Nutr. 86 (2007) 542.
[19] J. Yang, G.W. Xu, J. Chromatogr. B 813 (2004) 53.
[20] L.Z. Yi, J. He, FEBS Lett. 580 (2006) 6837.
[21] L. Eriksson, E. Johansson, N. Kettaneh-Wold, Multi- and Megavariate Data Analysis—Part 1: Basic Principles and Applications, 2001.
[22] C. Wang, H.W. Kong, Y.F. Guan, J. Yang, G.W. Xu, Anal. Chem. 77 (2005) 4108.
[23] L. Eriksson, E. Johansson, N. Kettanah-Wold, S. Wold, Umetrics AB, Malmo, Sweden, 1999.
[24] User's Guide to SIMCA-P, SIMCA-P+ version 11.0, Umetrics AB, Ume.
[25] K.L. Yuan, H.W. Kong, Y.F. Guan, G.W. Xu, J. Chromatogr. B 850 (2007) 236.
[26] X. Li, Z.L. Xu, X. Lu, X.H. Yang, G.W. Xu, Anal. Chim. Acta 633 (2009) 257.
[27] M.W. Ruddock, A. Stein, E. Landaker, C. Robert, J. Biochem. 144 (2008) 599.
[28] A. Noto, P. Zahradka, N.R. Ryz, N. Yurkova, X.P. Xie, C.G. Taylor, Metabolism 56 (2007) 142.
[29] Y. Itoh, Y.J. Kawamata, M. Harada, Nature 422 (2003) 173.
[30] A.R. Saltiel, C.R. Kahn, Nature 414 (2001) 799.