

## ORIGINAL ARTICLE

## Potential metabolite markers of schizophrenia

J Yang<sup>1,2,11</sup>, T Chen<sup>3,11</sup>, L Sun<sup>1,2</sup>, Z Zhao<sup>4,5</sup>, X Qi<sup>6</sup>, K Zhou<sup>1,2</sup>, Y Cao<sup>3</sup>, X Wang<sup>3</sup>, Y Qiu<sup>7</sup>, M Su<sup>7</sup>, A Zhao<sup>6</sup>, P Wang<sup>8</sup>, P Yang<sup>8</sup>, J Wu<sup>9</sup>, G Feng<sup>1,2</sup>, L He<sup>1,2,10</sup>, W Jia<sup>3</sup> and C Wan<sup>1,2</sup>

<sup>1</sup>Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Bio-X Center, Shanghai Jiao Tong University, Shanghai, China; <sup>2</sup>Institutes for Nutritional Sciences, Shanghai Institute of Biological Sciences, Chinese Academy of Sciences, Shanghai, China; <sup>3</sup>Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China; <sup>4</sup>Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN, USA; <sup>5</sup>Department of Psychiatry, Vanderbilt University School of Medicine, Nashville, TN, USA; <sup>6</sup>School of Pharmacy, Shanghai Jiao Tong University, Shanghai, China; <sup>7</sup>Department of Nutrition, University of North Carolina at Greensboro, North Carolina Research Campus, Kannapolis, NC, USA; <sup>8</sup>Wuhu No. 4 People's Hospital, Wuhu, China; <sup>9</sup>Instrumental Analysis Center, Shanghai Jiao Tong University, Shanghai, China and <sup>10</sup>Institutes of Biomedical Sciences, Fudan University, Shanghai, China

Schizophrenia is a severe mental disorder that affects 0.5–1% of the population worldwide. Current diagnostic methods are based on psychiatric interviews, which are subjective in nature. The lack of disease biomarkers to support objective laboratory tests has been a long-standing bottleneck in the clinical diagnosis and evaluation of schizophrenia. Here we report a global metabolic profiling study involving 112 schizophrenic patients and 110 healthy subjects, who were divided into a training set and a test set, designed to identify metabolite markers. A panel of serum markers consisting of glycerate, eicosenoic acid,  $\beta$ -hydroxybutyrate, pyruvate and cystine was identified as an effective diagnostic tool, achieving an area under the receiver operating characteristic curve (AUC) of 0.945 in the training samples (62 patients and 62 controls) and 0.895 in the test samples (50 patients and 48 controls). Furthermore, a composite panel by the addition of urine  $\beta$ -hydroxybutyrate to the serum panel achieved a more satisfactory accuracy, which reached an AUC of 1 in both the training set and the test set. Multiple fatty acids and ketone bodies were found significantly ( $P < 0.01$ ) elevated in both the serum and urine of patients, suggesting an upregulated fatty acid catabolism, presumably resulting from an insufficiency of glucose supply in the brains of schizophrenia patients.

*Molecular Psychiatry* (2013) 18, 67–78; doi:10.1038/mp.2011.131; published online 25 October 2011

**Keywords:** schizophrenia; biomarker; metabolomics; GC-TOFMS; NMR

## Introduction

Schizophrenia is a severe mental disorder affecting approximately 0.5–1% of the population worldwide.<sup>1</sup> As one of the most expensive medical illnesses, schizophrenia represents a serious burden on the health-care system.<sup>2</sup> Clinically, it has heterogeneous presentations, with positive and negative symptoms at different levels of prominence across time and across individuals.<sup>3</sup> The current diagnosis of schizophrenia remains subjective due to its complex

spectrum of symptoms, and the mechanism underlying the disease process has yet to be elucidated.

Recent clinical studies suggest that early intervention mitigates progression and improves therapeutic outcomes in the disease.<sup>4–6</sup> Establishment of biomarkers will enable early disease prevention, and thus improve prognosis. A niacin test proposed to measure reduced membrane arachidonic acid levels was suggested for the diagnosis of schizophrenia in 1980,<sup>7</sup> but has not found its way into clinical use so far owing to its low sensitivity and specificity.<sup>8–10</sup> Imaging modalities including functional magnetic resonance imaging, positron emission tomography and single-photon emission computerized tomography (SPECT), although frequently used in the pathological study of schizophrenia,<sup>11</sup> have not been able to offer a diagnostic solution with a high enough sensitivity and specificity. Recently, genetic variants at the major histocompatibility complex locus have been found significantly associated with this disease from several major genome-wide association studies,<sup>12–14</sup> suggesting an immune component in its

Correspondence: Dr W Jia, Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Center for Systems Biomedicine, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China or Dr C Wan, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Bio-X Center, Shanghai Jiao Tong University, 1954 Huashan Road, Shanghai 200030, China.  
E-mails: w\_jia@uncg.edu or clwan@sjtu.edu.cn

<sup>11</sup>These two authors contributed equally to this work.  
Received 23 January 2011; revised 15 August 2011; accepted 23 August 2011; published online 25 October 2011

pathogenesis. However, genetic research strategies, including linkage, association and genome-wide association studies, have yet to generate applicable genetic biomarkers for schizophrenia.

Global profiling strategies, such as unbiased proteomics or metabolomics with body fluids, including plasma/serum, urine or cerebrospinal fluid (CSF), may hold a significant potential for translating discriminating molecules into clinical biomarkers.<sup>15–17</sup> Metabonomics, defined as the quantitative measurement of the metabolite composition in a given biological compartment,<sup>18</sup> probes significant biochemical alterations and provides mechanistic insights into the pathophysiology of diseases.<sup>19–21</sup> In fact, metabonomic strategies have been widely used to characterize the human metabolic status in the field of central nervous system disorders, such as motor neuron diseases, Parkinson's disease, Huntington's disease, multiple sclerosis, schizophrenia and meningitis (for reviews see refs. 17, 22). Psychiatric disorders have been found to be associated with metabolic pathway disturbances,<sup>17,23</sup> which could well be reflected in metabonomic profiles. In 2004, metabonomic profiling of post-mortem brain tissues by the nuclear magnetic resonance (NMR) was used as supplement to genomics and proteomics, providing evidence for mitochondria dysfunction hypothesis of schizophrenia.<sup>24</sup> It was repeatedly reported that schizophrenia patients showed decreased levels of various polyunsaturated fatty acids and increased phospholipid breakdown in both central and peripheral tissue. As membrane phospholipid was proposed as the biochemical basis of neurodevelopment hypothesis of schizophrenia,<sup>25</sup> lipid metabolism was the focus of early metabonomic research. Bahn's group further reported significant alternations in levels of free fatty acids and phosphatidylcholine in gray and white matter of schizophrenia using ultra-performance liquid chromatography-mass spectrometry.<sup>26</sup> It deserves to be mentioned that these findings are best viewed as hypotheses generating rather than as definitive conclusions.

As brain tissues are rare and limited, CSF is an alternative as reflection of brain metabolic status, which can provide valuable biological specimens for the detection of altered biological molecules directly associated with neuropsychiatric disorders.<sup>27</sup> Holmes *et al.*<sup>5</sup> found an abnormal CSF biochemical profile in patients with schizophrenia as compared with normal controls, characterized by significantly altered levels of glucose, acetate, alanine and glutamine. They further found that these alterations in CSF of schizophrenia can be traced in initial prodromal state of psychosis, but not sufficient to predict clinical outcome.<sup>28</sup> However, CSF is only clinically accessible by means of a sophisticated procedure involving lumbar puncture; thus, peripheral blood samples are also considered for metabonomic study. Tsang *et al.*<sup>29</sup> analyzed the NMR spectra of plasma from a cohort of 21 pairs of monozygotic twins discordant for schizophrenia and 8 pairs of matched control twins.

Resonances from very-low-density lipoproteins, low-density lipoproteins and aromatic groups were found predominantly responsible for the differentiation of affected and unaffected twins from the normal controls. However, the affected twins were on anti-psychotic medication and it is not clear whether and how much the drugs contributed to the alterations in lipid profiles if any. Given the key role lipids play in metabolism, Kaddurah-Daouk *et al.*<sup>30</sup> used a targeted lipidomics platform to map the global plasma lipid profile of patients with schizophrenia. In their study, phosphatidylethanolamine and phosphatidylcholine were found decreased in schizophrenia patients at baseline compared with healthy controls, which supported the membrane phospholipid hypothesis. Thereafter, the plasma tryptophan and purine metabolites were investigated in schizophrenia by Yao's group using the metabolomics platform of high-pressure liquid chromatography coupled with electrochemical coulometric array detection. *N*-acetylserotonin was found increased in schizophrenia patients compared with healthy controls, and some metabolite interactions within the plasma tryptophan pathway were proposed to be altered and implicated in the pathogenesis of schizophrenia.<sup>31</sup> After identification of purine metabolites in the plasma, homeostatic imbalance of purine catabolism was found in schizophrenia patients,<sup>32</sup> which is consistent with the notion of free radical-mediated neurotoxicity in schizophrenia. In addition, guanine was found to discriminate schizophrenia patients from healthy controls with a moderate sensitivity of 0.68 and specificity of 0.67.

These studies exemplify metabonomic applications in schizophrenia research, providing valuable clues to understanding the mechanism of the condition. However, fewer of them were designed for the purpose of disease diagnosis. The lack of disease biomarkers to support objective laboratory tests still constitutes a bottleneck in the clinical diagnosis and evaluation of schizophrenia. To date, there has not been a comprehensive global evaluation of small-molecule metabolites using gas chromatograph time-of-flight mass spectrometer (GC-TOFMS) in the context of schizophrenia. This approach has comparative advantages in outstanding sensitivity as well as capability to identify metabolites over other analytical platforms. In this study, we enrolled a cohort of 112 schizophrenia patients and 110 normal healthy subjects as controls, utilizing a GC-TOFMS platform to explore potential diagnostic serum biomarkers. Urine samples were also analyzed using GC-TOFMS and <sup>1</sup>H-NMR to identify additional metabolite markers for schizophrenia.

## Materials and methods

### Subjects

We recruited 112 patients diagnosed as schizophrenic according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition criteria in Anhui

Province, China. Of these, 64 patients were on first-onset psychosis and 48 hospitalized for relapse after at least 1 month without any antipsychotic drugs. In all, 50 of the 112 patients withdrew from this project after enrolling. Serum and urine samples were collected from the patients at baseline before initiation of antipsychotic treatment. We excluded those patients who received any form of mood-stabilizing drugs in the recent 2 weeks before clinical samples were collected. Serum and urine samples were also collected from 110 normal controls in the same district. We excluded participants with metabolism disorders (such as type I or type II diabetes) and/or with heavy consumption of alcohol in both patients and controls.

The enrolled subjects were separated into a training set and a test set (Table 1). The training set included 62 patients whose serum were available after 4 weeks of antipsychotic treatment and 62 sex- and age-matched controls; the remaining patients and controls made up the test set. All samples were collected following the guidelines of the local ethics committee. Written informed consent was obtained from all participants. The overall sampling was finished within 1 year.

#### Biofluid specimen preparation

Venous blood and urine specimens were collected in polypropylene tubes in the early morning after overnight fasting. The fresh blood samples were placed at room temperature and clotted naturally after approximately 1 h. Serum were then obtained after centrifugation at 10 000 r.p.m. for 10 min. Serum and urine were stored at  $-80^{\circ}\text{C}$ .

#### GC-TOF spectral acquisition and data analysis

Serum and urine metabolites were analyzed using chemical derivatization following our previously published procedures with minor modifications.<sup>33</sup> A 100  $\mu\text{l}$  aliquot of serum sample was vortexed after being spiked with two internal standard solutions (10  $\mu\text{l}$  of L-2-chlorophenylalanine in water, 0.3  $\text{mg ml}^{-1}$ ; 10  $\mu\text{l}$  of heptadecanoic acid in methanol, 1  $\text{mg ml}^{-1}$ ). The mixed solution was extracted with

300  $\mu\text{l}$  of methanol/chloroform (3:1) and vortexed for 30 s. After storing for 10 min at  $-20^{\circ}\text{C}$ , the samples were centrifuged at 12 000 g for 10 min. A measure of 300  $\mu\text{l}$  of the supernatant was transferred to a glass sampling vial and vacuum-dried at room temperature; 50  $\mu\text{l}$  urine sample was extracted after being centrifuged at 12 000 g for 10 min, followed by the addition of 30 U urease. Urea was degraded for 15 min at  $37^{\circ}\text{C}$ . After the addition of two internal standard solutions as in urine, the mixture was extracted with 170  $\mu\text{l}$  methanol. The samples were then centrifuged at 12 000 g for 5 min. A measure of 200  $\mu\text{l}$  of the supernatant was transferred to a glass sample vial and vacuum-dried at room temperature. The serum and urine residues were derivatized using a two-step procedure, that is, 80  $\mu\text{l}$  of methoxyamine (15  $\text{mg ml}^{-1}$  in pyridine) was added to the vial and kept at  $30^{\circ}\text{C}$  for 90 min, and then 80  $\mu\text{l}$  of BSTFA (*N,O*-bis(trimethylsilyl) trifluoroacetamide) (1% TMCS (trimethylchlorosilane)) was added and heated to  $70^{\circ}\text{C}$  for 60 min. A measure of 1  $\mu\text{l}$  aliquot of the derivatized solution was injected in splitless mode into an Agilent 6890N gas chromatograph coupled with a Pegasus HT TOFMS (Leco Corporation, St Joseph, MI, USA). Electron impact ionization (70 eV) at full scan mode ( $m/z$  30–600) was used, with an acquisition rate of 20 spectra per s in the TOFMS setting.

The MS files acquired from GC-TOFMS analysis were exported in NetCDF format using the Chroma-TOF software (v.3.30; Leco Co., Oakland, CA, USA). CDF (channel definition format) files were extracted using custom scripts (revised Matlab toolbox hierarchical multivariate curve resolution, developed by Par Jonsson *et al.*<sup>34,35</sup>) using MATLAB 7.0 (The MathWorks Inc., Natick, MA, USA) for data pretreatment procedures such as baseline correction, denoising, smoothing, alignment, time-window splitting and multivariate curve resolution (based on a multivariate curve resolution algorithm). The resulting three-dimensional data set included sample information, peak retention time and peak intensities. Then, the internal standards and any other known artificial peaks, such as peaks caused by noise, column bleed and BSTFA derivatization procedure, were removed

**Table 1** Demographic details of patients involved in GC-TOF analysis

Group	N	Male%	Age	Height (cm)	Weight (kg)	BMI ( $\text{kg m}^{-2}$ )	N Smoker (cigarettes)
<i>Training set<sup>a</sup></i>							
Control	62	40.3	36.9 $\pm$ 9.3	164.9 $\pm$ 6.8	60.8 $\pm$ 8.6	22.0 $\pm$ 4.8	15 (6 $\pm$ 3)
SZ-BL	62	40.3	36.9 $\pm$ 11.9	161.9 $\pm$ 7.9	56.3 $\pm$ 8.7	21.5 $\pm$ 2.6	12 (16 $\pm$ 8)
<i>Test set</i>							
Control	48	0	28.3 $\pm$ 8.6	160.1 $\pm$ 4.6	53.2 $\pm$ 6.1	20.7 $\pm$ 2.0	0
SZ-BL	50	42	36.9 $\pm$ 12.6	161.2 $\pm$ 6.2	57.0 $\pm$ 7.9	21.8 $\pm$ 2.3	6 (16 $\pm$ 7)

Abbreviations: BMI, body mass index; GC-TOF, gas chromatograph time of flight; SZ-BL, schizophrenia patients at baseline. <sup>a</sup>Training set consisted of patients and controls matched for gender, age, height, weight and BMI. These factors were not considered in the test set. Smoking habit is noted with the number of smokers (N smoker) and daily consumption of cigarettes (within the parenthesis, average  $\pm$  s.d.).

from the data set. For multivariate statistical analysis, the resulting data were mean centered and unit variance scaled during chemometric data analysis using the SIMCA-P 11.5 software package (Umetrics, Umeå, Sweden). Principal component analysis, partial least squares-discrimination analysis and orthogonal projection to latent structures (OPLS) were carried out for group discrimination. The following statistics, based on the models, are discussed in detail throughout this article.  $R^2X$  is the cumulative modeled variation in the  $X$  matrix;  $R^2Y$  is the cumulative modeled variation in the  $Y$  matrix; and  $Q^2$  is the cumulative predicted variation in the  $Y$  variable or matrix, based on sevenfold cross-validation. The range of these parameters is 0–1, where 1 indicates a perfect fit. Random permutation of the  $Y$  observations, while keeping the  $X$  matrix intact, was used as a measurement of data overfitting.<sup>36</sup> The default seven-round cross-validation in the SIMCA-P software package was applied with one-seventh of the samples being excluded from the mathematical model in each round, in order to guard against overfitting. A completely separate external test set was used as a prediction set to assess the accuracy of the OPLS model. Based on the variable importance on a projection (VIP) with a threshold of 1.5 from an appropriate sevenfold cross-validated OPLS model, a number of variables were obtained responsible for the difference in the metabolic profiles between different groups. These variables were also validated at the univariate level using the non-parametric Wilcoxon–Mann–Whitney test with the critical  $P$ -value set at 0.05. A false discovery rate (FDR)<sup>37</sup> control was implemented to correct for multiple comparisons. The  $q$ -value in the FDR control was defined as the FDR analog of the  $P$ -value. The  $q$ -value of an individual hypothesis test is the minimum FDR at which the test may be called significant. In this study, the  $q$ -value was set at 0.2.

In addition, compound identification was performed by comparing the mass fragments with NIST 05 Standard mass spectral databases in the NIST MS search 2.0 (NIST, Gaithersburg, MD, USA) software. This revealed a similarity of more than 70%, which was then verified by available reference compounds. Logistic regression was fitted to find a classifier composite of differential metabolites between schizophrenic patients and normal controls. Areas under the receiver operating characteristic (ROC) curve (AUC) were calculated to evaluate the performance of this classifier. Non-parametric Wilcoxon–Mann–Whitney test, FDR control, logistic regression and ROC curves were carried out on the freely available software R (<http://www.r-project.org>) with corresponding packages.

#### *NMR spectra of urine and data analysis*

An aliquot of 600  $\mu$ l urine was thawed at room temperature. After centrifugation at 12 000  $g$  for 5 min, 470  $\mu$ l supernatant urine was mixed with 60  $\mu$ l  $D_2O$  from Merck (Darmstadt, Germany) 5  $\mu$ l trisodium

phosphate (0.4 mM) and 115  $\mu$ l phosphate buffer (0.3 M, pH 7.4). The pH of the mixtures was adjusted to pH 7–7.4 with NaOH (1 M) or HCl (1 M) and then centrifuged again at 12 000  $g$  for 5 min. A measure of 600  $\mu$ l of the supernatant was pipetted into a 5-mm NMR tube (Bruker Biospin, Karlsruhe, Germany).

All samples were run on a 400 MHz Bruker Avance  $\beta$  spectrometer.  $^1H$ -NMR spectra were acquired using standard 1D water presaturation pulse sequence of the form  $d1-\pi/2-acq.$ , where  $\pi/2$  represents a 90° hard pulse and  $d1$  is a relaxation delay (5 s). The resonance of  $H_2O$  ( $\delta \approx 4.7$  p.p.m.) was selectively irradiated with a continuous-wave low-power pulse during  $d1$ . For each sample, 32 transients were collected into 16K data points, with a spectral width of 12 p.p.m. Before Fourier transformation, the free-induction decay was multiplied by an exponential function corresponding to a line-broadening factor of 0.3 Hz in the frequency domain. All spectra were phase and baseline corrected and referenced to TSP ( $\delta = 0$  p.p.m.) manually on TOPSPIN 2.1.

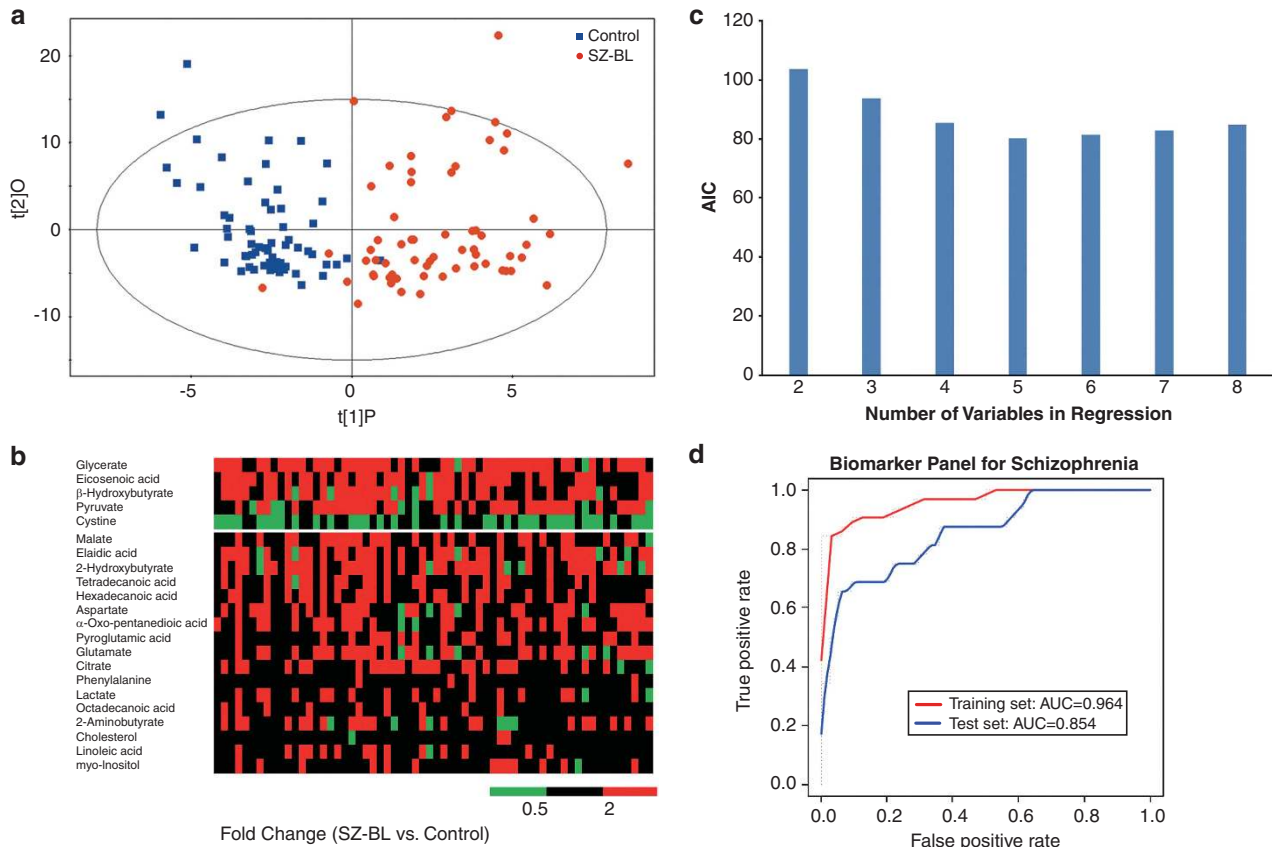
Regions from  $\delta = 0$ –0.6 p.p.m. (TSP) to  $\delta = 4.5$ –6.7 p.p.m. (water and urea) were excluded before analysis. Using AMIX 3.8.6, each NMR spectrum was reduced to integrated segments with a width of 0.04 p.p.m., giving a total of 181 integrated regions per spectrum. The spectrometric data were then converted into Microsoft Excel format. Univariate and multivariate statistical analysis were conducted as described above in the section of ‘GC-TOF spectral acquisition and data analysis’.

## Results

### *Metabolic profiles of schizophrenia patients*

*Serum metabolic profile.* After excluding internal standards, a total of 330 individual peaks were consistently detected in about 90% of the serum samples, which are used in the next analysis. For the training set, principal component analysis scores plot revealed little clustering trends between patients and controls, while partial least squares-discrimination analysis scores plot using three components showed difference between schizophrenia patients at baseline (SZ-BL) and normal controls ( $R^2Y = 0.842$ ,  $Q^2 = 0.578$ ). With  $Q^2$  intercepting the  $Y$  axis at  $-0.214$  in the 999 random permutations test, the supervised model was considered well guarded against overfitting (Supplementary Figure 1). To specify metabolic variations relevant to schizophrenia, a cross-validated OPLS model was constructed with satisfactory predictive ability using one predictive component and one orthogonal component in the training set ( $R^2Y = 0.716$ ,  $Q^2 = 0.472$ ).

Given the success of the OPLS model in classifying patients and controls (Figure 1a), a total of 39 paired retention time mass to charge ratio (RT- $m/z$ ) variables stood out relative to the VIP threshold (VIP > 1.5). In all, 22 differential metabolites were annotated using the available library, among which 20 metabolites



**Figure 1** Serum biomarker panel for schizophrenia. **(a)** The orthogonal projection to latent structures (OPLS) model was fitted to discriminate the schizophrenia patients at baseline (SZ-BL) and normal controls. **(b)** In all, 22 metabolites were identified with variable importance on a projection (VIP) > 1.5. **(c)** Logistic regression models were fitted with different numbers of metabolites. Akaike information criterion (AIC) of each model was shown here. The model built with five metabolites (glycerate, eicosenoic acid,  $\beta$ -hydroxybutyrate, pyruvate and cystine, upper part of **b**) had the highest predictive power. **(d)** Receiver operating characteristic (ROC) curves were drawn for the training and test sets. The area under the curve was 0.945 (95% confidence interval: 0.900–0.91) in the training set and 0.895 (95% confidence interval: 0.829–0.961) in the test set, indicating a ‘good’ clinical diagnosis efficiency for this set of biomarker metabolites.

being validated using reference standards (Table 2). In the 22 metabolites contributing most to the discrimination of schizophrenia patients from normal controls, five middle- and long-chain fatty acids (tetradecanoic acid, hexadecanoic acid, octadecanoic acid, oleic acid and eicosenoic acid) were found significantly elevated in patients.  $\beta$ -Hydroxybutyrate, a by-product of fatty acid metabolism, was found elevated with a fold change of 2.6. Cystine, a conjugation of two cysteines, was decreased in patients compared with normal controls. Amino acids, including glutamate, aspartate, serine and phenylalanine, were increased in patients compared with normal controls.

**Urine metabolic profile.** In the training set (Table 1), urine samples were available from 51 SZ-BL and matched controls. A total of 345 variables were used for statistical analysis using the same protocols as in the serum profile analysis. The cross-validated OPLS model was constructed using one predictive component and two orthogonal component

( $R^2 Y=0.578$ ,  $Q^2=0.352$ ), achieving a distinct separation between the metabolite profiles of the two groups (Figure 2a). Metabolites contributing most to the discrimination were screened with a VIP cutoff of 1.5. Short- and middle-chain fatty acids were found elevated in SZ-BL urines (Table 3).  $\beta$ -Hydroxybutyrate was found increased in SZ-BL urine compared with normal controls, which is consistent to that in SZ-BL serum. Cystine was found elevated in SZ-BL urine, which presents an opposite trend with that in SZ-BL serum.

We also employed the  $^1\text{H-NMR}$  platform to analyze the urine samples from 41 schizophrenia patients and their matched controls in the training set. Glucose was found elevated in SZ-BL by both analytical platforms (Table 4), although not significantly different between SZ-BL and controls in GC-TOFMS data. Chemical shift of 3-hydroxybutyrate was found increased in urine of schizophrenia patients, which is in accordance with the GC-TOFMS results. Other ketone bodies, acetone and acetoacetate, which were not detected by GC-TOFMS, were found elevated by NMR.

**Table 2** Differential serum metabolites between schizophrenia patients and normal controls

Pathway	Compound	VIP <sup>a</sup>	P-value <sup>b</sup>	FDR <sup>c</sup>	FC <sup>d</sup>
Fatty acids metabolism	Glycerate <sup>e</sup>	3.10	7.90E-10	6.06E-04	2.57
	Tetradecanoic acid <sup>e</sup>	2.77	2.90E06	6.06E-03	1.45
	Hexadecanoic acid <sup>e</sup>	2.61	4.10E-04	1.76E02	1.40
	Linoleate <sup>e</sup>	1.63	3.67E02	3.88E02	1.18
	Oleic acid <sup>e</sup>	3.01	1.36E-05	1.09E-02	2.09
	Octadecanoic acid <sup>e</sup>	1.78	7.39E-03	2.55E-02	1.14
	Eicosenoic acid <sup>f</sup>	3.26	8.56E-08	3.03E-03	1.96
	3-Hydroxybutyrate <sup>e</sup>	2.21	5.33E-05	1.27E-02	2.61
Carbohydrates metabolism	Pyruvate <sup>e</sup>	2.46	1.24E-05	1.03E-02	1.88
	Lactate <sup>e</sup>	1.94	2.71E-03	2.18E-02	1.24
	Citrate <sup>e</sup>	2.34	7.46E-06	7.88E-03	1.45
	2-Oxoglutarate <sup>e</sup>	2.57	1.82E-06	5.45E-03	1.59
	Malate <sup>e</sup>	3.07	2.21E-08	1.82E-03	1.57
Amino-acid metabolism	Glutamate <sup>e</sup>	2.35	4.26E-05	1.21E-02	1.63
	Aspartate <sup>e</sup>	2.60	3.44E-06	7.27E-03	1.38
	5-Oxoproline <sup>f</sup>	2.52	3.68E-07	3.64E-03	1.35
	Serine <sup>e</sup>	1.69	7.07E03	2.48E-02	1.13
	Phenylalanine <sup>e</sup>	1.96	4.26E-04	1.88E-02	1.14
	Cystine <sup>e</sup>	2.99	1.06E06	4.24E-03	-1.36
	2-Aminobutyrate <sup>e</sup>	1.71	2.44E-02	3.45E-02	1.28
	2-Hydroxybutyrate <sup>e</sup>	2.91	7.99E-08	2.42E-03	2.45
Inositol phosphate metabolism	myo-Inositol <sup>e</sup>	1.57	2.42E-02	3.39E-02	1.16

Abbreviations: FDR, false discovery rate; GC-TOF, gas chromatograph time of flight; OPLS, orthogonal projection to latent structures; SZ-BL, schizophrenia patients at baseline.

For metabolites from serum analyzed by GC-TOF.

<sup>a</sup>VIP shows variable importance in the projection obtained from the OPLS model with a cutoff of 1.5.

<sup>b</sup>P-values from non-parametric Wilcoxon–Mann–Whitney test.

<sup>c</sup>FDR was calculated with a *q*-value of 0.2. Paired *t*-test was used to investigate the difference between SZ-BL and SZ-4w.

<sup>d</sup>FC is a positive value fold change that indicates a relatively higher concentration present in SZ-BL (SZ-4w), while a negative value means a relatively lower concentration as compared to the controls (SZ-BL).

<sup>e</sup>Metabolites were verified by reference compounds.

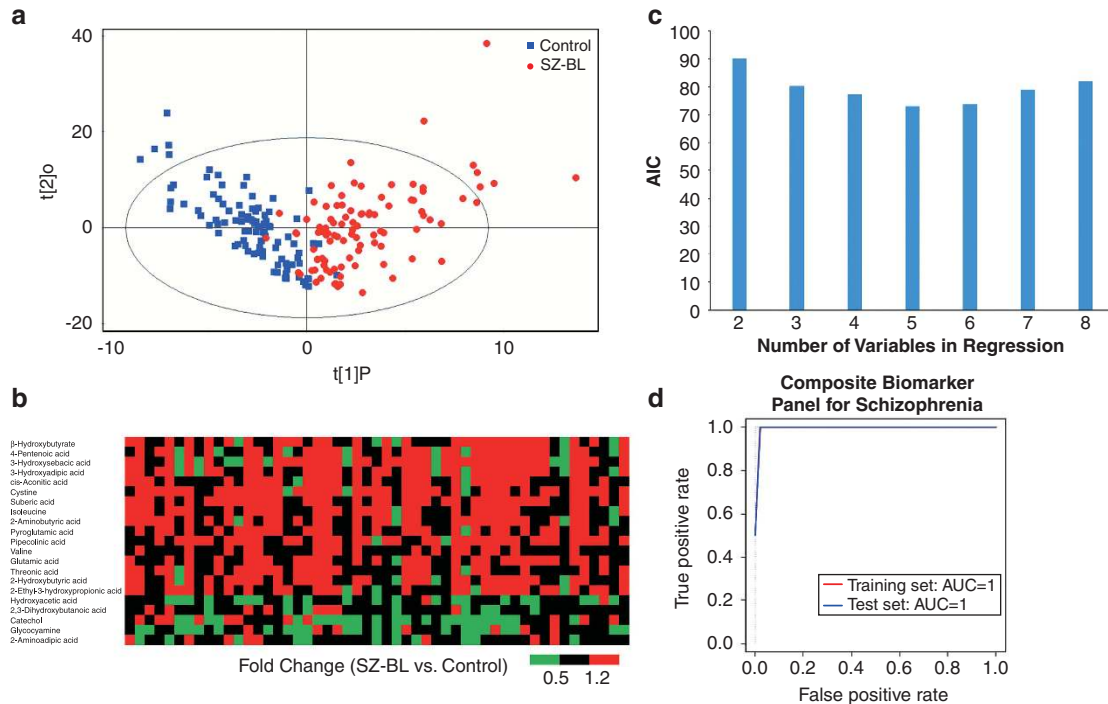
<sup>f</sup>Metabolites were identified using available library databases.

### Biomarker panel for diagnosis of schizophrenia

In the serum training set (Table 1), the sensitivity and specificity of the OPLS classifier model reached 95.2% and 98.4%, respectively (Supplementary Figure 2). The prediction package in the SIMCA-P software was applied to predict the test set with the OPLS model. The supervised variable *Y* of controls was designated 1 and that of schizophrenia patients 2. The *Y*-predicted scatter plot assigned samples to either the control or the schizophrenic group using an *a priori* cutoff of 1.5, that is, individuals with *Y*Pred (*Y* predicted) > 1.5 were considered as schizophrenic patients and those with < 1.5 as normal. A total of 41 out of 50 patients in the test set were correctly classified, reaching a satisfactory sensitivity of 82.0% and specificity of 89.6% (Supplementary Figure 2). These results suggest that this OPLS model constructed by the serum metabolites provides great potential for molecular diagnosis of schizophrenia.

ROC analysis is a classical methodology from signal detection theory and is now commonly used in clinical research.<sup>38</sup> Diagnosis based on quantification of fewer metabolites will be more convenient and

economical if they can provide sufficient information. To explore a simplified set of schizophrenia biomarkers, metabolites in the serum with higher fold change and VIP value were used as candidates (Table 2 and Figure 1b), although the pre-selecting marker candidates may result in positively biased cross-validation estimates. We calculated the AUC for each candidate as a classifier and sorted the AUC in descending order. Logistic regression was then fitted from first 2 to 8 variables (Figure 1c). According to the Akaike information criterion (AIC), the logistic regression with five variables explained most deviations between SZ-BL and normal controls using the smallest numbers of variables. In particular, the models with 6–8 variables disclosed that the five-variable model had the highest predictive power when we performed a stepwise, backward-elimination mode. These five metabolites were glycerate, eicosenoic acid,  $\beta$ -hydroxybutyrate, pyruvate and cystine. ROC curves of the serum panel for the training and test set were drawn with the odds ratio (Figure 1d). The AUC of this biomarker panel was 0.945 (95% confidence interval: 0.900–0.991) in the



**Figure 2** Composite biomarker panel for schizophrenia. (a) Scores plot of orthogonal projection to latent structures (OPLS) discriminating urine profiles of schizophrenia patients at baseline (SZ-BL) and normal controls. (b) In all, 21 metabolites identified from the OPLS model with variable importance on a projection (VIP) > 1.5. (c) Logistic regression models fitted with different numbers of metabolites. The smallest Akaike information criterion (AIC) was of the model with the five serum metabolites (glycerate, eicosenoic acid,  $\beta$ -hydroxybutyrate, pyruvate and cystine), the same as those in serum biomarker panel. The logistic regression with the addition of urine  $\beta$ -hydroxybutyrate to the serum panel had a similar AIC value to perfectly predict classification of schizophrenia or control. Thus, the urine  $\beta$ -hydroxybutyrate and the serum biomarker consisted a so-called ‘composite biomarker panel for schizophrenia’. (d) ROC curves of the composite biomarker panel in the training and test sets.

training set and 0.895 (95% confidence interval: 0.829–0.961) in the test set. According to the traditional academic points system, an area of 1 represents a perfect test, while an area of 0.5 represents a worthless test. Thus, we obtained a ‘good’ classifier of schizophrenia patients and normal controls.

The differential metabolites in urine, which reflected early changes in disease, could also supplement the serum metabolite markers. We calculated the AUCs of logistic regression models composed of the five metabolites in the serum panel added by each metabolite in Table 3. Most of the newly constructed panels have increased AUCs than the original serum panel (Supplementary Table 1). Of note, the composite panel by addition of a urinary metabolite, 3-hydroxybutyrate, to the serum panel achieved an AUC of 1.00 in both the training set and test set, meaning an equivalent diagnosis compared with the clinical interviews (Figure 2d).

#### Similar metabolic profile of first-episode and recurrent schizophrenia patients

Among the 112 schizophrenia patients, 64 were diagnosed on first episode and the others were recurrent. Serum metabolic profile of the first episode was found similar to that of the recurrent, as

multivariable partial least squares-discrimination analysis or OPLS models were not constructed with reliable model parameters, indicating that the serum metabolic profile of the first-episode subjects is not distinct from that of the recurrent subjects. Metabolites were either with  $P$ -values > 0.05 from Wilcoxon–Mann–Whitney test or screened out by FDR correction with  $q$  set at 0.2 in both the serum and urine data. As a result, the similar metabolic profiles of first-episode and recurrent patients were not able to generate identifiable metabolite markers to differentiate first-episode patients from those with recurrent disease.

## Discussion

### Potential diagnostic panel for schizophrenia

A series of research in genetics, proteomics as well as imaging have been designated to the discovery of markers in schizophrenia. However, it is still a challenge to gain a marker that could replace or even aid diagnosis of this severe disease. In this study, we acquired GC-TOFMS spectra of serum and urine metabolites from 112 schizophrenia patients and 110 healthy subjects. Within the identified differential metabolites, a serum biomarker panel consisting

**Table 3** Differential urine metabolites in schizophrenia patients

Pathway	Compound	VIP <sup>a</sup>	P-value <sup>b</sup>	FDR <sup>c</sup>	FC <sup>d</sup>
Fatty acids metabolism	Suberic acid <sup>e</sup>	2.05	1.44E-04	3.74E-03	1.59
	3-Hydroxysebacic acid <sup>e</sup>	1.59	3.53E05	1.87E03	5.55
	3-Hydroxyadipic acid <sup>e</sup>	2.35	3.34E04	4.36E-03	2.06
	2-Ethyl-3-hydroxypropionic acid <sup>e</sup>	1.53	5.65E02	5.30E-02	1.29
	4-Pentenoic acid <sup>e</sup>	2.14	6.07E-03	2.18E-02	1.54
	Threonic acid <sup>f</sup>	1.64	5.82E02	5.36E02	1.21
	2,3-Dihydroxybutanoic acid <sup>e</sup>	1.58	6.19E-03	2.24E-02	-1.3
	2-Hydroxybutyric acid <sup>f</sup>	1.55	3.23E-03	1.81E-02	1.41
	3-Hydroxybutyric acid <sup>f</sup>	1.6	5.83E-03	2.06E-02	1.37
	Hydroxyacetic acid <sup>f</sup>	2.09	1.90E-03	1.56E-02	-1.36
Carbohydrates metabolism	<i>cis</i> -Aconitic acid <sup>f</sup>	1.99	1.23E-02	3.30E-02	1.28
Amino-acid metabolism	Cystine <sup>f</sup>	1.55	1.20E-03	9.97E-03	1.54
	Valine <sup>f</sup>	1.72	3.56E-02	4.42E-02	1.07
	Isoleucine <sup>f</sup>	1.94	1.26E-03	1.06E-02	1.3
	Glutamate <sup>f</sup>	1.68	6.85E-03	2.43E-02	1.35
	Pyroglutamic acid <sup>f</sup>	1.83	2.60E-03	1.62E02	1.25
	Catechol <sup>e</sup>	1.77	5.82E04	7.48E-03	-1.83
	Pipecolinic acid <sup>f</sup>	1.76	1.51E-03	1.31E-02	1.65
	2-Aminoadipic acid <sup>e</sup>	1.51	6.58E-03	2.37E-02	-1.27
	2-Aminobutyric acid <sup>f</sup>	1.89	1.19E-02	3.12E-02	1.45
	Glycocyamine <sup>e</sup>	1.96	6.74E-04	8.10E-03	-1.89

Abbreviations: FDR, false discovery rate; GC-TOF, gas chromatograph time of flight; OPLS, orthogonal projection to latent structures; SZ-BL, schizophrenia patients at baseline.

For metabolites from urine analyzed by GC-TOF.

<sup>a</sup>VIP shows variable importance in the projection obtained from the OPLS model with a cutoff of 1.5.

<sup>b</sup>P-values from non-parametric Wilcoxon–Mann–Whitney test.

<sup>c</sup>FDR was calculated with a *q*-value of 0.2.

<sup>d</sup>FC is a positive value fold change that indicates a relatively higher concentration present in SZ-BL, while a negative value means a relatively lower concentration as compared to the controls.

<sup>e</sup>Metabolites were identified using available library databases.

<sup>f</sup>Metabolites were verified by reference compounds.

**Table 4** Metabolites identified by <sup>1</sup>H-NMR and GC-TOFMS

Metabolite	NMR			GC-TOFMS	
	Shift	P-value <sup>a</sup>	FC <sup>b</sup>	P-value	FC
Glucose	3.46	1.73E-04	1.79	2.87E-01	1.14
Citrate	2.66	5.58E-03	-1.34	2.75E-01	-1.19
Lactate	4.10	2.09E-08	-1.30	7.58E-03	1.48
Acetone	2.22	1.90E-03	1.74	Not found	
Acetoacetate	2.26	5.96E-02	1.23	Not found	
3-Hydroxybutyrate	1.22	6.47E-01	1.58	5.83E-03	1.37

Abbreviations: GC-TOFMS, gas chromatograph time-of-flight mass spectrometer; NMR, nuclear magnetic resonance; SZ-BL, schizophrenia patients at baseline.

<sup>a</sup>P-values from non-parametric Wilcoxon–Mann–Whitney test.

<sup>b</sup>FC is a positive value fold change that indicates a relatively higher concentration present in SZ-BL, while a negative value means a relatively lower concentration as compared to the controls.

of five metabolites (glycerate, eicosenoic acid, β-hydroxybutyrate, pyruvate and cystine) was established using logistic regression. This biomarker panel is a ‘good’ classifier for diagnosing schizophrenia in that AUC achieved 0.945 in the training set and 0.895 in the test set. In particular, the fact that patients and

controls in the test set were not matched in gender and age suggested an unbiased discrimination of this set of biomarkers in the general population. In addition, supplement of urine 3-hydroxybutyrate to the serum panel offers an equivalent diagnosis to the clinical interviews, with AUC achieving 1 in both



training and test set. Further, insights for schizophrenia mechanism could be obtained from pathways behind these biomarkers.

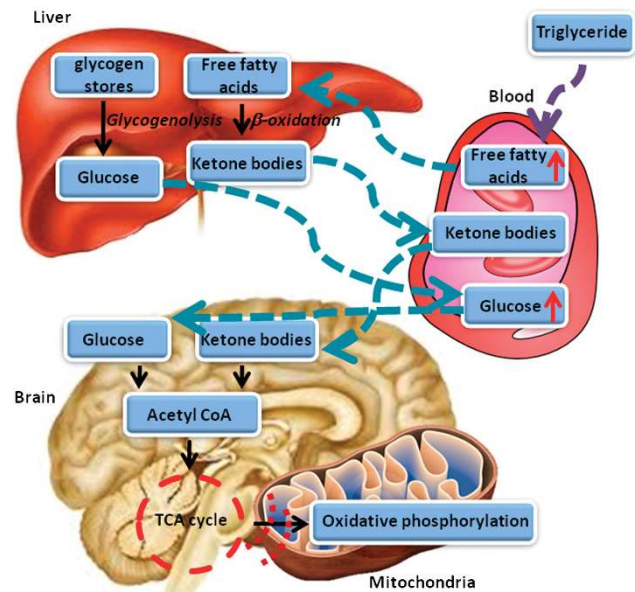
#### Metabolic evidences for energy metabolism disorder

Pyruvate is an important intermediate in glucose metabolism, connecting glycolysis and the tricarboxylic acid cycle. In adults, the brain accounts for 20% of the total body basal oxygen consumption and 25% of the total body glucose utilization, although constituting only 2% of body weight.<sup>39</sup> Glycolysis coupled with the tricarboxylic acid cycle fully oxidizes glucose and supplies energy for the brain. An increased supply of glucose in the blood has been found to enable better memory performance in schizophrenic patients.<sup>40</sup> Converging evidence<sup>41–44</sup> suggests that malfunction of glucose metabolism may be a causative factor for schizophrenia. A higher prevalence of glucose tolerance and insulin resistance in patients with first-onset schizophrenia has been attributed to abnormal glucose metabolism.<sup>45–47</sup> An increased level of pyruvate observed in this study indicates increased energy production and demand in schizophrenic subjects. A higher energy demand has been found to result from inefficiency in brain circuitry.<sup>48</sup>

Glycerate, eicosenoic acid and  $\beta$ -hydroxybutyrate are the metabolites produced from the degradation of glyceride and fatty acids. Owing to the elevated levels of glycerate, eicosenoic acid indicates an increased fatty acid catabolism as an alternative energy source in schizophrenia subjects. Especially, concentration of eicosenoic acid was significantly increased in the pretreatment of schizophrenic patients ( $P=8.56E-08$ , fold change = 1.96), which has also been reported as being elevated in the red blood cells of children suffering from autism as compared with typically developing controls.<sup>49</sup>

The brain requires continuous delivery of substrate by the blood to support energy consumption. Ketone bodies can and do in some naturally occurring conditions substitute, at least in part, for glucose.<sup>39</sup> Ketone bodies are produced primarily in the liver as a by-product of fatty acid mobilization. As the liver increases its rate of production of ketone bodies in association with increased fatty acid metabolism, it releases them into the blood for utilization by other tissues. This elevated serum fatty acids and  $\beta$ -hydroxybutyrate is presumed to be as a result of high level of glyceride mobilization in response to higher energy demand in the brain, which is supported by the accumulation of short- to middle-chain fatty acids and ketone bodies in urine (Figure 3). A long-term shortage of energy from glucose metabolism may push the brain to change the expression of genes related to fatty acid metabolism, resulting in higher concentration of transcripts in post-mortem brain tissue.<sup>24</sup>

Pathways related to energy metabolism have been found to be associated with schizophrenia based on



**Figure 3** Disordered energy metabolism in schizophrenia. The brain uses glucose in circulation as its main energy source, with ketone bodies as an alternative. In schizophrenic patients, brain energy supply is scarce owing to mitochondrial dysfunction. Hence, the brain is presumed to have to partially shift its energy supply towards ketone bodies, and fatty acid metabolism in the liver is then mobilized to produce the necessary ketone bodies.

evidence from proteins, transcripts and metabolites in the human brain.<sup>23,24</sup> Based on the serum and urine metabolite changes we observed, we hypothesized that systematic alternations on glycogenolysis and lipid metabolism occurred in the schizophrenia patients (Figure 3). Fatty acid mobilization, resulting from intrinsic inhibited glucose metabolism in the brain, accounts for the increase of free fatty acids in the serum and ketone bodies in the urine of schizophrenic patients. The increased consumption of ketone bodies also suggests that a shortage in glucose utilization may push the brain to shift part of its energy supply from glucose towards ketone bodies in schizophrenic patients. If the disease progresses, the brain cells have to increase the transcription of fatty acid metabolism, as evidenced by decreased glycolysis and increased glycogen and fatty acid catabolism transcripts reported in the post-mortem brain.<sup>24</sup>

In our study, free fatty acids such as tetradecanoic acid, hexadecanoic acid and octadecanoic acid were found significantly elevated in the serum of schizophrenia patients at baseline, which is not consistent with previously reported plasma lipids.<sup>30</sup> This might be due to the different classes of compounds (free fatty acids vs lipids) detected with the respective platforms. For example, GC-TOFMS can readily capture small-molecule fatty acids, but is not amenable to detecting phospholipids.

#### Cystine (cysteine) in antioxidant defense system

Reactive oxygen species (ROS) resulting from oxidative phosphorylation and metabolic transformations

cause oxidative damage in cellular DNA, RNA, proteins as well as lipids constructing various membranes.<sup>50</sup> The antioxidant glutathione (GSH; L- $\gamma$ -glutamyl-L-cysteinyl-glycine) is essential for the cellular detoxification of ROS in brain cells.<sup>51</sup> Cystine is the preferred form of cysteine for the synthesis of GSH in cells involved in the immune function, including macrophages and astrocytes (<http://www.hmdb.ca>). Cystine, as the oxidized form of cysteine in the extracellular space (serum), is a rate-limiting substrate for GSH synthesis. Maintenance of intracellular GSH drives cystine transport across plasma membrane under oxidative stress.<sup>52</sup> The depletion of serum cystine observed in this study may be indicative of a compromised GSH system in the brain under increased oxidative stress in neurological diseases. This observation is supported by the higher concentration levels of 2-hydroxybutyrate (and its amino derivative, 2-aminobutyrate) in the serum of the schizophrenia patients, which is a by-product of the reaction-producing cysteine from homocysteine. These results are consistent with a previous study in which transcripts related to GSH synthesis were found to be increased in human brain tissues.<sup>24</sup> In addition, 2-hydroxybutyrate has been found to be an early biomarker for insulin resistance and glucose intolerance in a non-diabetic population, indicating that it also has a close association with energy metabolism.<sup>53</sup>

The elevation of cystine in the serum and depletion in the urine is presumably due to the different metabolic/excretion rates in different biological compartments.

#### *Glutamate as neurotransmitter and/or energy source*

Glutamate is the predominant excitatory neurotransmitter of the vertebrate central nervous system.<sup>54</sup> It plays a critical role in synaptic maintenance and plasticity,<sup>55</sup> and also involves in learning and memory. Under normal conditions, brain glutamate concentration is highly regulated at both sides of the brain–blood barrier. As a result, the fluctuation of glutamate concentration in circulation brings little influence to glutamate in the brain. However, the alteration of serum glutamate level is of neurological importance as it is closely associated with the metabolism of  $\gamma$ -aminobutyric acid, the primary inhibitory neurotransmitter.

The elevated concentration of glutamate in the serum is presumably indicative of insufficient energy supply in the brain under schizophrenia, as glutamate also serves as an important fuel reserve. When insufficient glucose concentration or glycolytic flux occurs in the brain, glutamate is mobilized as alternative fuel to glucose.<sup>56</sup>

#### **Conclusion**

We used the GC-TOFMS and NMR analytical platforms to characterize the metabolic profiles of serum and urine from schizophrenic patients. A biomarker panel composing of five serum metabolites (glycerate, eicosenoic acid,  $\beta$ -hydroxybutyrate, pyruvate and

cystine) was established to discriminate schizophrenic patients from normal controls. A composite panel by the addition of urine  $\beta$ -hydroxybutyrate to the serum panel achieved AUC of 1 in both the training set and test set, meaning an equivalent diagnosis to clinical interviews. Elevated fatty acids and ketone bodies were found in both the serum and urine of patients, suggesting an upregulated fatty acid catabolism, presumably resulting from an insufficiency of glucose supply in the brains of schizophrenia patients. In addition, metabolic profile of schizophrenia patients on first episode was found similar to that of recurrent ones when we compared the metabolites in the serum and urine from the two groups.

#### **Conflict of interest**

The authors declare no conflict of interest.

#### **Acknowledgments**

This work was supported by the 973 Program (2006CB910600, 2010CB529600, 2007CB914700, 2007CB947300), the National Nature Science Foundation of China (31071159), the National Key Technology R&D Program (2006BAI05A09), the 863 Program (2009AA022701), the Shanghai Municipal Commission of Science and Technology Program (09DJ1400601), the National Key Project for the Investigation of New Drugs (2008ZX09312-003) and the Shanghai Leading Academic Discipline Project (B205).

#### **References**

- 1 Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Med* 2005; **2**: e141.
- 2 Freedman R. Schizophrenia. *N Engl J Med* 2003; **349**: 1738–1749.
- 3 MacDonald AW, Schulz SC. What we know: findings that every theory of schizophrenia should explain. *Schizophr Bull* 2009; **35**: 493–508.
- 4 McGorry PD, Yung AR, Phillips LJ, Yuen HP, Francey S, Cosgrave EM *et al*. Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. *Arch Gen Psychiatry* 2002; **59**: 921–928.
- 5 Holmes E, Tsang TM, Huang JT, Leweke FM, Koethe D, Gerth CW *et al*. Metabolic profiling of CSF: evidence that early intervention may impact on disease progression and outcome in schizophrenia. *PLoS Med* 2006; **3**: e327.
- 6 McGlashan TH, Zipursky RB, Perkins D, Addington J, Miller T, Woods SW *et al*. Randomized, double-blind trial of olanzapine versus placebo in patients prodromally symptomatic for psychosis. *Am J Psychiatry* 2006; **163**: 790–799.
- 7 Horrobin DF. Schizophrenia: a biochemical disorder? *Biomedicine* 1980; **32**: 54–55.
- 8 Puri BK, Easton T, Das I, Kidane L, Richardson AJ. The niacin skin flush test in schizophrenia: a replication study. *Int J Clin Pract* 2001; **55**: 368–370.
- 9 Tavares H, Yacubian J, Talib LL, Barbosa NR, Gattaz WF. Increased phospholipase A2 activity in schizophrenia with absent response to niacin. *Schizophr Res* 2003; **61**: 1–6.
- 10 Bosveld-van Haandel L, Knegtering R, Kluiters H, van den Bosch RJ. Niacin skin flushing in schizophrenic and depressed patients and healthy controls. *Psychiatry Res* 2006; **143**: 303–306.

- 11 Abou-Saleh MT. Neuroimaging in psychiatry: an update. *J Psychosom Res* 2006; **61**: 289–293.
- 12 Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D et al. Common variants conferring risk of schizophrenia. *Nature* 2009; **460**: 744–747.
- 13 Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I et al. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009; **460**: 753–757.
- 14 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
- 15 Teng PN, Bateman NW, Hood BL, Conrads TP. Advances in proximal fluid proteomics for disease biomarker discovery. *J Proteome Res* 2010; **9**: 6091–6100.
- 16 Cowan ML, Vera J. Proteomics: advances in biomarker discovery. *Expert Rev Proteomics* 2008; **5**: 21–23.
- 17 Kaddurah-Daouk R, Krishnan KR. Metabolomics: a global biochemical approach to the study of central nervous system diseases. *Neuropsychopharmacology* 2009; **34**: 173–186.
- 18 Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 1999; **29**: 1181–1189.
- 19 Zyromski NJ, Mathur A, Gowda GAN, Murphy C, Swartz-Basile DA, Wade TE et al. Nuclear magnetic resonance spectroscopy-based metabolomics of the fatty pancreas: implicating fat in pancreatic pathology. *Pancreatol* 2009; **9**: 410–419.
- 20 Boudonck KJ, Mitchell MW, Nemet L, Keresztes L, Nyska A, Shinar D et al. Discovery of metabolomics biomarkers for early detection of nephrotoxicity. *Toxicol Pathol* 2009; **37**: 280–292.
- 21 Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 2009; **457**: 910–914.
- 22 Kaddurah-Daouk R, Soares JC, Quinones MP. Metabolomics: a global biochemical approach to the discovery of biomarkers for psychiatric disorders. In: Turck C (ed). *Biomarkers for Psychiatric Disorders*. Springer: New York, USA, 2009, pp 1–34.
- 23 Khaitovich P, Lockstone HE, Wayland MT, Tsang TM, Jayatilaka SD, Guo AJ et al. Metabolic changes in schizophrenia and human brain evolution. *Genome Biol* 2008; **9**: R124.
- 24 Prabakaran S, Swatton JE, Ryan MM, Huffaker SJ, Huang JT, Griffin JL et al. Mitochondrial dysfunction in schizophrenia: evidence for compromised brain metabolism and oxidative stress. *Mol Psychiatry* 2004; **9**: 684–697, 643.
- 25 Horrobin DF. The membrane phospholipid hypothesis as a biochemical basis for the neurodevelopmental concept of schizophrenia. *Schizophr Res* 1998; **30**: 193–208.
- 26 Schwarz E, Prabakaran S, Whitfield P, Major H, Leweke FM, Koethe D et al. High throughput lipidomic profiling of schizophrenia and bipolar disorder brain tissue reveals alterations of free fatty acids, phosphatidylcholines, and ceramides. *J Proteome Res* 2008; **7**: 4266–4277.
- 27 Quinones MP, Kaddurah-Daouk R. Metabolomics tools for identifying biomarkers for neuropsychiatric diseases. *Neurobiol Dis* 2009; **35**: 165–176.
- 28 Huang JT, Leweke FM, Tsang TM, Koethe D, Kranaster L, Gerth CW et al. CSF metabolic and proteomic profiles in patients prodromal for psychosis. *PLoS One* 2007; **2**: e756.
- 29 Tsang TM, Huang JT, Holmes E, Bahn S. Metabolic profiling of plasma from discordant schizophrenia twins: correlation between lipid signals and global functioning in female schizophrenia patients. *J Proteome Res* 2006; **5**: 756–760.
- 30 Kaddurah-Daouk R, McEvoy J, Baillie RA, Lee D, Yao JK, Doraiswamy PM et al. Metabolomic mapping of atypical antipsychotic effects in schizophrenia. *Mol Psychiatry* 2007; **12**: 934–945.
- 31 Yao JK, Dougherty Jr GG, Reddy RD, Keshavan MS, Montrose DM, Matson WR et al. Altered interactions of tryptophan metabolites in first-episode neuroleptic-naive patients with schizophrenia. *Mol Psychiatry* 2010; **15**: 938–953.
- 32 Yao JK, Dougherty Jr GG, Reddy RD, Keshavan MS, Montrose DM, Matson WR et al. Homeostatic imbalance of purine catabolism in first-episode neuroleptic-naive patients with schizophrenia. *PLoS One* 2010; **5**: e9508.
- 33 Qiu Y, Cai G, Su M, Chen T, Zheng X, Xu Y et al. Serum metabolite profiling of human colorectal cancer using GC-TOFMS and UPLC-QTOFMS. *J Proteome Res* 2009; **8**: 4844–4850.
- 34 Jonsson P, Gullberg J, Nordstrom A, Kusano M, Kowalczyk M, Sjostrom M et al. A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS. *Anal Chem* 2004; **76**: 1738–1745.
- 35 Jonsson P, Johansson AI, Gullberg J, Trygg J, A J, Grung B et al. High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. *Anal Chem* 2005; **77**: 5635–5642.
- 36 Golland P, Liang F, Mukherjee S, Panchenko D. Permutation tests for classification. In: Auer P, Meir R (eds). *Learning Theory*, vol. 3559, Springer: Berlin/Heidelberg, 2005, pp 501–515.
- 37 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Statist Soc Ser* 1995; **B57**: 12.
- 38 Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recogn* 1997; **30**: 1145–1159.
- 39 Sokoloff L. Metabolism of ketone bodies by the brain. *Annu Rev Med* 1973; **24**: 271–280.
- 40 Newcomer JW, Craft S, Fucetola R, Moldin SO, Selke G, Paras L et al. Glucose-induced increase in memory performance in patients with schizophrenia. *Schizophr Bull* 1999; **25**: 321–335.
- 41 Olsen L, Hansen T, Jakobsen KD, Djurovic S, Melle I, Agartz I et al. The estrogen hypothesis of schizophrenia implicates glucose metabolism: association study in three independent samples. *BMC Med Genet* 2008; **9**: 39.
- 42 Hazlett EA, Buchsbaum MS, Kemether E, Bloom R, Platholi J, Brickman AM et al. Abnormal glucose metabolism in the mediodorsal nucleus of the thalamus in schizophrenia. *Am J Psychiatry* 2004; **161**: 305–314.
- 43 Martins-de-Souza D, Harris LW, Guest PC, Bahn S. The role of energy metabolism dysfunction and oxidative stress in schizophrenia revealed by proteomics. *Antioxid Redox Signal* 2011; **15**: 2067–2079.
- 44 Buchsbaum MS, Hazlett EA. Positron emission tomography studies of abnormal glucose metabolism in schizophrenia. *Schizophr Bull* 1998; **24**: 343–364.
- 45 Ryan MC, Collins P, Thakore JH. Impaired fasting glucose tolerance in first-episode, drug-naive patients with schizophrenia. *Am J Psychiatry* 2003; **160**: 284–289.
- 46 Spelman LM, Walsh PI, Sharifi N, Collins P, Thakore JH. Impaired glucose tolerance in first-episode drug-naive patients with schizophrenia. *Diabet Med* 2007; **24**: 481–485.
- 47 Fernandez-Egea E, Bernardo M, Donner T, Conget I, Parellada E, Justicia A et al. Metabolic profile of antipsychotic-naive individuals with non-affective psychosis. *Br J Psychiatry* 2009; **194**: 434–438.
- 48 Buchsbaum MS, Buchsbaum BR, Hazlett EA, Haznedar MM, Newmark R, Tang CY et al. Relative glucose metabolic rate higher in white matter in patients with schizophrenia. *Am J Psychiatry* 2007; **164**: 1072–1081.
- 49 Bu B, Ashwood P, Harvey D, King IB, Water JV, Jin LW. Fatty acid compositions of red blood cell phospholipids in children with autism. *Prostaglandins Leukot Essent Fatty Acids* 2006; **74**: 215–221.
- 50 Apel K, Hirt H. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* 2004; **55**: 373–399.
- 51 Dringen R, Hirrlinger J. Glutathione pathways in the brain. *Biol Chem* 2003; **384**: 505–516.
- 52 McBean GJ. Cerebral cystine uptake: a tale of two transporters. *Trends Pharmacol Sci* 2002; **23**: 299–302.
- 53 Gall WE, Beebe K, Lawton KA, Adam KP, Mitchell MW, Nakhle PJ et al. Alpha-hydroxybutyrate is an early biomarker of insulin

- resistance and glucose intolerance in a nondiabetic population. *PLoS One* 2010; **5**: e10883.
- 54 Smith QR. Transport of glutamate and other amino acids at the blood–brain barrier. *J Nutr* 2000; **130**(Suppl): 1016S–1022S.
- 55 McDonald JW, Johnston MV. Physiological and pathophysiological roles of excitatory amino acids during central nervous system development. *Brain Res Brain Res Rev* 1990; **15**: 41–70.
- 56 Miller AL, Hawkins RA, Veech RL. Decreased rate of glucose utilization by rat brain *in vivo* after exposure to atmospheres

containing high concentrations of CO<sub>2</sub>. *J Neurochem* 1975; **25**: 553–558.



**This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>**

Supplementary Information accompanies the paper on the Molecular Psychiatry website (<http://www.nature.com/mp>)