A metabonomic approach to early prognostic evaluation of experimental sepsis by $^1$H NMR and pattern recognition

Zhong-ying Lin$^a$, Ping-bo Xu$^{b,d}$, Shi-kai Yan$^c$, Hai-bing Meng$^b$, Gen-jin Yang$^a$, Wei-xing Dai$^a$, Xin-ru Liu$^a$, Jin-bao Li$^b$, Xiao-ming Deng$^b$**, Wei-dong Zhang$^{a,c}$*

This study proposes an NMR-based metabonomic approach to early prognostic evaluation of sepsis. Forty septic rats receiving cecal ligation and puncture (CLP) were divided into the surviving group and nonsurviving group on day 6, while 20 sham-operated rats served as the control group. Serum samples were collected from septic and sham-operated rats at 12 h after surgery and analyzed using $^1$H NMR spectroscopy. Orthogonal partial least squares (OPLS) were applied and showed clustering according to predefined groups, indicating that NMR-based metabolic profiling could reveal pathologic characteristics in the serum of sham-operated, surviving, and nonsurviving septic rats. In addition, six characteristic metabolites including lactate, alanine, acetate, acetoacetate, hydroxybutyrate, and formate, which are mainly involved in energy metabolism, changed markedly in septic rats, especially in the nonsurvivors. Using these metabolites, a predictive model for prognostic evaluation of sepsis was constructed using a radial basis function neural network (RBFNN) with a prediction accuracy of about 87% by test samples. The results indicated that the NMR-based metabonomic approach is a potential technique for the early prognostic evaluation of sepsis. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: metabonomics; sepsis; $^1$H NMR; serum; prognostic evaluation; OPLS; RBFNN; k-NN

INTRODUCTION

Sepsis is a systemic inflammatory response syndrome resulting from dysregulated immune responses of the host to injury or infection (1). Sepsis may deteriorate fast into multi-organ dysfunction syndrome, and eventually to death if efficient treatments are not performed immediately (2). Despite various advances in antibiotic and supportive therapies, the mortality rate of sepsis-related diseases remains high in intensive care units (ICUs) (3, 4) due to delayed evaluation of illness severity and inappropriate treatment of the condition (5). Accurate and timely evaluation of illness severity is therefore urgently needed to limit mortality, reduce costs, and improve outcome (6, 7).

The prognostic evaluation of sepsis is quite a challenging task for the clinician, and considerable efforts have been attempted to find a sensitive and specific predictive marker (8). C-reactive protein (9) and procalcitonin (10, 11) are two candidate biomarkers for sepsis that have received a lot of attention, but they still have a long way to go before being widely used in clinical practices; the pathological characteristics of sepsis cannot be fully described by such a limited panel of biomarkers, (12, 13). In fact, sepsis involves not only specifically expressed proteins but...
also metabolite changes. Some metabolites have already been proposed as biomarkers for prognostic evaluation of sepsis (14–16). However, due to the complexity of the molecular mechanisms underlying the inflammatory processes, there are various metabolites, not just one or two biomarkers, related to sepsis (17). Thus, it is possible to develop a predictive approach based on a holistic metabolic profiling to identify septic patients with poor outcome in the early stage.

Metabonomics is a well-developed platform for studying systems biology, leading to high-throughput screening processes in clinical diagnosis (18). It alters the traditional concepts of single biomarker analysis by aiming at detecting and using holistic metabolic patterns for clinical diagnosis. Proton nuclear magnetic resonance (1H NMR) spectroscopy and liquid chromatography combined with mass spectrometry (LC/MS) are now routinely applied for detecting changes in metabolic profiles (19, 20). 1H NMR spectroscopic analysis of biofluids offers a high potential for the understanding of biochemical processes associated with disease in an organism (21, 22). It provides information on both the structure and the composition of low-molecular-mass metabolites in biological fluids, and is a rapid and low-cost technique for exploring pathological metabolic processes. Serum is one of the most widely used fluids for diagnostic purposes, but it is very complex from the analytical point of view. 1H NMR spectroscopy could be used to determine a great number of endogenous metabolites in serum. However, the data from 1H NMR spectra of serum are quite complex involving signals originating from hundreds of metabolites. Thus, multivariate statistics and pattern recognition techniques are often used to monitor the levels of endogenous metabolites for disease diagnosis (23).

Metabonomics has been proved to be an effective tool for biomarker discovery and clinical diagnosis (24). In our previous study, a LC/MS-based metabolomic approach was successfully applied for the early prognostic evaluation of experimental sepsis (25). However, some low-molecular-weight polar metabolites, which might be very important for the prognostic evaluation of sepsis, could not be detected efficiently by MS due to their complex reversed-phase retention behavior. 1H NMR allows for unbiased analysis of many types of metabolites and can provide additional information on these molecules. Therefore, the present study used 1H NMR to further explore serum metabolic characteristics of sepsis in rats in an attempt to find more potential biomarkers for prognosis evaluation of sepsis.

**EXPERIMENTS**

**Animals and the CLP model**

Male specific-pathogen-free Sprague-Dawley rats (220–250 g; 6–8 weeks) were purchased from the Shanghai Experimental Animal Center of the Chinese Academy of Sciences (Shanghai, China). Rats were housed in specific pathogen-free conditions and allowed to acclimatize in our facility for 1 week before experiments. Experimental sepsis was induced by cecal ligation and puncture (CLP) as described previously (26). Briefly, a 2.5-cm incision was made in the lower right quadrant of the abdomen to expose the cecum. The distal two-thirds of the cecum (approximately 2.5 cm long) were ligated tightly with a 7th suture and punctured through once with a 14-, 18-, or 22-gauge needle. The cecum was re-placed into peritoneal cavity and the abdomen was closed with 4th suture. In sham surgical controls, the cecum was exposed but not ligated or punctured. All animals were administered 4 mL of sterile saline s.c. for fluid resuscitation, and given free access to normal diet and drinking after surgery. The postoperative survival duration and rate were recorded during the following 6 days.

**Determination of the sampling time**

Forty-eight rats were equally assigned to an untreated group, a sham-operated group (sampling at 12 h after sham operation), a CLP-6 h group (sampling at 6 h after CLP), a CLP-12 h group (sampling at 12 h), a CLP-18 h group (sampling at 18 h), and a CLP-24 h group (sampling at 24 h). Then 3.0 mL arterial blood was drawn from left carotid artery at preset time. 0.5 mL arterial blood was used for blood gas analysis after anticoagulation, and the remaining 2.5 mL was left clotted for 2 h at room temperature. After centrifugation at 3000 g for 20 min, the aliquot was separated for liver function test, biochemistry test, and metabonomic analysis. After euthanasia, the left lower lung and kidney specimens were collected, fixed in 10% formalin with PBS buffer (pH 7.2), and embedded in paraffin. Five-micrometer-thick sections were made and stained with hematoxylin and eosin (H&E) for morphologic analysis.

**Sampling**

Sixty rats were randomized into a sham-operated group (n = 20) and a CLP group (n = 40). According to the difference of survival duration during the following 6 days, CLP rats were divided into a surviving group (n = 19, survival duration exceeding 6 days) and a nonsurviving group (n = 21, survival duration between 24 h and 6 days). About 1.0 mL blood was collected by tail shearing method at 12 h after surgery and then left clotted for 2 h at room temperature. Sera were separated by centrifugation at 3000 rpm for 15 min. Four hundred microliters of each sample were mixed with 50 μL D2O for locking signal and with 50 μL phosphate buffer solution (0.2 M Na2HPO4/0.2 M NaH2PO4; pH 7.4) to minimize pH variations of the serum samples. Then, TPS (3-trimethylsilylpropanesulfonate sodium; 1 mM final concentration) was added as an internal standard. The samples were stored at −80°C until NMR analysis.

**1H NMR analysis**

NMR measurements were made on a Bruker AVANCE II 600 MHz spectrometer at 298 K. NMR spectra of the serum samples were acquired using a solvent presaturation pulse sequence to suppress the residual water resonance. Free induction decays (FIDs) were collected into 64k data points, with a spectral width of 7289 Hz and an acquisition time of 2.04 s, giving a total pulse recycle delay of 3.04 s. The data were zero filled by a factor of 2, and the FIDs were multiplied by an exponential weighting function equivalent to a line broadening of 0.3 Hz prior to Fourier transformation.

**Data analysis**

All spectra were baseline corrected using multipoint baseline correction method (27), and chemical shifts were adjusted with reference to TPS. The chemical shift region of 4.60–5.06 was removed to eliminate any spurious effects of variability on the suppression of the water resonance. Peak detection (freely accessed from http://www.wam.umd.edu/~toh/spectrum) and peak matching were performed using programs coded by...
Matlab, and a data matrix containing all peak intensities was thus generated. The data were logarithm transformed and centered for orthogonal partial least squares (OPLS) analysis, which is derived from the projection to latent structure regression method (28). The OPLS model was constructed using the NMR data as the X variables and the class information identifier as the Y variable. In OPLS, the pathologic differences in rats can be visualized in score plots and metabolites responsible for the differences could be statistically identified in regression coefficient plots using Hotelling’s T² test (29).

Modeling

To construct predictive models, two-thirds of the samples were used as the training set and the remaining as the test set. The training set was used to construct the predictive model, while the test set was used to evaluate the accuracy of the model prediction. Knowing that sample selection for the training set belong to. The standard k-NN algorithm was implemented as follows: (1) calculate the Euclidean distance between an unknown sample and each training sample; (2) select k samples from the training set that were nearest to the unknown sample; and (3) classify the unknown sample into the class to which the majority of the k samples belong. The parameter k was optimized using eightfold cross-validation method.

RBFNN can be described as a special feed-forward structure consisting of three layers: the input layer, the hidden layer, and the output layer. The input layer passes the input vectors to the hidden layer. The hidden layer, which consists of a number of neurons with each employing a radial basis function (RBF), executes nonlinear transformation on the input data and passes the results to the output layer. The most often used RBF is a Gaussian function that is characterized by center c_j and width r_j, and is given as \( h_j(x) = \exp(-||x - c_j||^2/r^2_j) \). The number of RBFs \( n_h \) and the width of spread \( r \), which could greatly affect the performance of the model, were optimized by experimenting with a number of trials using eightfold cross-validation method. All programs were coded in MATLAB 7.0 (The Mathworks Inc., Natick, MA, USA).

RESULTS

CLP model

The survival rate of the CLP rats depended on the size of the needle used to puncture the cecum (Fig. 1). In the present study, 14-, 18-, and 22-gauge needles were attempted to puncture, causing a low-, middle-, and high-grade CLP models with a mortality of 93, 57, and 7%, respectively, on day 6 after surgery, and finally the 18-gauge needle was selected to produce the CLP model since it provided an appropriate survival rate.

Optimal time for metabonomic prediction

To implement early outcome prediction, a time course study of physiological parameters and histological sections of the lung and kidney was carried out during sepsis. As shown in Table 1,

### Table 1. Results of blood examination in CLP-induced septic rats.

<table>
<thead>
<tr>
<th>Groups</th>
<th>PaO₂ (mmHg)</th>
<th>PaCO₂ (mmHg)</th>
<th>TB (μmol/L)</th>
<th>ALT (U/L)</th>
<th>AST (U/L)</th>
<th>Creatinine (μmol/L)</th>
<th>Urea (mmol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated</td>
<td>95.2 ± 4.1</td>
<td>40.2 ± 3.1</td>
<td>0.14 ± 0.04</td>
<td>30.9 ± 2.7</td>
<td>67.0 ± 16.1</td>
<td>16.9 ± 2.2</td>
<td>4.43 ± 0.64</td>
</tr>
<tr>
<td>Sham-operated</td>
<td>94.0 ± 4.8</td>
<td>41.0 ± 3.8</td>
<td>0.15 ± 0.06</td>
<td>54.0 ± 9.4</td>
<td>85.4 ± 21.3</td>
<td>19.1 ± 4.9</td>
<td>4.49 ± 0.66</td>
</tr>
<tr>
<td>CLP-6 h</td>
<td>94.9 ± 4.9</td>
<td>37.9 ± 4.1</td>
<td>0.15 ± 0.05</td>
<td>58.8 ± 10.5</td>
<td>90.6 ± 20.1</td>
<td>18.8 ± 3.2</td>
<td>5.09 ± 0.91</td>
</tr>
<tr>
<td>CLP-12 h</td>
<td>92.5 ± 4.2</td>
<td>36.5 ± 4.5</td>
<td>0.21 ± 0.08</td>
<td>66.9 ± 15.5</td>
<td>117.9 ± 43.5</td>
<td>25.8 ± 9.5</td>
<td>5.11 ± 0.88</td>
</tr>
<tr>
<td>CLP-18 h</td>
<td>86.9 ± 5.5abc</td>
<td>31.9 ± 5.6abc</td>
<td>0.71 ± 0.26</td>
<td>90.3 ± 34.1</td>
<td>266.4 ± 120.1abcd</td>
<td>36.1 ± 12.1abc</td>
<td>5.70 ± 1.04ac</td>
</tr>
<tr>
<td>CLP-24 h</td>
<td>82.1 ± 6.5bd</td>
<td>28.1 ± 6.2bd</td>
<td>0.84 ± 0.35</td>
<td>112.4 ± 36.2</td>
<td>316.9 ± 146.8bd</td>
<td>43.1 ± 23.6bd</td>
<td>7.25 ± 1.28bd</td>
</tr>
</tbody>
</table>

Data are expressed as mean ± SD (n = 8). Abbreviations: PaO₂, partial pressure of oxygen in arterial blood; PaCO₂, partial pressure of carbon dioxide in artery; TB, total bilirubin; ALT, alanine aminotransferase; AST, aspartate aminotransferase.

\( ^* p < 0.05 \)

\( ^{ab} p < 0.01 \) versus untreated group.

\( ^{ac} p < 0.05 \)
laboratory parameters reflecting hepatic, renal, and respiratory functions started to change at 12 h after CLP but did not reach significant abnormality until 18 h after CLP when compared with those of the untreated and sham-operated rats, which is consistent with the results of histological examination of the lung and kidney specimens (Figs 2 and 3). However, some marked alterations were clearly observed in metabolic profiles of septic rats at that time by $^1$H NMR (Fig. 3). The results suggested that changes in serum metabolic profiles of the rats were earlier than those of organ dysfunction, and it was possible to establish a predictive model at 12 h after CLP by the NMR-based metabonomic approach. Therefore, blood samples for metabonomic prediction were collected at 12 h after CLP in our study.

$^1$H NMR spectra analysis

Typical $^1$H NMR spectra of the serum samples are shown in Fig. 4. Resonances were assigned according to the international literature (32–34) and the Human Metabolome Database (available at http://www.hmdb.ca). The chemical shifts for the identified metabolites are listed in Table 2. By data preprocessing, a total of 221 peaks were resolved for each serum sample, and OPLS was performed on the normalized data to better summarize and investigate the changes and any correlation with sepsis, and to aid visualization and detection of patterns that might reflect or match the sepsis. Two orthogonal components, which explain 16.2% variation of the total, were calculated for the model to remove the variation in the NMR spectra unrelated to class information. Validation of the model was assessed by an eightfold cross-validation method, and the cross-validation predictive ability $Q^2(y)$ was 0.80, indicating a good predictability of the model, and $R^2$, which represents the total explained variation for the $X$, was about 8.1%. As shown in Fig. 5A, samples from different groups were well separated along the first PLS component, which indicates that NMR-based metabolic profiling could reveal characteristic pathological alterations in serum from sham-operated, surviving, and nonsurviving septic rats.
Since OPLS concentrates all discriminating information into the first component, it is sufficient to plot the regression coefficients for this component only. As shown in Fig. 5B, seven most important resonances responsible for the separation were selected by Hotelling’s $T^2$ test. These resonances are structurally postulated as hydroxybutyrate (1.20 ppm), lactate (1.34 ppm), alanine (1.49 ppm), acetate (1.94 ppm), acetoacetate (2.25 ppm), and formate (8.48 ppm). As seen in Table 2, the level of formate and five other metabolites changed markedly in the septic rats, especially in the nonsurviving ones.

Prognostic evaluation

Based on the promising results showing complete separation of the samples from the sham-operated, survivors, and nonsurvivors using OPLS, we proceeded with using the identified metabolites to construct predictive models for the prognostic evaluation of sepsis. Using the training set, two predictive models were built upon the identified metabolites by $k$-NN and RBFNN. To evaluate the prediction accuracy of these models, the test set was input into the predictive models. The predictive accuracy of $k$-NN

![Figure 4. Representative $^1$H NMR spectra of serum from (A) sham-operated, (B) surviving and (C) nonsurviving septic rats. Abbreviations: s, singlet; d, doublet.](image)

### Table 2. Assignments of the metabolites and their intensities in rat serum from shams, survivors, and nonsurvivors.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Group</th>
<th>$\delta^1$H (ppm)$^a$</th>
<th>Multiplicity$^b$</th>
<th>Sham ($n = 20$)</th>
<th>Survivor ($n = 19$)</th>
<th>Nonsurvivor ($n = 21$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxybutyrate</td>
<td>$^3$CH$_3$</td>
<td>1.20</td>
<td>d</td>
<td>$4.90 \pm 2.53$</td>
<td>$57.7 \pm 19.7^d$</td>
<td>$68.6 \pm 24.8^d$</td>
</tr>
<tr>
<td>Lactate</td>
<td>$^6$CH$_3$</td>
<td>1.34</td>
<td>d</td>
<td>$31.8 \pm 10.4$</td>
<td>$155 \pm 139^d$</td>
<td>$196 \pm 135^d$</td>
</tr>
<tr>
<td>Alanine</td>
<td>$^2$CH$_3$</td>
<td>1.49</td>
<td>d</td>
<td>$2.29 \pm 0.60$</td>
<td>$12.0 \pm 11.7^d$</td>
<td>$21.4 \pm 5.95^d,e$</td>
</tr>
<tr>
<td>Acetate</td>
<td>$^4$CH$_3$</td>
<td>1.94</td>
<td>s</td>
<td>$1.76 \pm 0.45$</td>
<td>$10.1 \pm 7.3^d$</td>
<td>$22.3 \pm 12.5^d,e$</td>
</tr>
<tr>
<td>Acetoacetate</td>
<td>$^7$CH$_3$</td>
<td>2.25</td>
<td>s</td>
<td>$4.62 \pm 0.17$</td>
<td>$13.2 \pm 10.5^d$</td>
<td>$16.8 \pm 16.1^d$</td>
</tr>
<tr>
<td>Formate</td>
<td>HCOO$^-c$</td>
<td>8.48</td>
<td>s</td>
<td>$5.11 \pm 0.26$</td>
<td>$3.34 \pm 0.67^d$</td>
<td>$1.93 \pm 0.37^d,e$</td>
</tr>
</tbody>
</table>

$^a$Chemical shifts are reported with reference to 3-trimethylsilylpropanesulfonyl (TPS) singlet resonance at 0.000 ppm.

$^b$Multiplicity definitions: s, singlet; d, doublet.

$^c$Data are expressed as mean $\pm$ SD.

$^d, e p < 0.01$ versus sham-operated group.

$^p < 0.01$ versus survivors.
In our previous study, a LC/MS-based metabonomic approach was developed to find potential biomarkers for the prognostic evaluation of sepsis, and six fatty acids were structurally identified (25). However, some low-molecular weight metabolites generated during sepsis that are neither retained in LC nor ionizable in MS, could not be detected efficiently by LC/MS. Alternatively, \(^1\)H NMR spectroscopy, which allows for an unbiased analysis of several types of metabolites, could provide useful information about them. Thus in the present study, \(^1\)H NMR spectroscopy was employed to further investigate serum metabolic profiles of the septic rats. The results showed that the levels of six characteristic metabolites that are mainly involved in energy metabolism, including lactate, alanine, acetate, acetoacetate, hydroxybutyrate, and formate, changed markedly in septic rats and especially in the nonsurvivors.

Lactate is one of the most important characteristic metabolites indicating a high risk of multi-organ failure, and its level was significantly increased in the septic rats, especially in the nonsurvivors. The reason for the increased lactate is quite complex (45), and might partly result from tissue hypoxia (46). In the presence of insufficient oxygen supply, pyruvate is disproportionately converted into lactate rather than entering into the Krebs cycle pathway. Defects in mitochondrial metabolic pathways might also contribute to lactate accumulation in the absence of tissue hypoxia (47, 48).

Alanine and acetate were also significantly increased in the septic rats, presumably resulting from enhanced pyruvate metabolism. In the septic rats, glucose metabolism is enhanced and the level of pyruvate increased accordingly, which could result in an increase in alanine through transamination (49). Pyruvate may also convert into acetyl CoA via oxidative decarboxylation, then to acetyl phosphate, and eventually form acetate. Thus, increased acetate might be due to the enhanced metabolism of pyruvate.

Acetoacetate and hydroxybutyrate are two ketone bodies, and their increase might be related to enhanced fatty acid oxidation in the septic rats. As the major source of energy, fatty acid oxidation is significantly enhanced to meet the energy requirement (50), and ketone bodies, which are important products of fatty acid oxidation, are increased and accumulated accordingly.

Formate is a potential alternative single-carbon source for the production of N5, N10-methylene-THF required for biosynthesis of nucleoic acid. The serum level of formate decreased markedly in the septic rats, especially in the nonsurvivors, which might have resulted from enormous consumption of formate due to significant increase in biosynthesis of purine nucleotide in the septic condition.

Using the metabolites discovered by \(^1\)H NMR, two predictive models for the prognostic evaluation of experimental sepsis were constructed by using \(k\)-NN and RBFNN. The results showed that the performance of RBFNN was superior to that of \(k\)-NN in terms of predictive accuracy. It is generally accepted that diseases and metabolites often behave in a very complicated nonlinear relationship, and RBFNN, as a strong nonlinear approximation algorithm, can model such a complex relationship. As a result, RBFNN was finally adopted to build the model for the early prognostic evaluation of sepsis. After validation by test samples, the prediction accuracy of this model reached about 87%, demonstrating that the NMR-based metabonomic approach with RBFNN was able to accurately predict the outcome of the septic rats in the early stage.
In summary, this study provides a metabonomic approach for prognostic evaluation of sepsis based on the integral information about serum metabolites. Being rapid, efficient, and less dependent on personal clinical experience, the approach constitutes a new concept in prognostic evaluation of sepsis, and offers a promising screening candidate tool for clinicians to predict the outcome of septic patients in ICU.

CONCLUSION

In this work, a metabonomic study of sepsis was performed using $^1$H NMR and pattern recognition techniques, which allows early holistic evaluation of pathological changes of sepsis. With the help of OPLS, six characteristic metabolites, which could be used for prognostic evaluation of sepsis, were structurally postulated. These metabolites include lactate, alanine, acetate, acetoacetate, hydroxybutyrate, and formate, most of which are involved in energy metabolism. Using these metabolites, a RBFNN model was constructed for early prognostic evaluation of sepsis that achieved a prediction accuracy of about 87%. The results indicate that the NMR-based metabonomic approach is a promising technique for early prognostic evaluation of sepsis with the advantages of being rapid, less costly, and more efficient, and it is expected that it will be developed as a supplementary tool for the clinical prognostic evaluation of sepsis.

Acknowledgements

The work was supported by the program for Changjiang Scholars and Innovative Research Team in University (PCSIRT), NCET Foundation, NSFC (30725045, the National 863 Program (2006AA02Z338), Shanghai Leading Academic Discipline Project (8906), and partly by the Scientific Foundation of Shanghai China (07DZ19728, 06DZ19717, 06DZ19005).

REFERENCES

40. Davis BH, Olsen SH, Ahmad E, Bigelow NC. Neutrophil CD64 is an improved indicator of infection or sepsis in emergency department patients. Arch Pathol Lab Med 2006; 130: 654–661.