Huiping Zhang
Yan Wang
Xue Gu
Junyi Zhou
Chao Yan

School of Pharmacy, Shanghai Jiao Tong University, Shanghai, P. R. China

Received August 25, 2010
Revised October 9, 2010
Accepted October 27, 2010

Research Article

Metabolomic profiling of human plasma in pancreatic cancer using pressurized capillary electrochromatography

The application of pressurized capillary electrochromatography (pCEC) coupled with ultra violet (UV) detection has been investigated for the production of global metabolite profiles from human plasma, and its capabilities of classifying pancreatic cancer patients. The pCEC separation of plasma samples was performed on a RP column with gradient elution. The applied voltage, detection wavelength and type of acid modifiers on separation of plasma samples were optimized with pooled quality control (QC) sample. The stability and the repeatability of the methodology were also determined by the repeat analysis of QC sample. The effects of different scaling methods on the results of orthogonal partial least-squares discrimination analysis (OPLS-DA) based on pCEC-UV data set were also investigated. The results of the current study clearly showed the different phenotypes of metabolites of pancreatic cancer patients and healthy controls based on pCEC-UV plasma profiles. OPLS-DA data are shown to provide a valuable means of convenient classification. This work indicated that pCEC-UV method can be used as a cost-effective and information-rich, while relatively simple and inexpensive approach for plasma profiling on disease metabolomics studies.

Keywords:
Chemometrics / Metabolomic profiling / Pancreatic cancer / Plasma / Pressurized capillary electrochromatography DOI 10.1002/elps.201000431

1 Introduction

Pancreatic cancer is one of the most fatal human malignancies. It was estimated that over 42 470 patients would be diagnosed with pancreatic cancer and 35 240 patients would die of pancreatic cancer in 2009 in the Unites States alone [1]. It was reported that only 10–15% of pancreatic cancer patients were suitable for resection, and the 5-year survival rate was only 10–29% after resection [2–4]. The high mortality rate is attributed to its aggressive growth property [5] and the lack of effective early diagnosis. Therefore, most patients have been diagnosed with advanced stage of pancreatic cancer. Accordingly, surgical cure is no longer a feasible option for most patients when the disease is diagnosed.

Correspondence: Professor Chao Yan, School of Pharmacy, Shanghai Jiao Tong University, Shanghai 200240, P. R. China
E-mail: chaoyan@unimicrotech.com
Fax: +86-21-3420-5908

Abbreviations: cLC, capillary LC; **MeOH,** methanol; **OPLS-DA,** orthogonal partial least-squares discrimination analysis; **PCA,** principal component analysis; **pCEC,** pressurized capillary electrochromatography; **QC,** quality control; **UV,** ultra violet

Currently, there is no approved modality for screening pancreatic cancer in the general population. The deep anatomic location of the pancreas makes routine physical examination ineffective, and magnetic resonance imaging is not cost-effective. Many serologic markers have been examined as potential screening tools, such as CA 19-9, MUC1, CEACAM1, MIC1 [6–8]. However, none of these markers has sufficient sensitivity and specificity when screening asymptomatic patients. It is urgent to develop better methods for early diagnosis of pancreatic cancer [9]. Metabolite profiling is a potential new type of diagnostic test that shows considerable promise.

Metabonomics is a branch of science concerned with the study of systems biology. It concerns the study of low molecular weight compounds (typically <1000 Da) in biofluids and other complex matrixes [10, 11]. Over 2100 endogenous metabolites are found in human biofluids currently [12]. Many metabolites are the final downstream products of genome and reflect best the operation of the biological system. To date, global metabolic profiling of human biofluids has been increasingly used as an effective tool for disease diagnosis [13, 14] to elucidate significant changes in tumor metabolism [15], to explore candidate 'biomarkers' from variance within a huge

Colour Online: See the article online to view Figs. 4, 5 and 7 in colour

number of endogenous metabolites [16, 17] and to characterize the biological pathways [10]. Recently, metabolic profiling of serum specimens of pancreatic cancer patients has revealed significant variations in the benign hepatobiliary disease and the pancreatic cancer [18, 19]. A comprehensive mass spectrometry based analytical platform (such as UPLC/MS and GC/MS) was established to detect multiple compounds previously unreported in plasma from pancreatic cancer patients [20]. These studies indicated that serum metabolic profiling might be useful for pancreatic cancer diagnosis. Metabolomic studies generally employ techniques such as NMR, GC/MS, HPLC/MS, UPLC/MS and CE. As none of these analytical techniques can separately resolve all the compounds in a metabolome, comprehensive metabolome analysis that will get more detailed metabolic profiling is a good choice [15, 21]. With regard to LC, a key area for further innovation in metabolic profiling is the use of higher resolution separation systems [22]. Capillary LC (cLC) provides higher resolution compared with conventional LC, and the two-dimensional LC separation combining dual separation technique also provides a route to increase metabolome coverage.

CEC is a hybrid technique of cLC and CE, offering the advantages of high efficiency, high selectivity, fast speed, low sample and solvent consumption, etc. [23, 24]. CEC is a powerful separation technique for neutral and charged compounds. However, CEC also suffers from some practical difficulties, such as bubbles formation and column drying-out during experiment. In a pressurized CEC (pCEC) system, supplementary pressure can be applied to the column in CEC to overcome these problems [25, 26]. Quantitative sample introduction into pCEC can be achieved through a rotary-type injector and solvent gradient elution can be realized [27]. These characteristics of pCEC provide better separation selectivity and repeatability than CEC. Due to the merits of the pCEC technique, it has the potential to become a powerful separation tool for complex mixtures such as biological fluids. pCEC has been used for metabolic profiling of rat urine [28], and it showed that pCEC was able to measure more urinary metabolites with good resolution than a conventional RPLC or cLC method. Ultra violet (UV) detection is generally available to a great population of scientists and perhaps the most rugged technique, which makes it an attractive alternative for plasma profiling [29–33].

In the present study, pCEC was used in the global metabolite profiling study of human plasma to provide a potential platform for the diagnosis of pancreatic cancer. The applied voltage, the detection wavelength and the type of acid modifiers on the separation of plasma samples were optimized. As the data quality is paramount for the analysis, we used a biological QC approach to investigate the stability of the analytical methodology and the repeatability of the plasma pretreatment method [34]. Besides, the effects of different scaling methods on the results of orthogonal partial least-squares discrimination analysis (OPLS-DA) were also investigated.

2 Materials and methods

2.1 Reagents and materials

Both methanol (MeOH) and ACN were of HPLC grade and obtained from Tedia (Fairfield, OH, USA), trifluoroacetic acid (TFA) was of HPLC grade (CNW, Germany), and distilled water was produced by the Milli-Q Reagent Water System (Millipore, MA, USA). All other chemicals and reagents were analytical grade from China National Pharmaceutical Group (Shanghai, China).

2.2 Samples

Plasma samples were obtained from 17 pancreatic cancer patients and 18 healthy volunteers from Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University (Shanghai, China). An informed consent was given by all the patients and volunteers, and approval was obtained from the local research ethics committee. The 17 pancreatic cancer patients, aged from 42 to 78 years old, including 9 male patients and 8 female patients were diagnosed through imaging modalities such as computerized tomography, magnetic resonance or endosonography and then histologically proven to be with early-stage pancreatic cancer without clinical evidence of distant metastasis, with CA 19-9 ranging from 1000 to 1500. The 18 healthy volunteers, aged from 40 to 75 years old, including 10 males and 8 females, were selected by a routine physical examination. The whole blood was collected into heparinized tubes. Then the samples were centrifuged at $3000 \times g$ for 10 min and stored at $-80^{\circ}C$ until use.

For quality control (QC), a pooled 'QC' sample was prepared by mixing 70 μL from each of the samples of healthy people before analysis [35–37]. The QC samples were also used in the optimization of separation and determination conditions.

2.3 Samples preparation

Each 70 μL plasma sample was mixed into 280 μL mixture of organic solvents (MeOH, ACN and acetone, 1:1:1; v/v/v) and vortexed for 1 min to precipitate the proteins. The mixture was kept at $4^{\circ}C$ for 10 min and then centrifuged at 13 000 rpm for 12 min at room temperature. The supernatant (without any particles at the bottom) was put into a clean tube and evaporated to dryness under nitrogen at room temperature. Then the dried residue was reconstituted (re-dissolved) in 70 μL of 50% MeOH v/v and vortexed for 2 min. After that, it is centrifuged again at 13 000 rpm for 10 min. The supernatant was transferred into a new tube and stored at $4^{\circ}C$ before pCEC-UV analysis.

For the QC sample, 150 μL plasma was mixed into 600 μL mixture of the same organic solvents and processed

in the same manner except that the dried residue was reconstituted in 150 μL of 50% MeOH v/v.

2.4 Gradient pCEC-UV analysis

Chromatographic separation was performed on a reversed-phase column (EP-150-30/50-5-C18, Global Chromatography) of 50 cm (of which 30 cm was packed) × 150 μm id packed with 5 μm C18 particles using the TriSep-2100 pCEC system (Unimicro Technologies, Pleasanton, CA, USA). The system was composed of a binary gradient pumps, a high-voltage power supply (−30 and +30 kV), a microfluid manipulation module with a six-port injection valve, a variable wavelength UV–vis detector (190–600 nm) and a data collection workstation of Unimicro Trisep-2003. Samples were injected into an external sample loop of 5 μL and then carried by the mobile phases, which were driven by the binary gradient pumps and entered a four-port split valve. The mobile phase flow was split into two ways, of which one was led to the capillary column under constant pressure controlled by a back-pressure regulator and another was led to the waste reservoir. A negative voltage was applied to the outlet of the column, and the inlet of the column was connected to the four-port split valve and grounded.

The separation was performed using gradient elution with 5% MeOH (v/v, 0.05% TFA) (A) and 95% (MeOH/ACN, 70:30, v/v) (0.05% TFA) (B) as mobile phases at a flow rate of 0.06 mL/min. The thiourea was used as a non-retention marker to measure the dead time. The flow rate in the capillary column was 1.19 mm/s and the spilt ratio was 95:1. The gradient condition was 0–5 min, 10% B; 5–10 min, 10–25% B; 10–45 min, 25–100% B; 45–80 min, 100% B. A voltage of −4 kV was applied to the capillary. The detection wavelength was set at 214 nm. The plasma samples used in the optimization of separation and determination conditions were QC samples. Besides, the QC samples were also used in the process of methodology validation to study the robustness of the analytical method. All the samples were analyzed once at a random order.

2.5 Data pretreatment and statistical analysis

The pCEC-UV raw data files were initially converted into NetCDF format using the Unimicro Trisep-2003 Workstation, and then imported into the scripts written in MATLAB 7.1 (The MathWorks, USA). The data preprocessing including baseline correction, filtering, peak detection, peak matching and normalization of the total sum of the chromatogram was performed. Data were only used for the period 10–68 min, as the non-retention peaks in 0–10 min and column re-conditioning period of 68–80 min was eliminated. Then, we treated the data with the pretreatment methods including mean centering, auto-scaling and pareto-scaling, respectively, to compare the influence of different methods on the analysis result. The resultant three-dimensional matrix encompassing peak indices (retention time), sample names (observation) and UV intensity (variable) was then exported to the SIMCA-P+ 12.0 software package (Umetrics, Umeå, Sweden) for multivariate statistical analysis.

Unsupervised principal component analysis (PCA) was first used to give an overview of plasma pCEC-UV data from cancer patients and healthy control people, which would reveal outliers, groups and trends in the data [38]. For further identifying the differentially expressed metabolites accountable for the separation between the different groups, more sophisticated OPLS-DA was carried out on the data set.

To ensure the reliability of the OPLS-DA model, it was validated with three different methods. First and foremost, the data set was randomly divided into training set and test set. Twenty-six samples (13 controls/13 patients) were used as the training set and the remaining nine samples (5 controls/4 patients) as the test set. The OPLS-DA model was trained with the training set and the prediction ability was tested with the test set. Besides, the seven-fold cross-validation with one-seventh of the samples excluded from the model and the 999 random permutation tests were also performed to avoid overfitting the model.

3 Results and discussion

3.1 Optimization of separation and determination conditions

3.1.1 Selection of detection wavelength and optimization of composition of the mobile phase

Assessing untargeted differences in biological systems requires the detection of as many metabolites as possible. In this study, the chromatograms at 214, 238, 254 and 265 nm wavelength were compared, and finally the wavelength of 214 nm (Fig. 1) was chosen because most analytes of interest showed absorbance peak at this wavelength and reflected the information of metabolites in the plasma as greatly as possible.

In order to find out the effects of solvent compositions on separation by pCEC, we investigated a range of different solvent compositions. Phase A was fixed to be 5% MeOH and phase B was chosen among 95% ACN, 95% MeOH and 95% mixture of ACN and MeOH, whereas the TFA concentration was kept 0.05% and the applied voltage at −4 kV. After the data from these various systems were analyzed with respect to resolution and selectivity, it was found that the 95% (MeOH/ACN, 70:30, v/v) (0.05% TFA)/5% MeOH (0.05% TFA) gradient system had the best peak shape, resolution, and selectivity, which was used for further experiment.

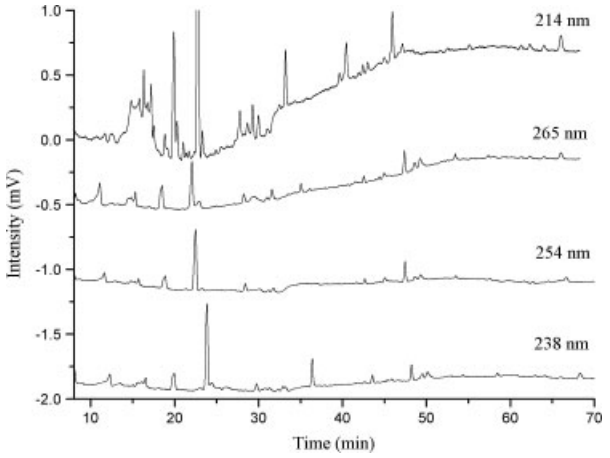

Figure 1. Electrochromatograms at different UV wavelengths. Experimental conditions: capillary column, 50 cm (packed length 30 cm) × 150 μm id packed with ODS (5 μm); mobile phase A was 0.05% TFA in 5% MeOH v/v and mobile phases B was 0.05% TFA in 95% (MeOH/ACN, 70:30, v/v); the gradient condition was 0–5 min, 10% B; 5–10 min, 10–25% B; 10–45 min, 25–100% B; 45–80 min, 100% B; applied voltage, −4 kV; total pump flow rate, 0.06 mL/min; injection volume, 52 nL.

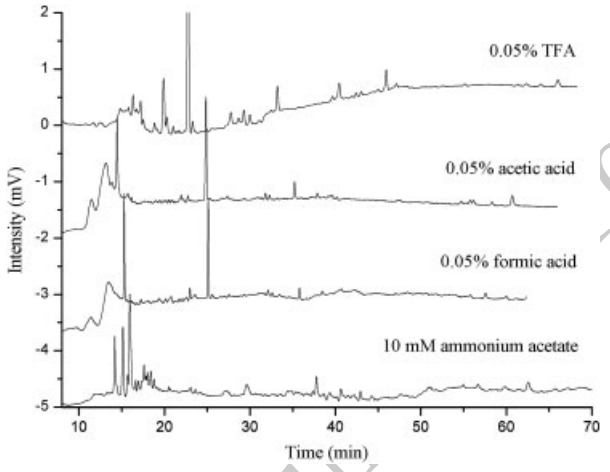

Figure 2. Effect of acid modifiers on separation of human plasma. Conditions: the gradient condition was 0–5 min, 10% B; 5–10 min, 10–25% B; 10–45 min, 25–100% B; 45–80 min, 100% B; applied voltage, −4 kV; total pump flow rate, 0.06 mL/min; injection volume, 52 nL.

3.1.2 Effect of acid modifiers

The effects of different acids such as formic acid, acetic acid, TFA and ammonium acetate on the separation of the plasma samples were investigated. We can see in Fig. 2 that more satisfactory peak shape and resolution were achieved when the pH of the mobile phase was adjusted with TFA compared with acetic acid, formic acid and ammonium acetate.

Furthermore, chromatograms with different concentrations of TFA varying from 0.01 to 0.1% were compared. Lower resolution was obtained when the TFA concentration

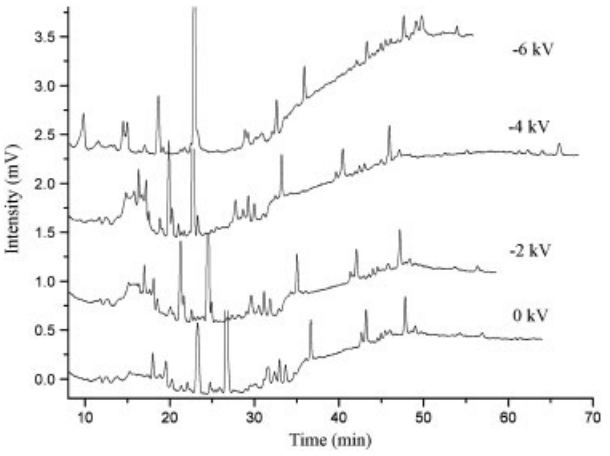

Figure 3. Effect of applied voltage on the pCEC separation of the human plasma. UV detection, 214 nm; other experimental conditions are the same as in Fig. 1.

was too low, while increasing the concentration above 0.1% led to greater background current which caused Joule heating. Therefore, 0.05% TFA was finally chosen to improve the peak shape and resolution.

3.1.3 Effect of applied voltage

Plasma is a very complex analytical system, as the sample contains both polar/ionizable metabolites and non-polar lipids covering a wide range of lipophilicity. It poses a real challenge to the separation technology. Compared with the pure HPLC, pCEC can improve the separation efficiency through changing the EOF and the electrophoretic mobility of charged solutes.

In this study, the effect of the applied voltage (from 0 to −6 kV) on the resolution and analysis time of the plasma sample was investigated. Figure 3 shows that the retention time of some peaks would become shorter when the higher voltage was applied. However, it also indicates that the resolution became better with the applied voltage increasing from 0 to −4 kV, whereas the resolution would become worse when the applied voltage was beyond −4 kV. Therefore, the applied voltage was chosen to be −4 kV, at which the best resolution with higher efficiency was obtained.

3.2 Methodology validation

Five independent samples were prepared from one QC sample and injected, respectively, to study the repeatability with respect to retention time and peak height. The result for eight common peaks selected to cover a range of retention time was shown in Table 1. Excellent repeatability for retention time was seen (variation in retention time was negligible with CV% values $<1.4\%$). The variation in the peak height was found to be broadly acceptable with CV% values ranging from 3.390 to 15.307%. The above results

Table 1. Repeatability in retention time and peak height of eight common peaks in QC plasma obtained by pCEC-UV ($n = 5$)

No.	Retention time		Peak height	
	Mean ± SD (min)	CV (%)	Mean ± SD (min)	CV (%)
1	16.112 ± 0.210	1.303	282.000 ± 37.000	13.120
2	19.309 ± 0.186	0.963	484.200 ± 16.407	3.390
3	23.035 ± 0.201	0.873	2124.800 ± 87.300	4.109
4	28.866 ± 0.0816	0.283	205.400 ± 16.410	7.999
5	32.123 ± 0.106	0.330	121.400 ± 12.178	10.03
6	35.700 ± 0.112	0.317	126.200 ± 19.318	15.307
7	47.848 ± 0.030	0.063	453.400 ± 57.292	12.685
8	67.655 ± 0.228	0.337	202.200 ± 29.903	14.798

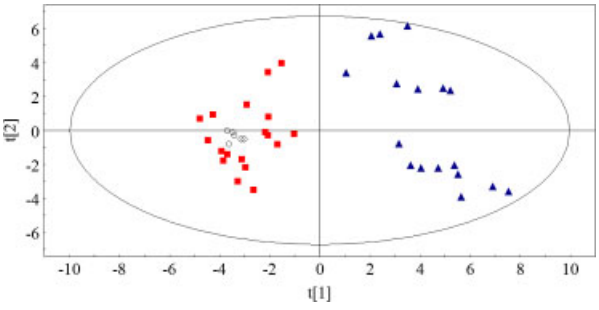

Figure 4. PCA score plot (Comp. 1 versus Comp. 2) of all samples analyzed by pCEC. (■) Health control; (▲) pancreatic cancer patient; and (◦) QC.

showed that the developed method has a reasonably good reproducibility [39].

Additionally, one QC sample was injected every seven samples to further study the stability of the analysis. A multivariate data analysis of pCEC results was performed using PCA and the data pretreatment was performed with auto-scaling. An overview of the data set was performed by PCA revealing outliers, groups and trends in the data. Figure 4 is the PCA scores plot, which showed a clear separation between pancreatic cancer patients and health control and displayed a tight clustering of the QC samples. The positions of the objects in the model plane stood for the relationship of the objects. Therefore, the objects close to each other had a similar multivariate profile, given the *k*-descriptors, whereas the objects far from each other had dissimilar properties. The tighter the clustering of the QC samples in scores plot, the more repeatable the runs were. This unsupervised analysis of PCA demonstrated a degree of platform stability prior to further analysis [35]. Besides, PCA was also applied on the six QC samples separately. In Fig. 5, the time series properties of the first component showed little shifts of the QC samples with time. The result further gave a good demonstration on the stability of the analysis for the duration of the run.

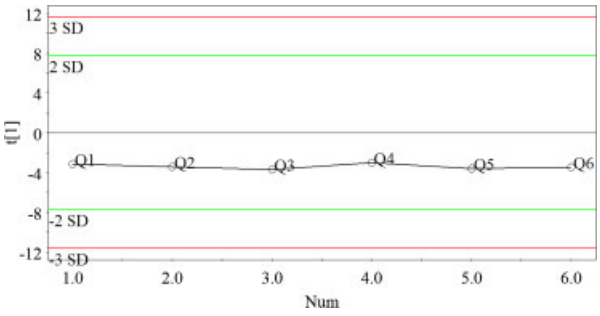

Figure 5. Time series plot of the first PCA component ($t[1]$ versus sample run order).

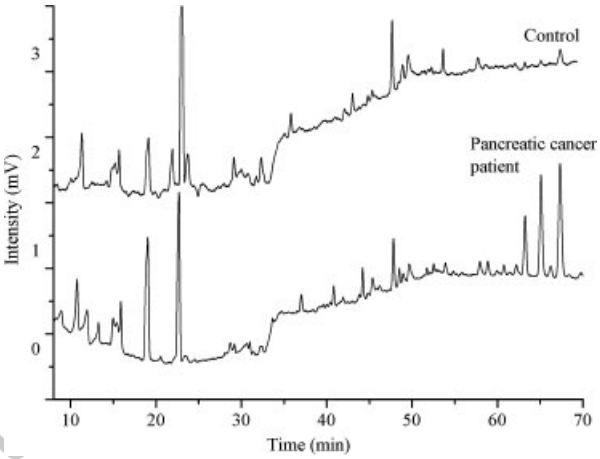

Figure 6. Typical electrochromatograms of a pancreatic cancer patient and the control generated by pCEC-UV analysis. Conditions: capillary column, 50 cm (packed length 30 cm) × 150 μm id packed with ODS (5 μm); mobile phase A was 0.05% TFA in 5% MeOH v/v and mobile phase B was 0.05% TFA in 95% (MeOH/ACN, 70:30, v/v); the gradient condition was 0–5 min, 10% B; 5–10 min, 10–25% B; 10–45 min, 25–100% B; 45–80 min, 100% B; applied voltage, −4 kV; total pump flow rate, 0.06 mL/min; applied voltage, −4 kV; UV detection, 214 nm; injection volume, 52 nL.

3.3 Metabolic profiles between pancreatic cancer patients and healthy controls

Plasma metabolic profiling was performed and a typical pCEC chromatography was presented in Fig. 6, some of the metabolic changes in patients could be found directly in it, indicating that there were probably different phenotypes of metabolites. The metabolic profiles contain abundant characteristic features, and the metabolic profile of each sample is unique, so it is quite difficult to find out the common difference between two sets of samples just through comparing the profiles of two individual samples. Hence, multivariate analysis can be used as an alternative method to further examine the changes in the metabolic fingerprint. Multivariate data analysis allows for the detection of variations in the levels of low concentration metabolites which are difficult to detect with the naked eye, yet are often important biomarkers indicating a

metabolic defect. In this study, the difference between pancreatic cancer patients and healthy controls was investigated by OPLS-DA model, which is a potent tool for the classification of metabolomic data, useful for classifying new samples, and at the same time allows identification of biomarkers for the studied diseases.

Before multivariate data analysis, data processing and data pretreatment should be performed for the raw analytical data [40]. Data processing is to eliminate the influence of background noise and retention time shifting on data analysis through deionizing, baseline correction and alignment. Scaling, a widely used method of pretreatment, constitutes another crucial step. Variables often have very different numerical ranges. A variable with a large range has a large variance, whereas a variable with a small range has a small variance. Since multivariate data analysis is based on maximum variance, they should be scaled. Different scaling methods emphasize different aspects of the data and alter the covariance structure of the data, affecting classification ability of the method and chosen biomarkers.

Auto-scaling and Pareto-scaling are the prevailing methods applied in the metabolomic MS-based studies. Here, we compared the models with the above two scaling methods, the mean-centered model and the model without pretreatment. The model parameters and the results of the OPLS-DA models with data preprocessed with different methods were summarized in Table 2. The model parameters for the explained variation R^2Y and the predictive capability Q^2 were all quite high, ranging from 0.906 to 0.978 for R^2Y and from 0.765 to 0.925 for Q^2, which indicated that the models were quite good. No overfitting was found according to the permutation validation. The auto-scaled model had the highest capability to explain the variation in the relatively high predictive capability, second only to the Pareto-scaled model which was also an excellent model. However, the mean-centered, unscaled model (Ctr) showed quite lower explanative and predictive capability, and the model without pretreatment (None) was really a poor model. From the results of the models, it could also be seen that the auto-scaled model had the best classification rate (33 out of 35), followed by the Pareto-scaled model (32 out of 35), whereas the model without pretreatment had a poor classification capability. The calculated sensitivity and specificity were 94.1 and 94.4% for the auto-scaled model and 88.2 and 94.4% for the Pareto-scaled model, respectively, based on a 95% confidence limit for class membership. The auto-scaled *T*-predicted score scatter plots for the OPLS-DA models were shown in Fig. 7. It was apparent that the samples were successfully classified by the models. Auto-scaling gave the best result for the data set, and Pareto-scaling could also be an alternative to auto-scaling for UV-based study.

Once differences between patterns have been established, further identification of metabolites responsible for differences in the profiles was one of the aims of metabolomics research. On the basis of the OPLS-DA results with a good group classification between pancreatic cancer and controls, a total of nine retention time variables were selected according to the variable importance in the projection (VIP) threshold (VIP >1.5, a stricter threshold can improve the reliability of differential variables). These differential variables are responsible for the deviated metabolic profiles. We did not identify exactly what the analytes were, as pCEC with UV detection could not provide molecular mass and structure information, which is a demerit of this tool. However, it provides a cost-effective and information-rich, while relatively simple and inexpensive approach for metabolomics studies. The aim of this study was to develop a simple and robust pCEC-UV method to enable discrimination of metabolic plasma profiles between pancreatic cancer and health controlled person with no prior knowledge of compounds being analyzed. The results clearly showed that based on pCEC-UV plasma profiling and OPLS-DA, different phenotypes of metabolites between pancreatic cancer and health control could be recognized and discriminated. To detect a disease through the whole profile instead of target analytes, pCEC-UV could be an economical, clinical tool once the method was exhaustively validated in humans. A further study in progress is the utility of pCEC coupled with MS which can further improve the sensitivity and the identification of metabolites in samples.

Table 2. Model parameters and results of OPLS-DA models

Data scaling method	Components	R^2Y (cum)[a)]	Q^2 (cum)[b)]	% Correct[c)]
None	1+1	0.696	0.565	25
Ctr	1+1	0.944	0.909	31
Pareto	1+1	0.958	0.925	32
Auto	1+2	0.978	0.914	33

a) R^2Y(cum) represents the cumulative sum of squares (SS) of all the *Y*'s explained by all extracted components.

b) Q^2 (cum) is an estimate of how well the model predicts the *Y*'s.

c) % Correct: classification success on the basis of a predicted *Y* cutoff of 0.5.

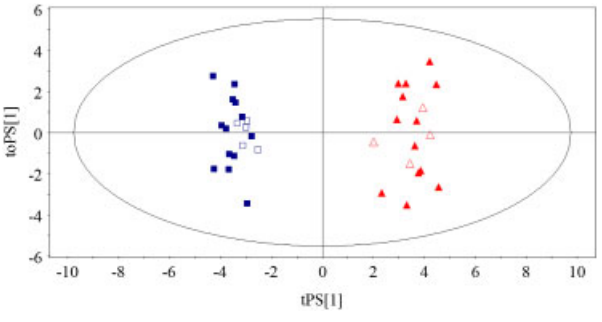

Figure 7. Auto-scaled OPLS-DA *T* predicted score scatter plots of all samples analyzed by pCEC: (■) health control in training set; (□) health control in test set; (▲) pancreatic cancer patient in training set; and (□) pancreatic cancer patient in test set.

4 Concluding remarks

pCEC is a high efficiency separation method with dual separation mechanism, namely chromatographic partition and electrophoretic mobility. The aim of this paper was to develop a pCEC-UV method to enable discrimination of plasma metabolic profiles in pancreatic cancer. The separation condition was optimized with plasma samples. After the voltage was applied to the system, the separation resolution became better than that of cLC. Based on the assessment on the QC samples, pCEC-UV provided an effective and repeatable method for global metabolite profiling of human plasma. Different scaling methods were also studied and compared for the pCEC-UV data. The results showed that pCEC-UV method combined with OPLS-DA could be used for discrimination between pancreatic cancer patients and healthy people and as a potential tool for disease. pCEC coupled with MS could be served as more powerful platform for disease biomarker discovery. Further study in this direction is in progress.

This work was supported by Fund of Transformation of Agriculture Science and Technology Achievements of Ministry of Science and Technology (2008002), Scientific Instrument Support Fund of Shanghai Science and Technology Commission (09142200202) and National Natural Science Foundation of China (20705018). National comprehensive technology platforms for innovative drug R&D (2009ZX09301-007).

The authors have declared no conflict of interest.

5 References

[1] Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Thun, M. J., *CA Cancer J. Clin.* 2009, *59*, 225–249.

[2] Trede, M., Schwall, G., Saeger, H. D., *Ann. Surg.* 1990, *211*, 447–458.

[3] Yeo, C. J., Cameron, J. L., Sohn, T. A., Lillemoe, K. D., Pitt, H. A., Talamini, M. A., Hruban, R. H., Ord, S. E., Sauter, P. K., Coleman, J., Zahurak, M. L., Grochow, L. B., Abrams, R. A., *Ann. Surg.* 1997, *226*, 248–257; discussion 257–260.

[4] Nitecki, S. S., Sarr, M. G., Colby, T. V., van Heerden, J. A., *Ann. Surg.* 1995, *221*, 59–66.

[5] Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., Thun, M. J., *CA Cancer J. Clin.* 2008, *58*, 71–96.

[6] Koopmann, J., Rosenzweig, C. N., Zhang, Z., Canto, M. I., Brown, D. A., Hunter, M., Yeo, C., Chan, D. W., Breit, S. N., Goggins, M., *Clin. Cancer Res.* 2006, *12*, 442–446.

[7] Gold, D. V., Modrak, D. E., Ying, Z., Cardillo, T. M., Sharkey, R. M., Goldenberg, D. M., *J. Clin. Oncol.* 2006, *24*, 252–258.

[8] Simeone, D. M., Ji, B., Banerjee, M., Arumugam, T., Li, D., Anderson, M. A., Bamberger, A. M., Greenson, J., Brand, R. E., Ramachandran, V., Logsdon, C. D., *Pancreas* 2007, *34*, 436–443.

[9] Lee, M. X., Saif, M. W., *JOP* 2009, *10*, 104–108.

[10] Nicholson, J. K., Connelly, J., Lindon, J. C., Holmes, E., *Nat. Rev. Drug Discov.* 2002, *1*, 153–161.

[11] Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R. N., Willmitzer, L., *Nat. Biotechnol.* 2000, *18*, 1157–1161.

[12] Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M. A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., Amegbey, G., Block, D., Hau, D. D., Wagner, J., Miniaci, J., Clements, M., Gebremedhin, M., Guo, N., Zhang, Y., Duggan, G. E., Macinnis, G. D., Weljie, A. M., Dowlatabadi, R., Bamforth, F., Clive, D., Greiner, R., Li, L., Marrie, T., Sykes, B. D., Vogel, H. J., Querengesser, L., *Nucleic Acids Res.* 2007, *35*, D521–D526.

[13] Vallejo, M., García, A., Tunón, J., García-Martínez, D., Angulo, S., Martin-Ventura, J. L., Blanco-Colio, L. M., Almeida, P., Egido, J., Barbas, C., *Anal. Bioanal. Chem.* 2009, *394*, 1517–1524.

[14] Wikoff, W. R., Gangoiti, J. A., Barshop, B. A., Siuzdak, G., *Clin. Chem.* 2007, *53*, 2169–2176.

[15] Qiu, Y., Cai, G., Su, M., Chen, T., Jia, W., *J. Proteome Res.* 2009, *8*, 4844–4850.

[16] Dunn, W. B., Broadhurst, D. I., Deepak, S. M., Buch, M. H., McDowell, G., Spasic, I., Ellis, D. I., Brooks, N., Kell, D. B., Neyses, L., *Metabolomics* 2007, *3*, 413–426.

[17] Wu, H., Xue, R., Dong, L., Liu, T., Deng, C., Zeng, H., Shen, X., *Anal. Chim. Acta* 2009, *648*, 98–104.

[18] Bathe, O. F., Shaykhutdinov, R., Weljie, A., McKay, A., Sutherland, F., Dixon, E., Vogel, H., ASCO Gastrointestinal Cancers Symposium, 2009, abstract ♯126.

[19] Beger, R. D., Schnackenberg, L. K., Holland, R. D., Li, D., Dragan, Y., *Metabolomics.* 2006, *2*, 125–134.

[20] Urayama, S., Zou, W., Brooks, K., Tolstikov, V., *Rapid Commun. Mass Spectrom.* 2010, *24*, 613–620.

[21] Beens, J., Brinkman, U. A., *Analyst* 2005, *130*, 123–127.

[22] Theodoridis, G., Gika, H. G., Wilson, I. D., *Trends Analyt. Chem.* 2008, *27*, 251–260.

[23] Leão Martins, J. M., Gago-Martinez, A., Dabek-Zlotorzynska, E., Aranda-Rodriguez, R., Lawrence, J. F., *J. Sep. Sci.* 2002, *25*, 342–344.

[24] Gago-Martínez, A., Piñeiro, N., Aguete, E. C., Vaquero, E., Nogueiras, M., Leão, J. M., Rodríguez-Vázquez, J. A., Dabek-Zlotorzynska, E., *J. Chromatogr. A* 2003, *992*, 159–168.

[25] Ru, Q. H., Luo, G. A., Fu, Y. R., *J. Chromatogr. A* 2001, *924*, 331–336.

[26] Ye, F. G., Xie, Z. H., Wu, X. P., Lin, X. C., *Talanta* 2006, *69*, 97–102.

[27] Yao, C. Y., Gao, R. Y., Yan, C., *J. Sep. Sci.* 2003, *26*, 37–42.

[28] Xie, G., Su, M., Li, P., Gu, X., Yan, C., Qiu, Y., Li, H., Jia, W., *Electrophoresis* 2007, *28*, 4459–4468.

[29] García-Pérez, I., Vallejo, M., Garcia, A., Legido-Quigley, C., Barbas, C., *J. Chromatogr. A* 2008, *1204*, 130–139.

[30] Barbas, C., Vallejio, M., García, A., Barlow, D., Hanna-Brown, M., *J. Pharm. Biomed. Anal.* 2008, *47*, 388–398.

[31] Soria, A. C., Wright, B., Goodall, D. M., Wilson, J., *Electrophoresis* 2007, *28*, 950–964.

[32] García-Pérez, I., Whitfield, P., Angulo, S., Legido-Quigley, C., Hanna-Brown, M., Barbas, C., *Electrophoresis* 2008, *29*, 3201–3206.

[33] Xie, B., Gong, T., Gao, R., Liu, J., Zuo, J., Wang, X., Zhang, Z., *J. Pharm. Biomed. Anal.* 2009, *49*, 492–497.

[34] Sangster, T., Major, H., Plumb, R., Wilson, A. J., Wilson, I. D., *Analyst* 2006, *131*, 1075–1078.

[35] Michopoulo, F., Lai, L., Gika, H., Theodoridis, G., Wilson, I., *J. Proteome Res.* 2009, *8*, 2114–2121.

[36] Gika, H. G., Theodoridis, G. A., Wingate, J. E., Wilson, I. D., *J. Proteome Res.* 2007, *6*, 3291–3303.

[37] Zelena, E., Dunn, W. B., Broadhurst, D., Francis-McIntyre, S., Carroll, K. M., Begley, P., O'Hagan, S., Knowles, J. D., Halsall, A., Wilson, I. D., Kell, D. B., *Anal. Chem.* 2009, *81*, 1357–1364.

[38] Wold, S., Esbensen, K., Geladi, P., *Chemometr. Intell. Lab. Syst.* 1987, *2*, 37–52.

[39] Centre for Drug Evaluation and Research (CDER), FDA Guidance for Industry, Bioanalytical Method Validation, May 2001.

[40] Van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., Van der werf, M. J., *Biomed. Chromatogr. Genomics* 2006, *7*, 142.

Address：489 Songtao Road, C01, Zhangjiang High-tech Park, Pudong, Shanghai 201203, P.R. China
E-Mail：info@unimicrotech.com，lemonyu@unimicrotech.com
Website：www.unimicrotech.com

