Study on Gender Difference Based on Metabolites in Urine by Ultra High Performance Liquid Chromatography/Time of Flight Mass Spectrometry

LU Guo†† WANG Jiangshan†† ZHAO Xinjie†† KONG Hongwei†† XU Guowang

National Chromatographic R. & A. Center Dalian Institute of Chemical Physics
the Chinese Academy of Sciences Dalian 116023 China

Abstract Metabolites in urine can illustrate the physical condition of an individual as a whole. Ultra high performance liquid chromatography/time of flight mass spectrometry UPLC/TOF-MS is a relative new technique for the separation of complex samples. The aim of this study is to assess the feasibility of metabolomics in gender difference in unrestricted conditions i.e. for healthy volunteers there are no strict controls such as food life style and the collection of urine samples. In this work 31 spontaneous urine samples were collected and analyzed by using UPLC/TOF-MS. Principal components analysis PCA and partial least squares discriminant analysis PLS-DA models were tested and compared in samples classification. The gender discrimination was highly improved and some gender related biomarkers were found by PLS-DA. These preliminary results suggested that UPLC/MS-based approaches coupled with pattern recognition show promise for metabolomics.

Key words metabolomics ultra high performance liquid chromatography UPLC mass spectrometry MS urine gender difference partial least squares discriminant analysis PLS-DA

CLC number O658 Document code A Article IC 1000-8712 200602-0109-05 Column Communications

Metabonomics is the identification and measurement of metabolic profile dynamics of holistic changes as the result of exposure to a toxin or drug to environmental effects or the onset of diseases. However, it is well known that physiological and some extrinsic differences such as food age and gender still have influences on the compositions of urine and blood. In order to be confident of perturbed metabolic profiles in such cases it is necessary to characterize the normal state of physiological variation without stimuli and define the boundaries of normal metabolism. Gender difference has been shown to be a consideration in drug development because the gender-related difference might lead to different pharmacokinetics/pharmacodynamics. Generally, in work carried out in gender difference for rats male rats have a greater metabolic activity than the females which results in a prolonged pharmacological activity for many drugs in female rats.

Received date 2005-11-22
First author†† LU Guo female. Corresponding author†† XU Guowang PhD professor Tel/Fax 041184379559 E-mail xugw@ dicp. ac. cn. Foundation itemSupported by the National Natural Science Foundation of China No. 90209048 the Knowledge Innovation Program of the Chinese Academy of Sciences K2002A11 KGCX2-SW-213.
compared with the males\textsuperscript{[20]}. Even so\textsuperscript{[3]} a report showed males were usually more susceptible to toxins\textsuperscript{[19]}. Studies have shown that the biochemical profiles of female urine samples differ significantly from those of males. For example\textsuperscript{[4]}, sex-related difference in the elimination of citrate in the urine has been determined and citrate is generally greater for the females\textsuperscript{[4,50]}. To know the influence of a given physiological factor on metabolite profile\textsuperscript{[6]} the best way is to get the other unrelated factors under control. For example\textsuperscript{[7]} to investigate the influence of gender difference\textsuperscript{[8]} we should get other factors such as food\textsuperscript{[9]}, light\textsuperscript{[10]} and temperature under control. For animals\textsuperscript{[11]} it is easy to perform a strict control experiment but for human beings\textsuperscript{[12]} there will be another case\textsuperscript{[13]} even for urine sample collection\textsuperscript{[14]} there will be some difficulty to collect 24 h urine samples to avoid the diurnal effects. In this case\textsuperscript{[15]} the metabolite profile of healthy human beings will contain many kinds of information including genetic difference\textsuperscript{[16]} gender difference\textsuperscript{[17]} age-related difference\textsuperscript{[18]} diet influence \textit{etc} and there will be difficult to find a distinct difference of metabolite profile caused by a given factor. In this work\textsuperscript{[19]} our aim was to screen out the information related to gender difference from the complex data provided by metabolite profiles of healthy human being urine samples. So there will be no restrict control on the volunteers\textsuperscript{[20]} the criterion for selecting volunteers was healthy without the considerations on food\textsuperscript{[21]} genetics and age \textit{etc}. For urine sample collection\textsuperscript{[22]} the protocol was to collect spontaneous urine without the consideration to collect 24 h or first void urine samples. Liquid chromatography/mass spectrometry \textsuperscript{[23]} LC/MS \textsuperscript{[24]} coupled with pattern recognition has been proved to be an effective method for metabolite detection\textsuperscript{[25]} metabolite profiling and data interpretation\textsuperscript{[26-100]}. Ultra high performance liquid chromatography\textsuperscript{[27]} UPLC\textsuperscript{[28]} employs 1.7 \textmu m porous particles resulting in higher peak capacity\textsuperscript{[29]} better resolution and sensitivity compared with traditional HPLC\textsuperscript{[30]} and has been applied in many researches\textsuperscript{[31-140]}. Since the data generated in metabolic fingerprint experiments are multisubject and multivariate\textsuperscript{[31]} it is necessary to utilize a wide range of statistical data reduction\textsuperscript{[32]} multivariate analysis\textsuperscript{[33]} such as principal components analysis\textsuperscript{[34]} PCA\textsuperscript{[35]} and partial least squares-discriminant analysis\textsuperscript{[36]} PLS-DA\textsuperscript{[37]}.

In this work\textsuperscript{[38]} we coupled UPLC/MS technology with multivariate statistical analysis in order to screen out gender difference related information from the data of metabolites profile. Two data interpretation methods\textsuperscript{[39]} unsupervised PCA and supervised PLS-DA mathematical models\textsuperscript{[40]} were used and the results of sample classification were compared and discussed.

1 Experimental

1.1 Chemicals

Acetonitrile was high performance liquid chromatographic\textsuperscript{[41]} HPLC\textsuperscript{[42]} grade\textsuperscript{[43]} Merck\textsuperscript{[44]} Germany\textsuperscript{[45]}. Formic acid was HPLC grade\textsuperscript{[46]} TEDIA\textsuperscript{[47]} USA\textsuperscript{[48]}. Potassium dihydrogenphosphate\textsuperscript{[49]} KH\textsubscript{2}PO\textsubscript{4} and all other chemicals were analytical reagents. Water was produced by Milli-Q ultra-pure water system\textsuperscript{[50]} Millipore\textsuperscript{[51]} Billerica\textsuperscript{[52]} USA\textsuperscript{[53]}.

1.2 Sample collection and preparation

The design of this study was to let volunteers in their natural situation to assess the feasibility of a metabolomics approach for gender difference under unrestricted conditions. Thirty-one spontaneous urine samples\textsuperscript{[54]} 13 males and 18 females with age range of 25 – 66 years\textsuperscript{[55]} were collected from healthy adult volunteers from the authors\textsuperscript{[56]} institute and there were no strict control on diet and lifestyle but excessive intake of alcohol was prohibited. After collection\textsuperscript{[57]} the samples were frozen immediately and stored at \textminus20 °C. Prior to analysis\textsuperscript{[58]} the samples were thawed at room temperature and were centrifuged under 13 000 r/min for 5 min\textsuperscript{[59]} then the supernatant was diluted five-fold with water and 5 \textmu L were analyzed by using UPLC/MS.

For spontaneous urine samples\textsuperscript{[60]} the concentrations of metabolites differ significantly due to the urine volume. To calibrate the difference of urine volume\textsuperscript{[61]} the levels of urinary creatinine\textsuperscript{[62]} a metabolite with a relative constant amount in urine\textsuperscript{[63]} were determined as described by Zheng \textit{et al} \textsuperscript{[150]}.

1.3 Chromatography

Chromatographic separations were performed on an Acquity C18 column\textsuperscript{[64]} 100 mm × 2.1 mm i. d. \textsuperscript{[65]} 1.7 \textmu m\textsuperscript{[66]} Waters Corp\textsuperscript{[67]} Milford\textsuperscript{[68]} USA\textsuperscript{[69]} using an Acquity ultra high performance liquid chromatography system\textsuperscript{[70]} Waters\textsuperscript{[71]} equipped with a
tunable ultraviolet UV detector set at 254 nm. The total flow rate was 0.25 mL/min. The column was maintained at 25 °C and the gradient program was started from 100% A and held for 0.3 min then changed to 100% B within 9 min and held at 100% B for 1 min then back to 100% A and held for 0.9 min solvent A aqueous solution of 0.1% formic acid solvent B acetonitrile. The eluate was directed to the mass spectrometer without split.

1.4 Mass spectrometry

Mass spectrometry was performed on a Micromass Q-TOF Waters MS Technologies Manchester UK orthogonal acceleration time of flight mass spectrometer equipped with an electrospray ionization source ESI operating in positive ion mode. The desolvation gas rate was set to 500 L/h at a temperature of 300 °C the cone gas rate was set at 25 L/h and the source temperature at 120 °C. The capillary voltage and the cone voltage were set at 3 200 V and 35 V respectively. Data were collected in centroid mode from m/z 50 – m/z 850 with a lock spray frequency of 5 s and data averaging over 10 scans.

1.5 Data collection

For each analysis a chromatogram which contains the information of retention times and related mass spectra can be obtained. The raw data were analyzed using the Micromass MarkerLynx Applications Manager version 4.0. MarkerLynx incorporates an ApexTrack-peak detection package which allows detection and retention time alignment of the peaks eluting in each chromatogram. The data were combined into a single matrix by aligning peaks with the same mass/retention time pair together from each data file in the dataset along with their associated intensities. Then these peak intensities were normalized to urinary creatinine. The resulting three-dimensional data peak number t m/z pair sample name and ion intensity normalized by creatinine levels were analyzed by PCA and PLS-DA SIMCA-P software version 10.0 Sweden. Mean centered Ctr was used for data scaling and centering.

2 Results and discussion

The aim of analytical technique used in metabolomics is to provide comprehensive quantitative and qualitative information of endogenous metabolites in biological samples. In the UPLC/TOF-MS analysis even co-eluted metabolites will be separated based on the difference of relative molecular masses by mass spectrometry. The characters of UPLC/TOF-MS enable the application of a fast gradient chromatographic separation without significant loss of information and make it a comprehensive and high throughput analytical technique in metabolomics. In this paper a 10 min fast gradient was applied to increase throughput. There are at least several thousands peaks related to metabolites in urine can be detected. A representative base peak intensity BPI chromatogram for a female urine sample is presented in Fig. 1. A set of 31 urine samples were analyzed using positive ion mode. In order to find out gender difference from the complex information provided by metabolite profile PCA and PLS-DA models were applied for analysis of these data. After normalized by the corresponding creatinine levels the variables from UPLC/MS spectrum were set as columns of a matrix and 31 samples were arranged in rows of the matrix. Then PCA
and PLS-DA analyses were performed on the matrix.

PCA is a well-known pattern recognition method in that most of the variance within a data set can be expressed by a small number of principal components PCs. Each PC is a linear combination of the original data parameters and independent of the other PCs. As an unsupervised method, PCA is usually used to find out whether any intrinsic clustering exists within a data set. The output of PCA results can generate two matrices named as scores and loadings. Scores are the co-ordinates for the samples in the established model. Each point in the score plot represents a single sample. The PC loadings define the way in which the original variables are linearly combined to form the new variables. In the loading plot each point represents a different chromatographic peak labeled by the given retention time and mass/charge ratio m/z.

Firstly, it is necessary to know whether the gender difference will lead to a distinct clustering of the data set. Those variables obtained from 31 urine samples were performed with PCA. From the score plot of PCA, it is not difficult to conclude that the PCA could not classify the samples according to gender difference. Moreover, the clustering results cannot be explained by age difference and sampling time difference. These results suggested that without a strict control protocol the influence of gender factor on urine metabolite profile will be obscured by the other factors and in the natural state there is not a predominate factor that can influence the pattern of metabolites profile. Simple unsupervised chemometrics methods obviously worked well for the data set with a limited number of well-defined classes. Because biological systems are seldom simple and many of the bio-fluid data sets generated within metabolomics require more sophisticated statistical data analysis. Consequently various adaptations of principal components partial least squares- and neural network-based methods can be used to optimize the classification. As a supervised method, PLS-DA is the combination of PLs method and DA method. This method can establish the optimal position to place a discriminant surface that separates classes best and generate the matrices of scores and loadings. In the present study, PLS-DA was used to investigate the gender difference to improve the class separation. A distinct separation of two groups was obtained with only one female sample entering the area of male. The loading plot

![Image of PCA and PLS-DA plots](image_url)

Fig. 2 a Score plot from PCA PC1 vs PC2 b score plot from PLS-DA 1 vs 2 c loading plot from PLS-DA models and the biomarkers labeled with mass/charge ratios and retention times triangle male square female.
Fig. 2-c indicates the most influential ions that are responsible for the separation between sample classes. That is, the ions having the greatest influence on the PLS-DA score plot are those furthest away from the main cluster of ions. In turn, these compounds might be the candidates for biomarkers. Some potential biomarkers are labeled in Fig. 2-c with mass/charge ratios and retention times. It should be mentioned here is that two compounds have the same retention time. There may be two causes that can account for this kind of result. Firstly, although ESI is a relatively soft ionization method in comparing with the electron impact ionization EI for many compounds the insource fragments still exist. In this case an ionized molecule will produce some fragment ions in ion source and not only the ionized molecule but also its fragments will reach the detector resulting two or more peaks with the same retention time. Secondly, some metabolites cannot be separated by the column and effuse in the same time. Fortunately, in both cases, the different metabolites with the same retention time or the insource fragments will be separated by MS based on the m/z difference.

The average levels of four potential biomarkers in male and female urine samples are shown in Fig. 3. The average level of potential biomarker 1 in male samples is higher than that in females. However, the average levels of potential biomarkers 2, 3, and 4 in females are far higher than those in males.

3 Conclusions

In this paper the UPLC/TOF-MS and multivariate statistical analysis have been used for urinary metabolite profiling and data interpretation. A rapid gradient chromatographic separation method was established. Both unsupervised PCA and supervised PLS-DA have been applied to sample classification. PLS-DA improves gender separation very well and can screen out some biomarkers that may account for gender difference. The results demonstrate that UPLC/MS combined with PLS-DA enables the discrimination of gender difference based on the metabolites of urine and suggest that UPLC/MS-based approaches coupled with pattern recognition may well provide useful insights into the biochemical differences between organisms and show promise for metabolomics.

References