A Preliminary Study of Oral Squamous Cell Carcinoma Using ^{1}H NMR-based Metabonomic Technique

Jing Lin ZHOU[1], Ju Kun SONG[1], Jing HUANG[2], Zheng Hong MA[2], Jing XUE[1], Li Ying XIAO[1], Wei LI[1]

Objective: *To investigate the differences in ^{1}H nuclear magnetic resonance (^{1}H NMR)-based metabonomic analysis between oral squamous cell carcinoma (OSCC) and healthy control by principal component analysis (PCA); also, to evaluate this method in distinguishing different sample classes, and to detect the relationships between genotypes and phenotypes.*

Methods: *^{1}H NMR resonances of plasmas were obtained from 20 OSSC subjects as well as 20 healthy subjects, and spectrally processed. Subsequent data were reduced by applying unsupervised PCA, one method of pattern recognition analysis, to detect the different spots corresponding to the low-molecular-mass metabolites.*

Results: *Some important micromolecular species in plasma, including myo-inositol, praline, creatinine, arginine, and aspartic acid displayed great dissimilarity between OSCC and healthy control.*

Conclusion: *The NMR metabonomic technique is feasible for studying the small molecule components of human blood plasma.*

Key words: *^{1}H NMR spectroscopy, metabonomics, oral squamous cell carcinoma*

Oral squamous cell carcinoma (OSCC) is one of the six most frequently occurring cancers[1], and one of the three most common malignancies in developing countries[2,3]. Surgical removal of the tumour usually results in facial distortion. Prevention and early detection or treatment of OSCC might significantly improve life quality for individuals at risk. Therefore, an accurate and non-invasive screening test for early detecton OSCC is clinically necessary.

Human blood plasma is the major place where metabolites are circulated around the body. Chemical analysis of plasma can provide a great amount of information at the molecular level relating to the biochemical status of an individual, which is very important for diagnostic purposes[4]. Further evaluation of plasma proteins associated with OSCC may be useful for further development of a panel of tumour markers that can aid the early detection of OSCC, potentially saving lives.

Genomics, transcriptomics, proteomics, and metabonomics currently constitute systems biology. Recently, metabonomic analysis of human biofluids has developed rapidly in clinical research[5]. The application of high-throughput genomic and proteomic platforms for global analysis of the production of genetically modified organisms is increasingly available. Multivariate datasets are being produced from the enormous response. ^{1}H nuclear magnetic resonance (^{1}H NMR)-based metabonomics and pattern recognition analysis (PRA) play an important role in this field. They can be used for description and recognition of phenotype changes. Metabolic

1 Key Lab of Oral Biomedical Engineering of Ministry of Education, Sichuan University, Chengdu, P.R. China.

2 West China College of Pharmacy, Sichuan University. Chengdu, P.R. China.

Corresponding author: Dr Wei LI, Key Lab of Oral Biomedical Engineering of Ministry of Education, Sichuan University, Chengdu, 610041, P.R. China. E-mail: leewei2000@sina.com

maps provide a framework for studying the consequences of genotype changes and the relationships between genotypes and phenotypes.

For almost two decades, ^{1}H NMR spectroscopy has been used as an "open" system to study the temporal changes in the biochemical composition of biofluids, in response to adverse toxic events. Many of these *in vivo* studies have reported changes in individual metabolites correlating with toxicological changes. ^{1}H NMR can be used to analyse chemical structures at the molecular level and provide the global profiling of biological samples as the characteristic metabolic fingerprint. Principal components analysis (PCA) is one of the 'unsupervised methods' and a suitable chemometric analysis of information from the data of resonances. It can reduce the complexity and volume of information content down to a suitable level[6] and might be a reliable method for screening for biomarkers of organ- or tissue-specific chemically-induced lesions. The present study aimed to investigate the differences in metabolites between OSCC and healthy subjects with ^{1}H NMR.

Materials and Methods

Subjects

The subjects were inpatients of the Hospital of West College of Stomatology, Sichuan University. The study was consistent with the ethical principles of Good Clinical Practice and the Declaration of Helsinki. All the subjects provided written informed consent, and were required to receive standard clinical laboratory examinations, such as haematology assessments, hepatitis B surface antigen, hepatitis C and HIV antibody determination, breast X-ray, CT, and cardiogram. Exclusion criteria were: 1) abnormalities in clinical chemistry or haematology; 2) hepatitis B or C; 3) AIDS; 4) acute or chronic pneumonia; 5) any sorts of heart disease; 6) any sorts of inborn disease; 7) drug or alcohol dependence. In addition, OSCC subjects were excluded if they were: 1) metastasised OSCC; 2) terminal OSCC; 3) received chemotherapy and radiotherapy. Twenty healthy subjects (48–69 years old) and 20 OSCC subjects (42–70 years old) were selected.

All subjects abstained from eating for 12 hours and avoided hazardous physical activities before blood samples were collected. All samples were collected at approximately 9 a.m at room temperature.

Samples

Blood samples were collected into tubes containing lithium heparin. The blood cells were separated from plasma by centrifugation. All the plasma samples were stored at -20°C prior to 1H NMR analysis.

^{1}H NMR spectroscopy

All the plasma samples were freeze-dried before analysis and reconstituted into 500 µl D_2O (1 mg/ml). To obtain a deuterium lock signal for the NMR spectrometer, ^{1}H NMR spectra were obtained immediately after thawing of each sample at 300 K with a Bruker Avance DRX600 spectrometer (Bruker, Coventry, UK), at 600.13 MHz with a 5-mm broad-band-inverse (BBI) probe and 5-mm NMR tube. The chemical shift was referenced to the solvent since TSP is not a suitable reference in plasma samples[7]. Spectra were achieved by the use of a selective presaturation pulse sequence (Bruker Analytik, Rheinstetten, Germany) for water suppression (δH = 4.7 ppm). For each sample, ^{1}D ^{1}H NMR spectra were collected into 131072 data points over a spectral width of 7788.16211 Hz.

Data-reduction of NMR spectra

All ^{1}H NMR spectra (δ10–0 ppm) were manually phased and corrected by baseline distortion into digital values using MestReC (version 4.4.1.0). Each spectrum was divided into 199 contiguous segments between 10 and 0 parts (ppm) from the spectral region and all segments were integrated. Each segment width of 0.04 ppm was chosen. The region of spectral 6.5–4.5 ppm was assigned a zero integral value in order to exclude the effect of water signal. The data were normalised to total spectral area in Excel (Microsoft, USA), and then exported to the SIMCA-P software package (version 11.0.0.0, Umetrics, Sweden).

Statistical analysis

PCA is a bilinear decomposition method that can be used to check these data and reduce the complexity and volume of information content down to a suitable level. PCA models were performed using SMICA software. The data described above were input into the SIMCA-P software for statistical analysis.

Results

^{1}H NMR spectra of plasma from OSCC and healthy subjects

^{1}H NMR spectra of plasma demonstrated great complexity of information as shown in Fig 1. The region of the aromatic (δ 6.0–8.5 ppm) had no important metabol-

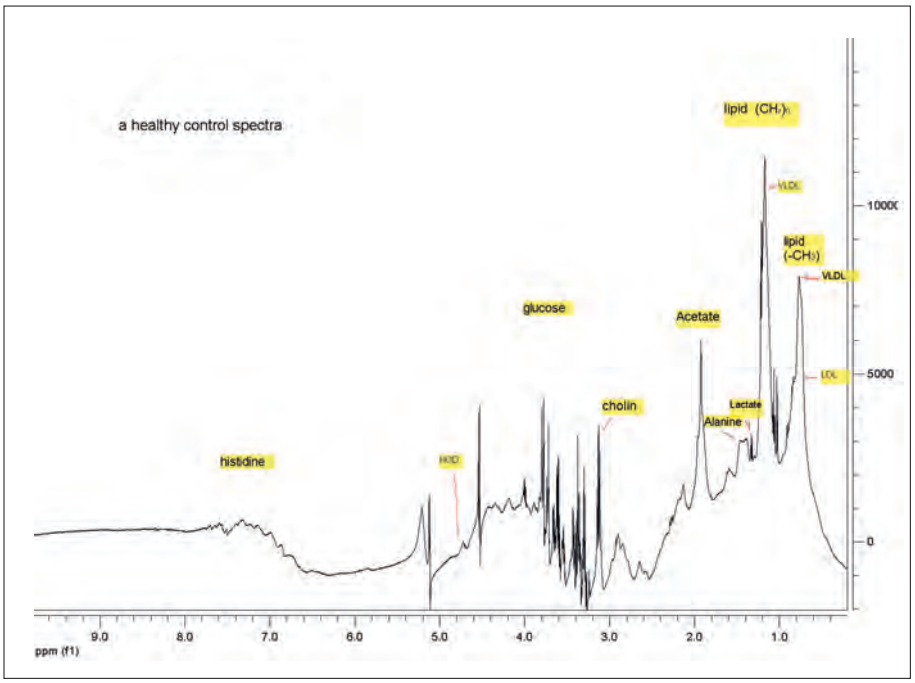

Fig 1 600 Hz 1D ^{1}H NMR spectra for the plasma obtained from a healthy subject. HDL, high-density lipoprotein; LDL, low-density lipoprotein; VLDL, very-low-density lipoprotein; HOD, water or D_2O signal.

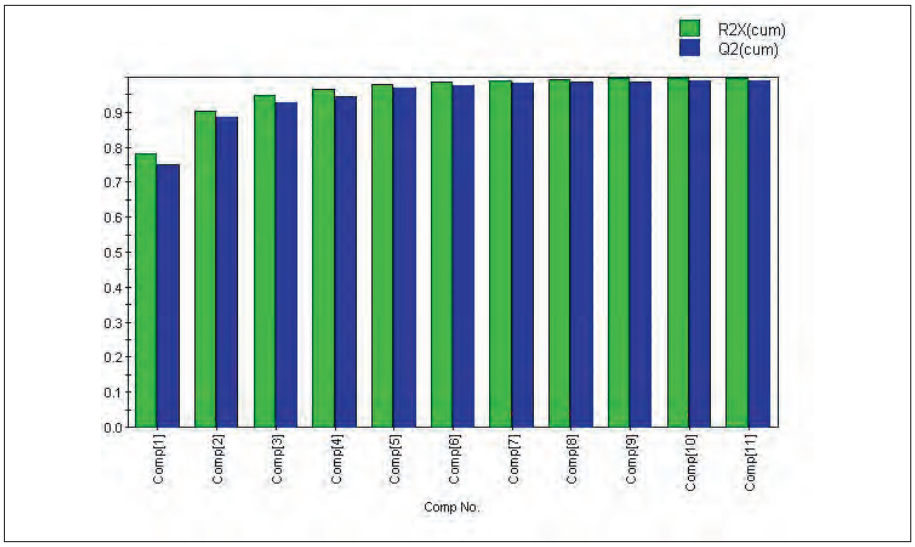

Fig 2 The cumulation of new principal components. R2X (green bars) is the percentage of variation of the training set explained by the model; Q2 (blue bars) is the percentage of the variation of the training set that can be predicted by the model.

ic signals. The chemical shift of δ 4.7 ppm was the water signal. The metabolite signals from δ 0 ppm to δ 4.5 ppm might be important region, containing the region of the alkene (δ 4–5.5 ppm), the alkyne (δ 1.8–3.5 ppm) and the aliphatic (δ 0-2.8 ppm). The 1H NMR spectra of plasma could display the fingerprint of metabolites.

PCA of ^{1}H NMR spectra of plasma from OSCC and healthy subjects

All the ^{1}H NMR spectra data were normalised in Excel with appropriate weighting coefficients and input into SIMCA-P software. The principal components (PC) variables were created, and each PC was orthogonal with all other PCs (Fig 2). Eleven PCs were obtained to represent the original data. The first PC accounted for 78.1% of the variance and the second PC 12.3%; the

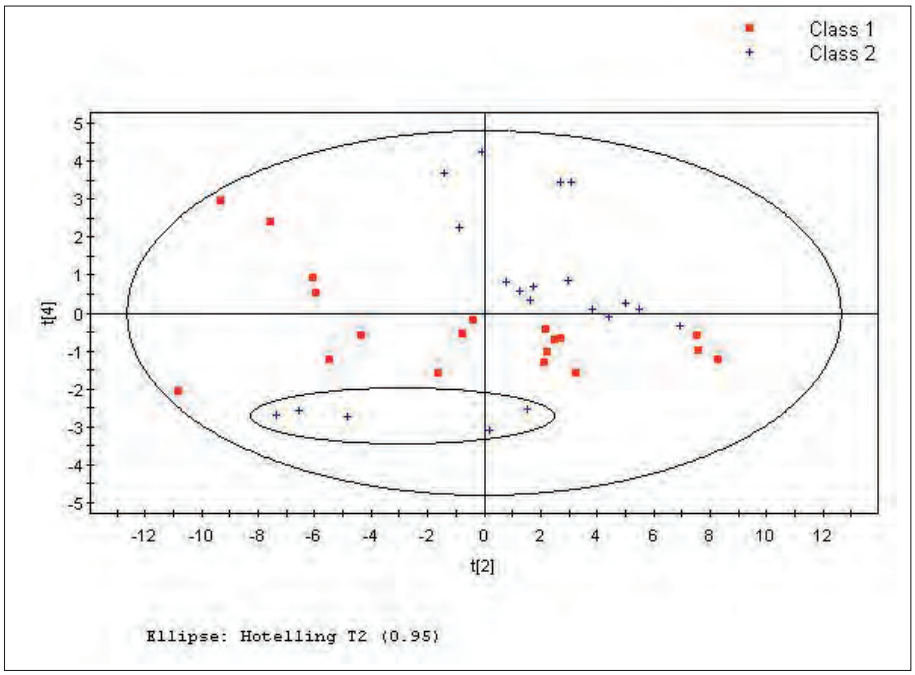

Fig 3 Scores plot result from principal component analysis clearly illustrating separation among the different sample classes.

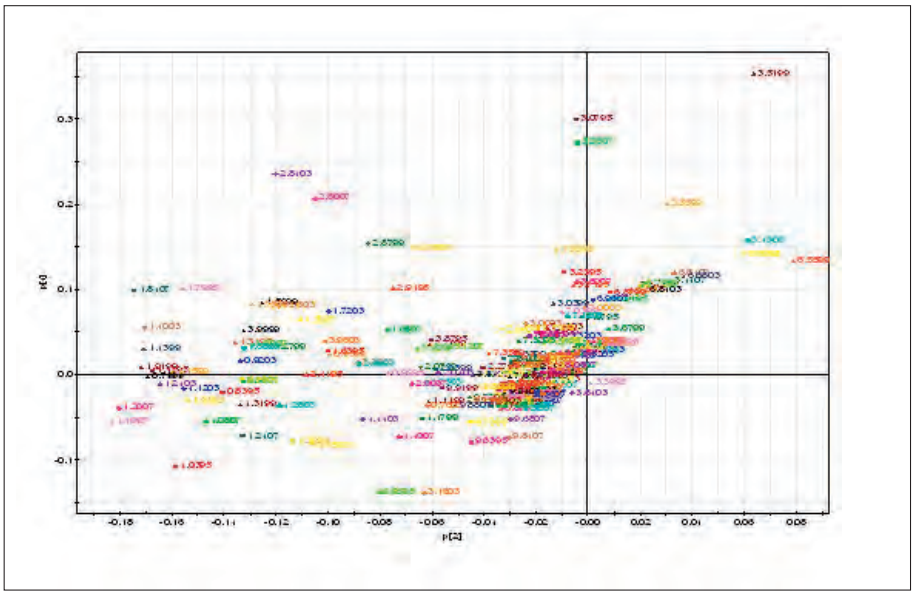

Fig 4 The loadings plots from the contributions of the dataset to the PCs. Numbers illustrate chemical shifts of compounds responsible for variation along p2 and p4.

third PC and the rest accounted for 9.6%. It was indicated that the first two PCs provided the maximum information content of the data from ^{1}H NMR spectra (Fig 2). The first component (t1) explains the largest variation of the space, followed by the second component (t2).

PCA analysis of the spectra data showed all of the scores were inside the 95% confidence ellipse based on Hotelling T2. Separation of OSCC subjects (red dots) from healthy subjects (blue crosses) is shown in Fig 3. The points for healthy subjects were displayed in relatively tight clusters while those for the OSCC subjects were spread out. A group of five blue crosses for the controls could not be distinguished from OSCC, but they were still different from OSCC. The disease histories of the five controls revealed that they were odontogenic cyst patients. Therefore, the PCA method could well have revealed the character of each group.

The eigenvector coefficients were termed loadings that comprise the loadings matrix, and were given the contributions of the dataset to the PCs. The scores

weighted averages of the variables with weights p2 in the first dimension and p4 in the second dimension (Fig 4). The loadings plot displayed the statistics of integrated data from NMR spectra corresponding to the chemical shift of the molecules. As shown in Fig 4, dense spots were located in one group; some spots were distant from the scope of the group. Considering the separation of OSCC from the controls in Fig 3, the distant spots in Fig 4 might mainly reflect the difference between OSCC and the controls data sets. These different chemical shifts were responsible for separation of OSCC plasma from the controls. For instance, chemical shift values 3.5199 ppm, 3.3599 ppm, 3.0785 ppm, 3.2807 ppm, 2.8403 ppm and 2.8007 ppm may correspond to the protons of myo-inositol, proline, creatinine, arginine, aspartic acid, as they are dispersed in the plot, and glucose, perhaps illustrating the decreased anabolism of glucose in tumour tissues, and may be directly related to glycosylated proteolipids.

Discussion

This study preliminarily showed that ^{1}H NMR-based metabonomic technique to analyse plasma of humans was capable of distinguishing between OSCC subjects and healthy controls. Furthermore, it was capable of elaborating the global profiling of human plasma. With multivariate statistical analysis, some substances revealed the characteristics of malignant tumours. However, plasma is very complex, in physico-chemical terms, in lipoprotein, proteins, small organic molecules, and ions. They have a variety of possible molecular interactions, such as chemical exchange processes including metal complexation, enzyme-mediated biotransformations, and small-molecule–macromolecular binding. Any abnormal metabolites of plasma, especially low-molecular-mass metabolites and some important macromolecular species, would also be revealed. Further research should study the relationships between genotype and phenotype of OSCC. It is necessary to develop an effective blood screening clinical technique for early OSCC detection.

The ^{1}H NMR-based metabonomic technique, combined with genomic findings, might be of appreciable value in studying the molecular properties of abnormal lipoproteins extracted from plasma. It can be expected to be useful in the diagnosis of human cancers in future.

References

1. Melrose RJ. Premalignant oral mucosal diseases. J Calif Dent Assoc 2001;29:593-600.
2. Ko YC, Huang YL, Lee CH et al. Betel quid chewing, cigarette smoking and alcohol consumption related to oral cancer in Taiwan. J Oral Pathol Med 1995;24:450-453.
3. Chen CL, Chi CW, Liu TY. Hydroxyl radical formation and oxidative DNA damage induced by areca quid in vivo. J Toxicol Environ Health A 2002;65:327-336.
4. Nicholson JK, Wilson ID. High resolution proton magnetic resonance spectroscopy of biological fluids. Prog NMR Spectrosc 1989;21:449-501.
5. Lindon JC, Holmes E, Nicholson JK. So what's the deal with metabonomics? Anal Chem 2003;75:384A-391A.
6. Lindon J, Holmes E, Nicholson J. Pattern recognition methods and applications in biomedical magnetic resonance. Prog NMR Spectrosc 2001;39:1-40.
7. Tang H, Wang Y, Nicholson JK, Lindon JC. Use of relaxation-edited one-dimensional and two dimensional nuclear magnetic resonance spectroscopy to improve detection of small metabolites in blood plasma. Anal Biochem 2004;325:260-272.