Applied Surface Science 255 (2008) 1084–1087

Contents lists available at ScienceDirect

Applied Surface Science

journal homepage: www.elsevier.com/locate/apsusc

ToF-SIMS PC-DFA analysis of prostate cancer cell lines

M.J. Baker [a], E. Gazi [b], M.D. Brown [b], N.W. Clarke [b], J.C. Vickerman [a], N.P. Lockyer [a,*]

[a] Centre for Instrumentation and Analytical Science, Manchester Interdisciplinary Biocentre, University of Manchester, Manchester M1 7DN, UK
[b] The Paterson Institute for Cancer Research, University of Manchester, Christie Hospital NHS Trust, Wilmslow Road, Manchester M20 4BX, UK

ARTICLE INFO

Article history:
Available online 18 May 2008

Keywords:
Cancer
Cell lines
Multivariate analysis
Discrimination
C_{60}^+

ABSTRACT

Three closely related cancer cell lines have been analysed with ToF-SIMS using a C_{60}^+ primary ion beam. Principal component-discriminant function analysis (PC-DFA) has been applied for spectral classification. Various spectral pre-processing methods are discussed and assessed for optimum discrimination of this data set. The sum-normalised PC-DFA spectral model produced sensitivities as high as 83.3% and specificities as high as 100% at the 99% confidence limit. At this confidence limit only one errant spectrum was misclassified. The resulting loadings plots suggest that a range of lipid and amino-acid related signals are responsible for the cell line discrimination.

1. Introduction

Prostate cancer (CaP) is the most common cancer and second most common cause of cancer-related death of men in the United Kingdom. Predicting clinical outcome and therefore prescribing the most effective treatment regime is particularly challenging for this disease. Early diagnosis has increased the percentage of patients who survive more than 5 years after their diagnosis [1]. Confirmation of initial diagnosis involves a biopsy and use of the Gleason grading system based on glandular architecture [2]. This histological methodology is subject to inter-observer variability, which limits its use in clinical decision-making [3]. New diagnostic tools based on an understanding of the underlying biochemistry are required. Recently advances have been made in the application of spectroscopic and spectrometric techniques to cancer research [4–7]. These techniques can supply information on carcinogenesis at the molecular level. Molecular differentiation of cell lines and tissue may aid development of new bio-spectroscopic diagnostic tools and drug targets for the identification and treatment of cancer.

Previous work in the area of cancer discrimination using time-of-flight-secondary ion mass spectrometry (ToF-SIMS) was performed by Kulp et al. [6] for the discrimination of three human breast cancer cell lines. The authors report the discrimination by using lysed cells and a mass selection of *m/z* 58–500 with peaks attributable to gallium and PDMS being removed. Previous work on prostate cancer has demonstrated that ToF-SIMS imaging of elemental and molecular ions provides a basis for new insights into specific biochemical compositions of the PC-3 cell line [8]. This study aimed to determine whether ToF-SIMS could discriminate between CaP cell lines based upon biomolecular differences in surface chemistry. The cell lines used in this study were LNCaP (prostate cancer cells derived from lymph node metastasis), non-malignant PNT2 (transfected normal prostate epithelial cells with genome of SV40 virus) and PC-3 (prostate cancer cells derived from bone metastasis). Principal component-discriminant function analysis (PC-DFA) was used to probe the molecular differentiation and construct a model to discriminate between cell lines. This approach reported here uses ions in the mass range *m/z* 1–1500 and freeze-dried cells that are structurally intact as determined by secondary electron microscopy.

2. Experimental

2.1. Cell culture

Cells were cultured on 5 mm × 5 mm silicon substrates (Agar Scientific, UK), sterilised in 70% ethanol for 1 h. All three cell lines were cultured in the same medium, Ham's F-12, 7% FCS and 2 mM L-glutamine to ensure that the chemometric classification was not influenced by different growth media. Cultures were grown at 37 °C in a humidified atmosphere of 5% CO_2 in air. All reagents were purchased from Sigma–Aldrich (Poole, UK). All tissue culture media were obtained from Invitrogen (Paisley, UK). After culture to 80% confluence the cells were removed from their growth media and washed for 3 s in phosphate buffer solution (PBS) and for 60 s

* Corresponding author.
E-mail address: M.J.Baker@manchester.ac.uk (N.P. Lockyer).

0169-4332/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.apsusc.2008.05.256

in deionised water prior to flash freezing in liquid nitrogen cooled methyl-2-butane (Sigma–Aldrich, UK). The samples were then freeze-dried for 24 h at 1×10^{-3} mbar.

2.2. ToF-SIMS analysis

ToF-SIMS analysis was performed on a BioToF-SIMS system [9] with a 40-keV C_{60}^{+} primary ion beam system (Ionoptika Ltd., UK). The samples were divided into machine replicates consisting of a sample from each cell line. These replicates were analysed on different days to remove any effect of instrument drift on spectral classification. Fifteen samples, from 15 different cultures of each cell line were analysed, with three different areas of each sample analysed in positive and negative ion mode. This resulted in a total of 90 spectra from each cell line. Each spectrum was taken over a $500\mu m \times 500\,\mu m$ field-of-view, corresponding to a primary ion dose density of no more than 1×10^{10} cm^{-2}. Charge compensation was employed using 25 eV electrons between ion pulses. Spectra were acquired over a mass range of 1–1500 Da.

2.3. Data processing

To aid multivariate analysis spectral data was binned to $1 \pm 1/2$ Da using software written in house. Each raw data file contains 168 592 data points. Binning the data reduces this to 1500 data points and eases the computational intensity required to perform multivariate analysis on a large data set. Data pre-processing and multivariate analysis was performed using Matlab (Mathworks Inc., MA, USA) and in house software. PC-DFA was used to build a diagnostic model of ToF-SIMS spectra against which a blind test set was projected into and classified according to cell line. PCA is an unsupervised method of identifying patterns in data, and expressing the data in such a way that it highlights similarities and differences [10]. Before the spectra were inputted to the PC-DFA model a quality test was employed based upon the level of sodium in the spectrum. The presence of high levels of residual sodium can influence the relative intensity of other species in the ToF-SIMS analysis through the matrix effect [11]. Wagner et al. discarded spectra where the intensity of the sodium ion peaks was greater than 1% of the total intensity of the selected protein peaks for the adsorbed protein films under analysis. However, a cell contains approximately 13 ± 4 mM sodium [12], therefore for this specific experiment spectra with a sodium ion intensity greater than 10% of the total ion intensity were discarded. This resulted in the model being built from 36 PC-3 spectra, 20 LNCaP spectra and 32 PNT2-C2 spectra for the positive model and 42 spectra from each cell line for the negative model. A fifth of the total data set, randomly selected, was retained to blind test the model and validate the discrimination. Error ellipses with 95% and 99% confidence are added to the discriminant function plots using error_ellipse.m written by AJ Johnson and obtained from Matlab central file exchange. Covariance matrices were calculated from the discriminant function analysis scores matrix for each cell line and the centroid was defined as the mean of the discriminant function analysis scores matrix for each cell line.

Sum-normalisation, autoscaling and mean centring pre-processing steps were each applied and their influence on the accuracy of the multivariate analysis evaluated using measures of sensitivity and specificity. In the context of this work, *sensitivity* is the probability of a positive test among patients with disease—in other words, the ability of the model to diagnose. *Specificity* is the probability of a negative test among patients without the disease—the ability of the model to not misdiagnose.

Table 1
Correctly classified spectra at 99% confidence limit for different pre-processing techniques of the positive and negative ion spectral models

	Sum-normalisation	Autoscaling	Mean centring
Positive model	94.1% (16/17)	47.1% (8/17)	23.5% (4/17)
Negative model	66.7% (18/27)	48.1% (13/27)	40.7% (11/27)

3. Results and discussion

3.1. Pre-processing

Sum normalisation normalises the total area of the spectrum to one and expresses the peaks as a ratio of that sum, mean centring deducts the mean spectrum from each spectrum focusing upon deviations from the mean rather than data as a whole and autoscaling subtracts the mean spectrum from each spectrum and divides the resulting spectrum by the standard deviation. Autoscaling has the effect of increasing the importance of lower intensity peaks in the spectrum and as such it enhances the noise. Table 1 shows the percentage of correctly classified spectra of the positive and negative ion blind test set at the 99% confidence limits for the different pre-processing techniques when PC-DFA was applied. The effectiveness is assessed by the number of spectra correctly classified at the 99% confidence limits of the blind test data set.

The optimum pre-processing technique for the positive and negative spectral model is sum-normalisation, this is the pre-processing technique used for this study.

3.2. Prostate cancer cell line discrimination

The positive spectra model uses 44 principal components and 2 discriminant functions, the negative ion model uses 45 principal components and 2 discriminant functions. The maximum number of discriminant functions available for use is the number of groups minus one. The optimal number of principal components for DFA was determined by assessing the accuracy of the resulting PC-DFA model. Fig. 1 (A) shows a discriminant function projection plot of the positive ion model with the blind test set projected in to test the discriminatory power of the model and (B) a discriminant function plot of the negative ion model with the blind test set projected in to test the discriminatory power of the model

Discriminant function 1 of the positive ion model is discriminating the PC-3 cell line from the PNT2-C2 cell line and the LNCaP cell line, essentially separating bone metastases from lymph node metastases and normal prostatic epithelial cells. This is an important observation, as it shows potential for discriminating prostate cancers based upon their severity. Bone metastatic prostate cancer is directly responsible for considerable mortality and morbidity. Discriminant function 2 of the positive ion model is discriminating the PNT2-C2 cell line from the PC-3 cell line and the LNCaP cell line, discriminating between normal prostatic epithelial cells (non-cancerous) and the metastatic cell lines (cancerous). Clearly this too has important implication for the use of ToF-SIMS to shed new light onto surface chemical processes linked to carcinogensis. Discriminant function 1 of the negative ion model is discriminating the PNT2-C2 cell line from the PC-3 and LNCaP cell lines and discriminant function 2 is separating the PC-3 cell line from the PNT2-C2 and LNCaP cell lines. Table 2(A) shows the sensitivities and specificities for the different cell lines of the positive ion model and Table 2(B) of the negative ion model at the 95% (inner ellipse) and 99% (outer ellipse) confidence limits.

At the 99% confidence limit of the positive spectra model, only one spectrum was misclassified. A spectrum from the PNT2-C2 cell

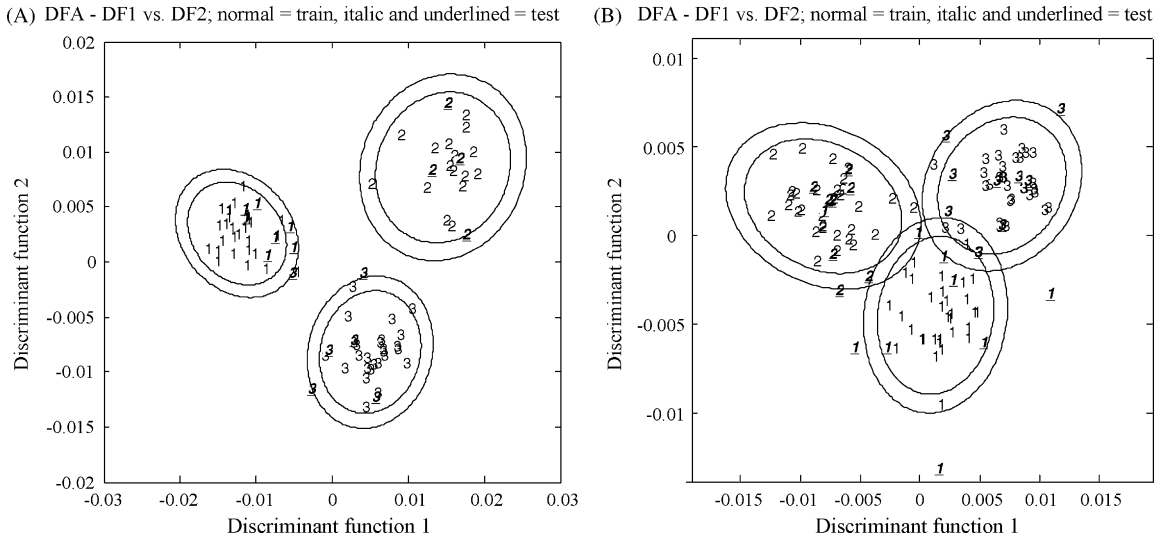

Fig. 1. (A) Discriminant function projection plot of the positive ion model and (B) discriminant function projection plot of the negative ion model with the training set (normal font) and the blind set (italic and underlined) showing confidence limits of 95% (inner ellipse) and 99% (outer ellipse). 1 = PC-3, 2 = LNCaP, 3 = PNT2-C2.

line was classified as a PC-3 spectrum. At the 95% confidence limit no spectra were misclassified. The other spectra lie outside the ellipse and are therefore not counted as being correctly classified, but as they do not lie within the ellipse of any other group they are also not counted as being misclassified. As such the specificity of the model is only different for the PC-3 cell line between confidence limits. The specificity for the model as a whole is high and therefore, it could be said, that the model has the ability to not misdiagnose. The sensitivity differs for the PC-3 and LNCaP cell lines between the confidence limits, with the greatest difference being seen in the PNT2-C2 group. The sensitivity for the model at the 99% confidence limit shows that at this confidence limit the model has the ability to diagnose. The sensitivities and specificities of the negative spectra model are lower than that of the positive spectra model showing that the positive spectra model has greater discriminatory power. However, the complementary ion signals that form the basis of the negative spectral model provide additional information on the underlying biochemical differences.

Each discriminant function has an associated loadings plot which contains the mass spectral ions upon which the model is discriminating. As this is a directed use of ToF-SIMS to analyse the cell surface, the cellular membrane, the mass spectral peaks identified can be assigned to major membrane lipids and amino acids. Loadings plots for cells are difficult to interpret. The cell is a complex biological entity and as such not all the small molecules have been unequivocally assigned using mass spectrometry. Other studies have faced similar problems [6], a focused study with tandem MS would greatly benefit this area of research.

The loadings plots point towards a difference in phospholipid content with the cancerous cell lines discriminating upon phosphocholine or sphingomyelin peaks and the transformed normal cell line discriminating upon phosphoethanolamine. Elevated phosphocholine levels are associated with cellular proliferation an aspect of cancer [13].

Table 2
Sensitivities and specificities at 95 and 99% confidence limits for (A) the positive spectra model and (B) the negative spectra model

Cell line	Confidence limit (%)	Sensitivity (%)	Specificity (%)
A			
PC-3	95	71.4	100
	99	100	90
LNCaP	95	100	100
	99	100	100
PNT2-C2	95	50	100
	99	83.3	100
B			
PC-3	95	44.4	100
	99	44.4	100
LNCaP	95	66.7	94.4
	99	88.9	94.4
PNT2-C2	95	66.7	100
	99	77.8	100

4. Conclusions

This pilot study demonstrates that a multivariate ToF-SIMS methodology is capable of discriminating cancer cell lines based on surface chemical analysis of cryo-fixated cells. A PC-DFA model has been developed that can successfully predict the origin of previously unseen C_{60} ToF-SIMS data based on ions representative of specific biomolecules, including phospholipids and proteins. As well as classifying spectra from cancerous and non-cancerous cell lines, this model accurately discriminates spectra from cancer cell lines of different metastatic sites. This capability is particularly relevant in the context of prostate cancer as the clinically most important decision is often not whether the patient has cancer, but whether a localised disease will metastasis and prove fatal. By providing biochemical information relevant to the metastatic pathway(s), ToF-SIMS may prove a valuable research tool in a number of clinical areas including the development of diagnostic tests, drug targets and patient management.

Acknowledgments

The authors would like to acknowledge the Royal Society of Chemistry, EPSRC Life Sciences Interface Programme, BBSRC and the Association for International Cancer Research for funding. We thank Alex Henderson and Roy Goodacre (The University of Manchester) for use of software routines

References

[1] S. Frankel, Screening for Prostate Cancer, Prostate Cancer IV 361 (2003) 1122–1128.

[2] D.F. Gleason, in: M. Tannenbaum (Ed.), Urological Pathology: The Prostate, Lee and Febiger, Philadelphia, 1977, p. 171.
[3] J.B. Latouff, F. Saad, Brit. J. Urol. Int. 90 (2002) 694–699.
[4] E. Gazi, J. Dwyer, P. Gardner, A. Ghanbari-Siakhali, A.P. Wade, J. Myan, N.P. Lockyer, J.C. Vickerman, N.W. Clarke, J.H. Shanks, C. Hart, M. Brown, J. Pathol. 201 (2003) 99–108.
[5] E. Gazi, M. Baker, J. Dwyer, N.P. Lockyer, P. Gardner, J.H. Shanks, R.S. Reeve, C.A. Hart, N.W. Clarke, M.D. Brown, Eur. Urol. 50 (2006) 750–761.
[6] K.S. Kulp, E.S.F. German, M.G. Knize, D.L. Shattuck, E.J. Nelson, L. Wu, J.L. Montogomery, K.J. Wu, Anal. Chem. 78 (2006) 3651–3658.
[7] P. Crow, B. Barrass, C. Kendall, M. Hart-Pietro, M. Wright, R. Persad, N. Stone, Brit. J. Cancer 92 (2005) 2166–2170.
[8] E. Gazi, N.P. Lockyer, J.C. Vickerman, P. Gardner, J. Dwyer, C.A. Hart, M.D. Brown, N.W. Clarke, J. Miyan, Appl. Surf. Sci. 231–232 (2004) 452–456.
[9] R.M. Braun, P. Blenkinsopp, S.J. Mullock, C. Corlett, K.F. Wiley, J.C. Vickerman, N. Winograd, Rapid Commun. Mass Spectrom. 12 (18) (1998) 1246–1252.
[10] J.E. Jackson, A Users Guide to Principal Components, Wiley, New York, 1991.
[11] M.S. Wagner, B.J. Tyler, D.G. Castner, Anal. Chem. (2002) 1824–1835.
[12] S. Chandra, D.R. Lorey 2nd, Cell. Mol. Biol. 47 (3) (2001) 503–518.
[13] J. Kurhanewicz, D.B. Vingeron, S.J. Nelson, Neoplasia 2 (1–2) (2000) 166–189.