

2015

Raman Micro-Spectroscopy for Rapid Screening of Oral Squamous Cell Carcinoma

Luis Felipe Carvalho

Universidade do Vale do Paraíba, Brazil

Franck Bonnier

Technological University Dublin, Franck.Bonnier@tudublin.ie

Kate O'Callaghan

Trinity College Dublin

See next page for additional authors

Follow this and additional works at: <https://arrow.tudublin.ie/radart>



Part of the [Biophysics Commons](#), and the [Toxicology Commons](#)

Recommended Citation

Carvalho, L., Bonnier, F., O'Callaghan, K., O'Sullivan, J., Flint, S., Byrne, H. & F. Lyng. (2015). Raman Micro-Spectroscopy for Rapid Screening of Oral Squamous Cell Carcinoma. *Experimental Molecular Pathology*, 98(3), 502-509. doi:10.1016/j.yexmp.2015.03.027

This Article is brought to you for free and open access by the Radiation and Environmental Science Centre at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, vera.kilshaw@tudublin.ie.

Funder: Science Foundation Ireland/National Biophotonics and Imaging Platforms

Authors

Luis Felipe Carvalho, Franck Bonnier, Kate O'Callaghan, Jeff O'Sullivan, Stephen Flint, Hugh Byrne, and Fiona Lyng

RAMAN MICRO-SPECTROSCOPY FOR RAPID SCREENING OF ORAL SQUAMOUS CELL CARCINOMA

Luis Felipe CS Carvalho^{a,1*}, Franck Bonnier^{a,2}, Kate O'Callaghan^b, Jeff O'Sullivan^b, Stephen Flint^b, Hugh J. Byrne^a Fiona M. Lyng^{c,d}

^aFOCAS Research Institute, Dublin Institute of Technology, Kevin Street, Dublin 8

^bSchool of Dental Science, Dublin Dental University Hospital, Trinity College Dublin, Dublin 2

^cRadiation and Environmental Science Centre, FOCAS Research Institute, Dublin Institute of Technology, Kevin Street, Dublin 8

^dSchool of Physics, Dublin Institute of Technology, Kevin Street, Dublin 8 Ireland

Email of corresponding author: luisfelipecarvalho@hotmail.com

Abstract

Raman spectroscopy can provide a molecular-level fingerprint of the biochemical composition and structure of cells with excellent spatial resolution and could be useful to monitor changes in composition for dysplasia and early, non-invasive cancer diagnosis (carcinoma in situ), both ex-vivo and in vivo. In this study, we demonstrate this potential by collecting Raman spectra of nucleoli, nuclei and cytoplasm from oral epithelial cancer (SCC-4) and dysplastic (pre-cancerous, DOK) cell lines and from normal oral epithelial primary cell cultures, in vitro, which were then analysed by principal component analysis (PCA) as a multivariate statistical method to discriminate the spectra. Results show significant discrimination between cancer and normal cell lines. Furthermore, the dysplastic and cancer cell lines could be discriminated based on the spectral profiles of the cytoplasmic regions. The principal component loading plot, which elucidates the biochemical features responsible for the discrimination, showed significant contributions of nucleic acid and proteins for nucleolar and nuclear sites and variation in features of lipids for the cytoplasmic area. This technique may provide a rapid screening method and have potential use in the diagnosis of dysplasia and early, non-invasive oral cancer, the treatment of which involves much less extensive and complex surgery and a reduction in associated co-morbidity for the patient.

¹ Present address: Laboratory of Biomedical Vibrational Spectroscopy, Institute of Research and Development, Universidade do Vale do Paraíba (UNIVAP), São José dos Campos-SP, Brazil

² Present address: Université François-Rabelais de Tours, Faculty of Pharmacy, EA 6295 Nanomédicaments et Nanosondes, 31 avenue Monge, 37200 Tours, France

Keywords: Raman Micro-spectroscopy, oral cancer, oral dysplasia, optical biopsy, oral cell lines, Principal Components Analysis

• Corresponding Author: luisfelipecarvalho@hotmail.com

1 INTRODUCTION

Oral cancer is one of the more common cancers worldwide, and in non-developed countries is next in prevalence to skin and breast cancer. Tumours are primarily located around the tongue, floor of the mouth, gingiva and buccal mucosa (Kelly *et al.*, 2014; Cabral *et al.*, 2011). Although clinical diagnosis of late stage, invasive oral cancer is relatively straightforward, the clinical challenge remains to detect early cancerous lesions, which are significantly harder to identify. Surgical treatment of early oral squamous cell carcinoma involves much less extensive and complex surgery and leads to a significant reduction in co-morbidity. Currently, the gold standard for diagnosis is histological analysis of tissue biopsies. However, numerous studies are being conducted in order to improve histopathological techniques with the ultimate goal to obtain a non-invasive, high throughput detection method for screening suspect cancer cells (Deer, 2014).

In this context, optically-based spectroscopic techniques are prominent among potential candidate non-invasive and rapid diagnostic methods. These methods are widely used as a tool for analysis of many biological tissues and the technique has been referred to as an “optical biopsy” (Kallaway *et al.*, 2013; Evers *et al.*, 2012; Nawaz *et al.*, 2011; Carvalho *et al.*, 2010) or “spectral cytopathology” (Schubert *et al.*, 2010) because of its capacity to show features of underlying pathological tissues when compared with normal samples. Raman spectroscopy is one of the most popular techniques with demonstrated potential in studies examining different types of cancer, including lung cancer (Wang *et al.*, 2014), neural cancer (Tanahashi *et al.*, 2014), breast cancer (Marro *et al.*, 2014; Damayanti *et al.*, 2013) cervical cancer (Lyng *et al.*, 2007), gastric cancer (Luo *et al.*, 2013), skin cancer (Mittal *et al.*, 2013) and head and neck cancer (Singh *et al.*, 2012).

The power of the method is primarily related to the molecular and biochemical signature it provides of the biological sample when the fingerprint region is analysed. The Raman effect is a result of inelastic scattering of monochromatic electromagnetic radiation due to interaction with molecular vibrations, such that the frequency of the scattered radiation is different from the incident radiation (Raman, 1928). The Raman spectrum of a given molecule consists of a series of peaks or bands, each corresponding to a characteristic vibrational mode of that molecule. Each molecule exhibits a characteristic spectrum, and thus the Raman spectrum can provide a "fingerprint" of a substance from which the molecular composition can be determined (Nawaz *et al.*, 2011; Lyng *et al.*, 2007). Changes in molecular composition due to the presence of disease are often subtle, requiring the use of multivariate statistical analysis (Byrne *et al.*, 2014). In this way, Raman spectroscopy has been demonstrated to elucidate biochemical signatures not only from cancer studies but from other different pathological conditions in tissues and cells. Carvalho *et al.*, (2010), using Raman spectroscopy to discriminate buccal inflammatory processes and normal tissues, revealed that the discrimination is possible using Soft Independent Modelling of Class Analogy (SIMCA) methods of analysis.

Although there are increased numbers of studies involving Raman spectroscopy as an optical biopsy method, as the technique is based on optical microscopy, it can also be extended to the analysis of cytological samples at a cellular and subcellular level (Bonnier *et al.*, 2010). Thus, interactions of, for example, anticancer agents and nanoparticles can be explored at a subcellular level (Dorney *et al.*, 2012; Nawaz *et al.*, 2011). Of clinical relevance, Raman spectroscopy has been explored for routine cytological screening for cervical cancer (Bonnier *et al.*, 2014, Vargis *et al.* 2012). [However only a few studies have been published](#)

on oral squamous cell carcinoma and Raman micro-spectroscopy (Lasalvia et al., 2015; Su et al., 2012; Guze et al., 2011). The present study demonstrates the potential of Raman spectroscopy to provide and differentiate the biochemical signature from sub-cellular regions of oral cell lines, such as cytoplasm, nucleus and nucleolus, in a label free manner. Also, we illustrate the potential of the technique to discriminate between cancer, pre-cancerous dysplastic cells and normal oral cell lines.

2 MATERIALS AND METHODS

2.1 ORAL CELL LINES

To determine the efficacy of Raman Spectroscopy in distinguishing between the cellular states observed during carcinogenic transformation, three different types of oral cell lines: SCC-4 (malignant cell line), DOK (dysplastic cell line) and Primary cells (normal oral epithelial cell line) were utilised.

2.1.1 SCC-4 and DOK cell cultures. It is important to note that both pathological cell lines were originally from tongue, one of the sites of highest incidence of oral squamous cell carcinoma.

The SCC-4 and DOK (HPA cultures, UK) cell lines were cultured in Dulbecco Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) FBS, penicillin/streptomycin (100U/100ug) and L-glutamine (2mM). For the DOK cell line, hydrocortisone (5ug/ml) was also added. The medium was pre-warmed at 37°C before incubation of the cells. Cells were cultured until 90% confluency in a humidified environment at 5% CO₂, before being passaged. When confluent, the cells were washed in pre-warmed phosphate buffered saline (PBS) (0.01M phosphate buffer, 0.154 M sodium chloride) and incubated with trypsin-EDTA

(0.5% trypsin, 0.02% EDTA) for 5 min at 37°C. Fresh, pre-warmed medium was added to deactivate trypsin, and the suspended cells were centrifuged at 250g for 5 min. The supernatant was discarded and the cell pellet was resuspended in a sufficient volume of fresh pre-warmed medium, counted and passaged at a ratio of 1:3.

2.1.2 Primary cell culture

Human oral mucosa was recovered at the Dublin Dental University Hospital from patients undergoing routine third molar extraction in the Department of Oral and Maxillofacial Surgery. The sample was immediately placed into pre-warmed collection medium [DMEM, penicillin/streptomycin (100U/100µg), amphotericin B (2.5µg/ml)] for 10 min before washing the tissue three times with pre-warmed 1X PBS and placing it in 0.17% trypsin overnight at 4°C. The following day, the sample was washed with 1X PBS and connective tissue was removed using a scalpel. The tissue was cut into small pieces (1mm x 2mm) and the small sections were placed in pre-treated T25 flasks (CELL+, Sarstedt), each with a small coating of keratinocyte growth medium (KGM). These were left to adhere to the flask for 1-2h and the flask was subsequently flooded with KGM. Once sufficient growth of cells from tissue was achieved (2-3 weeks), KGM was replaced with Epilife medium (Invitrogen) to select for epithelial cell growth. Cells were passaged using 0.05% trypsin and spun at 250g for 10min at 4°C and cultured to 90% confluency.

2.1.3 Sample preparation

To facilitate Raman spectroscopy measurements, cells were detached from the flasks using 0.025% Trypsin–EDTA at 37°C and pelleted at 250g for 5 min at room temperature. The supernatant was removed and cells were counted and seeded at a density of 5×10^4

cells/calcium fluoride (CaF₂) disc in a multiwell plate and maintained, as previously described, until a monolayer of cells was stably growing on the disc. The cells were then fixed with 10% neutral buffered formaldehyde for 5 min, washed with 1X PBS and stored in 0.9% physiological saline solution prior to capture of the Raman spectrum.

2.2 RAMAN SPECTROSCOPY MEASUREMENTS

The study was conducted with a Horiba Jobin-Yvon LabRam HR800 instrument using a 532 nm laser as the source in a backscattering geometry, and a 300 lines/mm grating, providing a dispersion of $\sim 1.5\text{cm}^{-1}$ per pixel. The laser power was approximately 35 mW at the sample. Spectra were taken in the range from 600 cm^{-1} to 1800 cm^{-1} with a confocal hole diameter of $100\text{ }\mu\text{m}$. A 100x water immersion objective (LUMPlanF1, Olympus, N.A.: 1.0) was used to focus the laser on the sample, immersed in distilled water, providing a spatial resolution of $\sim 1\text{ }\mu\text{m}$ (Fullwood *et al.*, 2014; Bonnier *et al.*, 2012 and Bonnier *et al.*, 2010). Water immersion has been demonstrated to reduce any photothermal damage to the cells during measurement, and the signals were observed to be stable and reproducible (Bonnier *et al.*, 2012). For each cell line, 20 cells were analyzed. For each cell, three different subcellular regions were analyzed, resulting in one spectrum for each subcellular region of each cell. Thus, for each cell line, 20 nucleolar, 20 nuclear and 20 cytoplasmic spectra were recorded, each for a period of 2 x 20 seconds.

2.3 DATA PREPROCESSING

Data preprocessing was performed using Matlab (Mathworks, USA). Before statistical analysis, a Savitsky-Golay filter (5th order, 7 points) was applied to smooth the spectra. As it has been demonstrated that, in water immersion, the background to the Raman spectrum is

simply the spectrum of the overlying water (Bonnier et al., 2011), no background was subtracted. All spectra were (vector) normalized to remove point to point intensity variations and facilitate comparison of all spectra.

2.4 DATA ANALYSIS

Principal Components Analysis (PCA) is a method of multivariate analysis broadly used with datasets of multiple dimensions (Varmuza, 2009; Korenius *et al.*, 2007; German *et al.*, 2006). It allows the reduction of the number of variables in a multidimensional dataset, although it retains most of the variation within the dataset. The order of the principal components (PCs) denotes their importance to the dataset. PC1 describes the highest amount of variation, PC2 the second highest, and so on (Kelly *et al.*, 2011; Martin *et al.*, 2007). A PCA Scatter plot groups similar datasets (spectra) according to the loadings of the PCs and can be used to distinguish different datasets (samples). The loadings represent the variance for each variable (wavenumber) for a given PC. Analysing the loadings of a PC can give information about the source of the variability inside a dataset, in the case of spectroscopy, derived from variations in the molecular components contributing to the spectra.

PCA was employed in this study to highlight the variability existing in the spectral data set recorded for the different subcellular regions, and to differentiate the spectroscopic signatures of different cell lines.

3 RESULTS AND DISCUSSION

3.1 SINGLE CELL ANALYSIS (Normal primary cell culture)

Analysis of the subcellular regions, nucleolus, nucleus and cytoplasm, from the primary cell culture (Figure 1 A), was completed and the different subcellular regions were compared using PCA analysis. Figure 1 (B, C, D) shows the average spectra for each subcellular region, the standard deviation being illustrated by the shaded region. Visually, it is possible to note the similarities, but few differences, between the average spectra of the nucleolus and nucleus. That of the cytoplasm is, however, clearly different. Notable differences, highlighted by the boxes in figures 1 B, C and D, include the peak at $\sim 700\text{cm}^{-1}$ observed in cytoplasm, the single peak around $\sim 800\text{cm}^{-1}$ observed in the nucleolus and nucleus, and also the profiles of the spectra between 1200cm^{-1} and 1400cm^{-1} . These features can be associated with specific chemical moieties as indicated in Table S1 of the Supplementary Information, and their differing relative strengths are an indication of differing chemical composition of the subcellular region analysed. For example, the peak at $\sim 720\text{cm}^{-1}$ can be associated with lipidic components, which are relatively strong in the cytoplasm, while the peak at $\sim 790\text{cm}^{-1}$ is associated with nucleic acids, predictably relatively strong in the nucleus and nucleolus. The changes in the spectral region from 1200 to 1400 cm^{-1} are more complex and subtle, illustrating the need for more sophisticated multivariate data analysis techniques such as PCA (Yasser et al., 2014; Bonnier and Byrne, 2012).

Using PCA clustering, the nucleolus, nucleus and cytoplasm were clearly differentiated and spatially located. In confocal operation, the focal depth is $\sim 1\text{-}2\ \mu\text{m}$, and thus it can be expected that while sampling the nucleus or nucleolus, the overlying plasmatic membrane and cytoplasm may contribute to the observed signal. Nevertheless,

good discrimination of the respective subcellular regions is achievable. The nuclear (nucleolus and nucleus) regions and cytoplasm are effectively separated according to PC1 as shown in figure 2A. Similarly, it can be seen that the nucleolus and nucleus are differentiated according to PC3, although no significant differences in the average spectra are immediately obvious in Figure 1 B, C. No significant differentiation is observed according to PC2, and the loadings plot of this component is dominated by spectral features of water. A detailed analysis of the peaks of the loadings of PC1 and PC3 is provided in Supplementary Information S3.1. The analysis shows that it was possible to separate the subcellular regions from each other based on their biochemical content, even though the nucleolus and nucleus are molecularly quite similar (Bonnier and Byrne, 2012).

The spectral loadings from PC1 and PC3 are respectively illustrated in Figures 2B and 2C. With respect to PC1, the spectra from the nucleolus and nucleus are clustered on the right (positive values) and the spectra from cytoplasm clustered on the left (negative values) of Figure 2 B. Thus, positive peaks in the loading of PC1 correspond to chemical species which are more abundant in the nucleolus and nucleus, and negative to species which are relatively more abundant in the cytoplasm (Notinger, 2007). Clearly, the main features associated with the nucleolus and nucleus are related to nucleic acids, and protein structural components, whereas the features associated with the cytoplasm are derived from lipids and carbohydrate components, which could be related to organelle membrane content.

With respect to PC3, the spectra from the nucleolus are clustered on the top right quadrant (positive values) and the spectra from the nucleus are clustered on the bottom right (negative values) of Figure 2A. Thus, positive peaks in the loading of PC3 shown in

Figure 2C correspond to species which are more abundant in the nucleolus, and negative to species which are relatively more abundant in the nucleus (Notinger, 2007). Thus, the main features associated with the nucleolus are related to nucleic acids, whereas the features associated with the nucleus are related to protein structural components, indicating a higher density of these components in the respective subcellular regions.

3.2 SQUAMOUS CELL CARCINOMA vs DYSPLASTIC CELLS vs NORMAL CELLS

In individual cell lines, Raman spectroscopy, coupled with multivariate statistical analysis, has thus been demonstrated to be a powerful tool to discriminate different subcellular regions based on their biochemical content. The technique can also be extended to differentiate between different cell lines, representing different oral pathologies.

A similar spectral screening was conducted for the three cell lines, SCC-4, DOK and primary mucosal epithelial cells and PCA was used for statistical analysis of the Raman spectroscopic data from the respective subcellular regions for each cell line. In all cases, PC1 was discarded, as it was dominated by water contributions and the analysis was therefore based on PC2, PC3 and PC4 in order to discriminate the cell lines according to each subcellular region. The variances explained by each PC for the respective cellular regions are tabulated in Table 2.

The peak at $\sim 720\text{cm}^{-1}$ can be associated with lipidic components, which are relatively strong in the cytoplasm, while the peak at $\sim 790\text{cm}^{-1}$ is associated with nucleic acids, predictably relatively strong in the nucleus and nucleolus. The changes in the spectral region from 1200 to 1400 cm^{-1} are more complex and subtle, illustrating the need for more

sophisticated multivariate data analysis techniques such as PCA (Yasser et al., 2014; Bonnier and Byrne, 2012).

3.2.1 Nucleolus

The scatter plot of Figure 3A shows the PCA of the Raman spectral data of the nucleolar regions of the three cell lines, comparing PC2 and PC3. Normal cells (blue) score predominantly negative with respect to PC2, while abnormal cells (SCC-4 (red) and DOK (green)), both score predominantly positively, giving a 90% discrimination according to PC2, calculated as the ratio of the number of spectra which are “correctly” clustered negatively, to those which are “incorrectly” clustered positively. The loading plot for this PC (Figure 3B), therefore shows positive peaks corresponding to species which are more prominent in the spectra of nucleoli of abnormal cells, while the negative peaks are related to spectra of normal cells. A detailed analysis of the features of the loadings is provided in Supplementary Information S3.2.1. In summary, the nucleolar region contains the main percentage of nucleic acids (Chen *et al.*, 2014) compared to the other subcellular regions, so we can argue that, from a spectroscopic point of view, every peak related to vibrational modes of nucleic acids could be important as a biomarker for discrimination of the cell lines in the nucleolus region.

Positive features include bands at: $788/811\text{ cm}^{-1}$, related to O-P-O stretching of DNA and nucleic acids in general, 1420 cm^{-1} related to CH_2 deformation from lipids and carbohydrates, and 1680 cm^{-1} , related to amide I vibrational modes of proteins. Negative peaks include 1220 cm^{-1} , related to A (adenine), T (thymine) vibrational modes from nucleic acids (Notinger, 2007).

3.2.2 Nucleus

Figure 4A shows the PCA scatter plot for the Raman spectral data from the nuclear region of the three cell lines. In this case, there is little or no differentiation of the data according to PC2, but there is some degree of differentiation between the cell lines according to PC3, the normal cell lines being predominant distributed on the negative axis, while the diseased SCC-4 and DOK data points are primarily distributed on the positive side, giving an 82.5% differentiation. The loading plot for this PC (Figure 4B) shows positive peaks corresponding to biochemical species which are more abundant in the nucleus of abnormal cell lines, whereas the negative peaks were related to spectra from the normal primary cells. A detailed analysis of the features of the loadings is provided in Supplementary Information S3.2.2. The subcellular region of the nucleus contains a large percentage of nucleic acids compared to the cytoplasm, although some changes to lipid content is observed. Thus, a combination of nucleic acid/lipid content has spectroscopic significance as a biomarker to discriminate the normal and abnormal cell lines in the nuclear region. In this case, little or no differentiation is observed between the two abnormal cell lines, however.

Positive features include mostly bands at: 1065 cm^{-1} , related to PO_2^- of nucleic acids, and 1680 cm^{-1} , related to the Amide I vibrational modes from proteins (Chen et al., 2014). Negative peaks mostly include: $788/811\text{ cm}^{-1}$, related to O-P-O stretching of DNA and nucleic acids in general, 1301 cm^{-1} related to CH_2 twist of lipids and 1420 cm^{-1} related to CH_2 deformation from lipids and carbohydrates (Notinger, 2007).

3.2.3 Cytoplasm

The scatter plot of Figure 5A shows PCA of the cytoplasmic regions of the three cell lines. Some degree of differentiation between the normal and abnormal cell lines according to PC3 is apparent giving 90% of discrimination. Similarly, PC2 differentiates between the spectra of the dysplastic and cancer cell lines giving 77.5% of discrimination (Figure 5A).

Analysing PC3, the main features include positive bands at $788/811\text{ cm}^{-1}$, related to O-P-O stretching of DNA, 1301 cm^{-1} related to CH_2 twist of lipids and 1420 cm^{-1} related to CH_2 deformation from lipids and carbohydrates. Negative peaks include: 937 cm^{-1} , related to the C-C backbone stretch of the α -helix of proteins and the C-O-C vibrations of carbohydrates, and 1658 cm^{-1} , related to Amide I -proteins (Figure 5B) (Notinger, 2007). The spectral profile of the cytoplasm subcellular region is dominated by vibrational modes from lipids and carbohydrates, which could be related to any physiologic subcellular transport and organelle membranes and both constituents could be important as biomarkers for discrimination of the cell lines in cytoplasm region. Notably, the cytoplasm was the only subcellular region in which some discrimination between dysplastic and cancer cells was observed according to the loading plot of PC2 (Figure 5B).

The analysis demonstrates that the technique may not only be utilized to differentiate normal and abnormal cells, but also dysplastic and cancer cells, based on cytoplasmic biomarkers. It is notable that the dysplastic and cancer cells are discriminated only based on cytoplasmic, rather than nucleolar or nuclear biomarkers, suggesting that the cells are relatively similar in biochemical content in these regions and the pathological differences are largely manifest in the cytoplasmic regions. Further analysis is required to explore this further, however.

Notably, the cytoplasm was the only subcellular region in which some discrimination between dysplastic and cancer cells was observed according to the loading plot of PC2 (Figure 5B). Positive features include mostly bands at: 788 cm^{-1} , related to O-P-O stretching of DNA and nucleic acids in general, 1301 cm^{-1} related to CH_2 twist of lipids and 1420 cm^{-1} related to CH_2 deformation from lipids and carbohydrates. Negative peaks mostly include: 937 cm^{-1} , related to C-C backbone stretching in α -helical proteins and C-O-C vibrations in glycosylated carbohydrates, $1220\text{ -}1284\text{ cm}^{-1}$ related to $=\text{CH}$ bend of lipids and 1640 cm^{-1} related to the Amide I band of proteins (Notinger, 2007).

Variations between the samples as reported here have a number of possible origins, both intrinsic and due to measurement conditions. As the measurements were performed in immersion geometry, and no explicit removal of it was carried out, the water spectrum emerges as one of the most significant principal components, in all measurements. In the case of the normal cells, the water contribution to the variance of the difference between nucleus, nucleolus and cytoplasm, described by PC2, as the differences between the subcellular regions are dominantly those associated with the biochemical make up and described by PC1 (~70% variance). In the case of the comparison of the same subcellular regions of different cells, slight variations in size and shape contribute significant variations in the contributions of the overlying water, represented by PC1 (~70 % variances). This variance is larger than the more subtle differences in biochemical makeup of the respective subcellular regions between the cell types. Cancer cells are continually undergoing transformation to enable survival, replication and metastasis and potentially changes observed in the Raman spectra may be a direct observation of some of these changes. These changes are driven through a number of cellular mechanisms and can manifest as

posttranslational modification (chemical processes resulting in the cleavage or modification of proteins) e.g. glycosylation, phosphorylation or epigenetic changes (heritable changes not due to changes in the DNA sequence) e.g. histone modification, DNA methylation. The sensitivity of the analysis used in this study is unlikely to identify such specific changes in cells, but both posttranslational modification and epigenetic changes have been observed utilizing surface-enhanced Raman spectroscopy (Barhoumi & Halas, 2011; Sundararajan et al 2006) and variation between the cell types used are potentially the result of an accumulation of these cellular changes (>75%). Notably, however, more sophisticated multivariate classification techniques can be applied to achieve higher degrees of sensitivity and specificity for real clinical applications (Lasalvia et al., 2015; Knipfer et al., 2014; Lyng et al., 2007).

5 CONCLUSION

It has been demonstrated that Raman micro-spectroscopy can differentiate each subcellular region (nucleolus, nucleus and cytoplasm), based on biomolecular content, revealing specific spectral biomarkers for each region. Although the spectral differences are subtle, the use of an unsupervised multivariate analytical method, PCA, allows a detailed, high content analysis of the spectral markers associated with the biochemical differences. The technique is completely label free, requires little sample preparation and is nondestructive.

Importantly, it was also possible to discriminate the pathological cell lines from the normal primary cells based on differences in the spectral profiles of either the nuclear

(>90%) or nucleolar (>80%) regions of the respective cell lines. Interestingly, the dysplastic and cancer cells were only discriminated in the analysis of the cytoplasmic regions.

Although the study presented is a proof of concept, and more elaborate analysis of normal and dysplasia/cancer cytological samples from the same patients will be required to explore intra and inter patient variabilities, the model cells used in this study are well established oral cancer models and are representative of the stage of the disease described and are indicative of clinical samples. The inclusion of isolated primary cells which are of a normal phenotype are an acceptable comparator in this study. Further such studies of a range of clinical samples are planned.

Notably, as an optically-based microscopic technique, the Raman spectroscopic analysis is readily adaptable to clinical applications. The development of fiber optic Raman probes ultimately offer the promise of in vivo, optically-directed biopsy in extensive areas of leukoplakia and intraoperative applications in determining tumour resection margins. In the short term, existing microscope based Raman technologies might also be adopted for chair-side diagnosis based on exfoliated cytological samples which could yield a rapid outcome and a significant reduction in the number of false negatives/positives along with early cancer detection, as evidenced by the observed changes in established cancer and dysplastic lesions.

6 ACKNOWLEDGEMENTS

Luis Felipe CS Carvalho was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – Brazil) for post doctoral research under Science Without Borders (237749/2012-2).

This research was supported in part by Science Foundation Ireland (12/IP/1494) and the National Biophotonics and Imaging Platform (NBIP) Ireland funded under the Higher Education Authority PRTL (Programme for Research in Third Level Institutions) Cycle 4, co-funded by the Irish Government and the European Union Structural Fund.

7 REFERENCES

Barhoumi A, Halas NJ (2011). Detecting Chemically Modified DNA Bases Using Surface Enhanced Raman Spectroscopy. *J Phys Chem Lett.* 2(24): 3118–3123.

Bonnier F, Ali SM, Knief P, Lambkin H, Flynn K, McDonagh V, *et al.* (2012). Analysis of human skin tissue by Raman microspectroscopy: Dealing with the background. *Vibrat Spec* 61:124–132.

Bonnier F, Mehmood A, Knief P, Meade AD, Hornebeck W, Lambkin H, *et al.* (2010). In vitro analysis of immersed human tissues by Raman microspectroscopy. *J Ram Spec* 42(5):888 - 896.

Bonnier F, Byrne HJ (2012). Understanding the molecular information contained in Principal Component Analysis of Vibrational Spectra of Biological Systems. *Analyst* 137:322-332.

Bonnier F, Knief P, Lim B, Meade AD, Dorney J, Bhattacharya K, *et al.* (2010). Imaging live cells grown on a three dimensional collagen matrix using Raman microspectroscopy. *Analyst* 135:3169-3177.

Bonnier F, Traynor D, Kearney P, Clarke C, Knief P, Martin C, *et al.* (2014). Processing ThinPrep Cervical Cytological Samples for Raman Spectroscopic Analysis *Analytical Methods*, 6, 7831-7841.

Byrne HJ, Ostrowska KM, Nawaz H, Dorney J, Meade AD, Bonnier F, *et al.* (2014). Vibrational Spectroscopy: Disease Diagnostics and Beyond, in *Optical Spectroscopy and Computational Methods in Biology and Medicine*, Challenges and Advances in Computational Chemistry and Physics, 14: 355-99.

Cabral L, de Carvalho LF, Salgado J, Brandão AA, Almeida JD (2010). Gingival squamous cell carcinoma: a case report. *J Oral Maxillofac Res* 1:1(3):e6.

Carvalho LF, Bitar RA, Arisawa EA, Brandão AA, Honório KM, Cabral LA, *et al* (2010). Spectral region optimization for Raman-based optical biopsy of inflammatory lesions. *Photomed Laser Surg* 28 Suppl 1:S111-7.

Chen Y, Dai J, Zhou X, Liu Y, Zhang W, Peng G (2014). Raman spectroscopy analysis of the biochemical characteristics of molecules associated with the malignant transformation of gastric mucosa. *PLoS One* 7;9(4):e93906.

Damayanti NP, Fang Y, Parikh MR, Craig AP, Kirshner J, Irudayaraj J (2013). Differentiation of cancer cells in two-dimensional and three-dimensional breast cancer models by Raman spectroscopy. *J Biomed Opt* 18(11):117008. Dorney J, Bonnier F, Garcia A, Casey A, Chambers G, Byrne HJ (2012). Identifying and localizing intracellular nanoparticles using Raman spectroscopy. *Analyst* 137: 1111-1119.

Evers D, Hendriks B, Lucassen G, Ruers T (2012). Optical spectroscopy: current advances and future applications in cancer diagnostics and therapy. *Future Oncol* 8(3):307-20.

Fullwood LM, Griffiths D, Ashton K, Dawson T, Lea RW, Davis C, *et al.* (2014). Effect of substrate choice and tissue type on tissue preparation for spectral histopathology by Raman microspectroscopy. *Analyst* 139(2):446-54

German MJ, Hammiche A, Ragavan N, Tobin MJ, Cooper LJ, Matanhelia SS, *et al.* (2006). Infrared spectroscopy with multivariate analysis potentially facilitates the segregation of different types of prostate cell. *Biophys J* 90(10):3783-95.

Guze K, Short M, Zeng H, Lerman M, Sonis S (2011). Comparison of molecular images as defined by Raman spectra between normal mucosa and squamous cell carcinoma in the oral cavity. *J Raman Spectrosc* 42(6):1232-1239.

Kallaway C, Almond LM, Barr H, Wood J, Hutchings J, Kendall C, *et al.* (2013). Advances in the clinical application of Raman spectroscopy for cancer diagnostics. *Phot Photodyn Ther* 10(3):207-19.

Kelly K, Johnson-Obaseki S, Lumingu J, Corsten M (2014). Oncologic, functional and surgical outcomes of primary transoral robotic surgery for early squamous cell cancer of the oropharynx: a systematic review. *Oral Oncol* S1368-8375(14)00128-6.

Kelly JG, Trevisan J, Scott AD, Carmichael PL, Pollock HM, Martin-Hirsch PL, *et al.* (2011). Biospectroscopy to metabolically profile biomolecular structure: a multistage approach linking computational analysis with biomarkers. *J Proteome Res* 10(4):1437-48.

Knipfer C, Motz J, Adler W, Brunner K, Gebrekidan MT, Agaimi A, *et al* (2014). Raman difference spectroscopy: a non-invasive method for identification of oral squamous cell carcinoma. *Bio Opt Exp* 5(9): 3252-3265.

Lasalvia M, D'Antonio P, Perna G, Capozzi V, Mariggìò MA, Perrone D, *et al* (2015). Discrimination of different degrees of oral squamous cell carcinoma by means of Raman microspectroscopy and atomic force microscopy. *Anal Methods* 7:699-707.

Laurikkala KJ, Juhola M (2007). On principal component analysis, cosine and euclidean measures in information retrieval. *Inf Sci* 177(22):4893–4905.

Luo S, Chen C, Mao H, Jin S (2013). Discrimination of premalignant lesions and cancer tissues from normal gastric tissues using Raman spectroscopy. *J Biomed Opt* 18(6):06700412.

Lyng FM, Faoláin EO, Conroy J, Meade AD, Knief P, Duffy B, *et al.* (2007). Vibrational spectroscopy for cervical cancer pathology, from biochemical analysis to diagnostic tool. *Exp Mol Pathol* 82(2):121-9.

Marro M, Nieva C, Sanz-Pamplona R, Sierra A (2014). Molecular monitoring of epithelial-to-mesenchymal transition in breast cancer cells by means of Raman spectroscopy. *Biochim Biophys Acta* 1843(9):1785-95.

Martin FL, German MJ, Wit E, Fearn T, Ragavan N, Pollock HM (2007). Identifying variables responsible for clustering in discriminant analysis of data from infrared microspectroscopy of a biological sample. *J Comput Biol* 14:1176–1184.

Mittal R, Balu M, Krasieva T, Potma EO, Elkeeb L, Zachary CB, *et al.* (2013). Evaluation of stimulated Raman scattering microscopy for identifying squamous cell carcinoma in human skin. *Lasers Surg Med* 45(8):496-502.

Nawaz H, Bonnier F, Meade AD, Lyng FM, Byrne HJ (2011). Comparison of subcellular responses for the evaluation and prediction of the chemotherapeutic response to cisplatin in lung adenocarcinoma using Raman spectroscopy. *Analyst* 136(12):2450-63.

Nottingham I (2007). Raman biospectroscopy cell based biosensors. *Sensors* 7:1343-1358.

Peer D (2014). Precision medicine - Delivering the goods? *Cancer Lett* S0304-3835(14)00218-3.

Schubert JA, Bird B, Papamarkakis K, Miljković M, Bedrossian K, Laver N, *et al* (2010). Spectral cytopathology of cervical samples: detecting cellular abnormalities in cytologically normal cells. *Lab Invest* 90(7):1068–1077.

Singh SP, Deshmukh A, Chaturvedi P, Krishna CM (2012). Raman spectroscopy in head and neck cancers: toward oncological applications. *J Cancer Res Ther* 1:S126-32.14.

Su L, Sun YF, Chen Y, Chen P, Shen AG, Wang XH, *et al* (2012). Raman spectral properties of squamous cell carcinoma of oral tissues and cells. *Laser Physics* 22(1): 311–316.

Sundararajan N, Mao D, Chan S, Koo TW, Su X, Sun L, *et al* (2006). Ultrasensitive detection and characterization of posttranslational modifications using surface-enhanced Raman spectroscopy. *Anal Chem*. 78(11):3543-50.

Raman CV, Krishnan KS (1928). A new type of secondary radiation. *Nature* 121: 501-502.

Tanahashi K, Natsume A, Ohka F, Momota H, Kato A, Motomura K, et al (2014). Assessment of tumor cells in a mouse model of diffuse infiltrative glioma by Raman spectroscopy. *Biomed Res Int* 860241.

Vargis E, Tang YW, Khabele D, Mahadevan-Jansen A (2012). Near-infrared Raman Microspectroscopy Detects High-risk Human Papillomaviruses. *Transl Oncol* 5(3):172-9.

Varmuza K (2009). Introduction to multivariate statistical analysis in chemometrics. Taylor & Francis Group ed. New york: CRC Press. 59–102.

Wang L, Zhang Z, Huang L, Li W, Lu Q, Wen M, et al (2014). Evaluation of Raman spectroscopy for diagnosing EGFR mutation status in lung adenocarcinoma. *Analyst* 139(2):455-63.

Yasser M, Shaikh R, Chilakapati MK, Teni T (2014). Raman spectroscopic study of radioresistant oral cancer sublines established by fractionated ionizing radiation. *PLoS One* 19;9(5):e97777.

Table 1: Band assignments of Raman peaks (Notinger, 2007)

| Band (cm⁻¹)/Subcellular region | Nucleolus <i>Vibrational mode/Structural component</i> | Nucleus <i>Vibrational mode/Structural component</i> | Cytoplasm <i>Vibrational mode/Structural component</i> |
|--------------------------------------------------|-------------------------------------------------------------------|-----------------------------------------------------------------|-------------------------------------------------------------------|
| 717 | | | CN+(CH ₃) ₃ str. - Lipids |
| 729 | A – Nucleic acids | | |
| 760 | Ring breath Try - Proteins | | |
| 811-14 | O-P-O str. RNA – Nucleic acids | | |
| 937 | | | C-C Bk str. α-helix – Proteins C-O-C glycos. – Carbohydrates |
| 1065 | PO ₂ ⁻ str. - Nucleic acids | PO ₂ ⁻ str. – Nucleic acids | Chain C-C str. – Lipids |
| 1301 | | CH ₂ twist - Lipids | CH ₂ twist – Lipids |
| 1320-4 | G – N ac./ C-H – Proteins | | |
| 1342 | | A, G – N ac./ C-H – Proteins | C-H def. – Carbohydrates |
| 1441 | G, A, CH def – Nucleic acids | G, A, CH def – Nucleic acids | C-H def. – Carbohydrates/Lipids |
| 1487 | | | C-H def. – Carbohydrates/Lipids |
| 1658 | | Amide I – Proteins | |

Table 2: Variances explained by each PC for the respective cellular regions

| PC/Region | Nucleolus% | Nucleus% | Cytoplasm% |
|-------------|------------|----------|------------|
| PC1 (water) | 76.08 | 76.69 | 63.42 |
| PC2 | 5.42 | 6.30 | 17.72 |
| PC3 | 4.05 | 4.05 | 6.34 |

Figure captions

Figure 1: **A** - Optical Micrograph of a primary cell cultured from human mucoperiosteal tissue. Mean Raman spectra from **(B)** Nucleolus **(C)** Nucleus **(D)** Cytoplasm of primary cell line with the standard deviation denoted by the shaded region.

Figure 2: **A** - PCA Scatter plot of Raman spectra of primary cell line Nucleolus (red), Nucleus (green) and cytoplasm (blue). **B** - Spectral loading of PC1, differentiating cytoplasm (negative) from nucleus/nucleolus (positive) of primary cell line. **C** - Spectral loading of PC3, differentiating nucleus (negative) from nucleolus (positive) of primary cell line.

Figure 3: **A** - PCA scatter plot differentiating the Raman spectra of the nucleoli of the primary (normal) cell line (blue), from those of the abnormal cell lines (SCC-4 (red) and DOK (green)). **B** - Spectral loading of PC2, differentiating the Raman spectra of the nucleoli of the primary (normal) cell line (negative), from those of the abnormal cell lines (positive).

Figure 4: **A** - PCA scatter plot differentiating the Raman spectra of the nuclei of the primary (normal) cell line (blue), from those of the abnormal cell lines (SCC-4 (red) and DOK (green)). **B** - Spectral loading of PC3, differentiating the Raman spectra of the nuclei of the primary (normal) cell line (negative), from those of the abnormal cell lines (positive).

Figure 5: **A** - PCA scatter plot differentiating the Raman spectra of the cytoplasm of the primary (normal) cell line (blue), from those of the abnormal cell lines (SCC-4 (red) and DOK (green)). **B** - Spectral loading of PC2 and PC3, differentiating the Raman spectra of the nuclei of the primary (normal) cell line (negative), from those of the abnormal cell lines (positive).









