2017

RAMAN SPECTROSCOPIC ANALYSIS OF ORAL CELLS IN THE HIGH WAVENUMBER REGION

Fiona Lyng
Technological University Dublin, fiona.lyng@tudublin.ie

LUIS FELIPE C.S CARVALHO
luisfelipecarvalho@hotmail.com

Franck Bonnier
Technological University Dublin, Franck.Bonnier@tudublin.ie

See next page for additional authors

Recommended Citation

Authors

Fiona Lyng, LUIS FELIPE C.S CARVALHO, Franck Bonnier, CLÁUDIO TELLEZ, LAURITA DOS SANTOS, KATE O´CALLAGHAN, Jeff O'Sullivan, LUIS EDUARDO S. SOARES, and Hugh Byrne

RAMAN SPECTROSCOPIC ANALYSIS OF ORAL CELLS IN THE HIGH WAVENUMBER REGION

LUIS FELIPE C.S. CARVALHO[1,2*], FRANCK BONNIER[3], CLÁUDIO TELLEZ[2], LAURITA DOS SANTOS[2], KATE O´CALLAGHAN[4], JEFF O´SULLIVAN[4], LUIS EDUARDO S. SOARES[2], STEPHEN FLINT[4], AIRTON A. MARTIN[2], FIONA M. LYNG[5,6], HUGH J. BYRNE[1]

[1]FOCAS RESEARCH INSTITUTE, DUBLIN INSTITUTE OF TECHNOLOGY, KEVIN STREET, DUBLIN 8, IRELAND
[2]UNIVERSIDADE DO VALE DO PARAÍBA, SÃO JOSÉ DOS CAMPOS-SP, BRAZIL
[3]UNIVERSITE FRANÇOIS-RABELAIS DE TOURS, FACULTE DE PHARMACIE, EA 6295 NANOMEDICAMENTS ET NANOSONDES, 31 AVENUE MONGE, 37200 TOURS, FRANCE.
[4]DUBLIN DENTAL SCHOOL AND HOSPITAL, TRINITY COLLEGE DUBLIN, DUBLIN 2, IRELAND
[5]RADIATION AND ENVIRONMENTAL SCIENCE CENTRE, FOCAS RESEARCH INSTITUTE, DUBLIN INSTITUTE OF TECHNOLOGY, KEVIN STREET, DUBLIN 8, IRELAND
[6]SCHOOL 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 signature of the biochemical composition and structure of cells with submicrometer spatial resolution and could be useful to monitor changes in composition for early stage and non-invasive cancer diagnosis, both *ex-vivo* and *in vivo*. In particular, the fingerprint spectral region (400–1,800 cm^{-1}) has been shown to be very promising for optical biopsy purposes. However, limitations for discrimination of dysplastic and inflammatory processes based on the fingerprint region have been demonstrated. In addition, the Raman spectral signal of dysplastic cells is one important source of misdiagnosis of normal versus pathological tissues. The high wavenumber region (2,800–3,600 cm^{-1}) provides more specific information based on N-H, O-H and C-H vibrations and can be used to identify the subtle changes which could be important for discrimination of samples. In this study, we demonstrate the potential of the high-wavenumber spectral region in this context by collecting Raman spectra of nucleolus, nucleus and cytoplasm from oral epithelial cancer (SCC-4) and dysplastic (DOK) cell lines and from normal oral epithelial primary cells, *in vitro*, in water immersion, which were then analyzed by principal components analysis as a method to discriminate the spectra. Analysis was performed before and after digital subtraction of the bulk water signal. In the normal cell line, the three subcellular regions are well differentiated before water subtraction, although the discrimination of the two nuclear regions is less well defined after water subtraction. Comparing the respective subcellular regions of the three cell lines, before water subtraction, the cell lines can be discriminated using sequential PCA and Feature Discriminant Analysis with up to ~ 100% sensitivity and 97% specificity for the cytoplasm, which is improved to 100% sensitivity and 99% specificity for the nucleus. The results are discussed in terms of discrimination comparing the CH vibrational modes of nucleic acids, proteins and lipids. The potential role of the OH vibrations, considering free water and confined water, in the discrimination of cell cultures and pathological processes are also discussed.

Keywords: Raman spectroscopy, high-wavenumber, optical biopsy, oral cancer, medical diagnostics

1. INTRODUCTION

Recently, a range of new diagnosis tools based on photonic technologies has been developed. One of these tools is the optical or spectroscopic biopsy. Optical biopsy refers to techniques in which the light-tissue interaction is analyzed and information concerning the state of the tissue is obtained either ''in vivo'' or ''ex vivo''. Optical spectroscopy techniques such as fluorescence, infrared absorption and Raman scattering can be employed. Holmstrup et al. [1] and Singh et al. [2]

have described many molecular interaction features in cells and tissues that cannot be accessed by conventional histopathology that can be probed by optical techniques.

In vibrational spectroscopy (infrared absorption or Raman scattering), the spectral region between 400 and 1,800 cm⁻¹ (fingerprint region) is considered to contain the most relevant biochemical information concerning biological tissues [3-6]. Many vibrational band frequencies of amino acids, nucleic acids, proteins, lipids, glucose, and other carbohydrates fall within this region, providing a complex fingerprint of the biochemical composition, usually requiring multivariate analysis [7-8]. One interesting possibility is to use the so-called high wavenumber region, 2,800–3,600 cm⁻¹, as it can provide more specific information based on N-H, O-H and C-H vibrations of lipids and proteins, which could complement analysis of the fingerprint region. Carvalho et al. [9] have demonstrated that analysis of the high wavenumber region yielded better discrimination of inflammatory oral tissues, highlighting the importance of the OH vibrations. Barroso et al. [10] have more recently demonstrated discrimination between oral cancer and healthy tissue based on water content of freshly excised tongue specimens, determined by Raman spectroscopy in the high wavenumber region. This dramatic signature difference can be used to determine the location of the tumor border in oral cancer surgery [11]. Santos et al. used Raman spectroscopy to characterize melanoma and benign melanocytic lesions suspected of melanoma, based on the stretching vibrations of lipids in the high-wavenumber region [12].

In this study, we further explore the diagnostic potential of the high-wavenumber spectral region at a subcellular level by collecting Raman spectra of nucleolus , nucleus and cytoplasm from normal oral epithelial primary cells, and from oral epithelial cancer (SCC-4) and dysplastic (DOK) cell lines, *in vitro*, which were then analyzed by principal components analysis (PCA) and a combination of PCA and Feature Discriminant Analysis, as multivariate statistical methods to analyse and discriminate the spectra. In this region, the discrimination comparing the CH vibrational modes of nucleic acids, proteins and lipids will be demonstrated. The role of the OH vibrations, considering free water and confined water, also will be explored, in terms of cell culture discrimination and how the water could affect a pathologic process.

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. It is important to note that both pathological cell lines are originally from tongue, one of the sites of highest incidence of oral squamous cell carcinoma.

2.1.1 SCC-4 and DOK cell cultures.

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 to 37°C before incubation of the cells. Cells were cultured until 90% confluency in a humidified environment at 5% CO_2, 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. The primary cell culture is considered representative of "normal" pathology.

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 x 10^4 cells/calcium fluoride (CaF_2) disc in a multiwell plate and maintained 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 ~1.5cm^{-1} per pixel. The laser power was approximately 35 mW at the sample. Spectra were taken in the range from 2600 to 3800 cm^{-1} with a confocal hole diameter of 100 ⌈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 ~1μm [13-15]. 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 [14]. For each cell line, 20 cells were analyzed, and 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 the instrument software, LabSpec. 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 [15], water background subtraction was performed by non-negatively constrained least squares analysis (NCLS) using an in house Mathlab algorithm. After water subtraction, all spectra were subjected to a linear baseline correction (rubber band with 2 nodes at each extremity of spectra) followed by a vector normalization 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 [17-18]. 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 [19, 20] 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 within a dataset, in the case of spectroscopy, derived from variations in the molecular components contributing to the spectra [21].

Supervised learning methods such as discriminant analyses enable classification of unknown samples using prior knowledge of the category from training sets. To predict the qualitative affiliation of data points to a defined group (for instance cell type), the coupling of PCA with Factorial Discriminant Analysis (PCA-FDA) exploits the calculated PC scores to better evaluate the discrimination rate achieved based on a data set [22]. In the present study, 100 fold cross validation using 10 dimensions has been employed to validate the observations made with 2/3 of the total spectra randomly selected as part of the calibration set, while the remaining 1/3 was used as the validation set. Classification values obtained for each iteration (or loop) of the cross validation have been gathered to deliver an overall confusion matrix representative of the data discrimination on the high wavenumbers region of the spectra [22].

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. PCA-FDA was employed to further discriminate the subcellular regions of the different cell lines, demonstrating the diagnostic potential of the technique.

3. RESULTS

3.1. Single cell-line analysis

Spectroscopic analysis of the subcellular regions, nucleolus, nucleus and cytoplasm, from the primary cell culture (Figure 1 A), was performed after water subtraction and spectral normalization and the different subcellular regions were compared. Figure 1 (B, C, D) shows the average spectra for each subcellular area, the standard deviation being illustrated by the shaded region. Notably, the thinner cytoplasm also has a much more variable spectrum. In all spectra, the range from ~3100-3400 cm^{-1} can include residual environmental water, as well as contributions of other –OH groups, N-H stretching vibrations, Amide A bands of proteins and bound or semi-bound water [23, 24]. Characteristic lipid/protein features are observable in the range 2800-3000cm^{-1}, and, visually, the mean spectrum of the cytoplasm is clearly different from that of the other subcellular compartments. The spectral features in this range can be associated with specific chemical compounds, as indicated in Table 1, and their differing relative strengths are an indication of differing chemical composition of the subcellular area analysed. For example, the peak at 2851 cm^{-1} can be associated with CH_2 stretching in lipidic components, fatty acids and phospholipids, while 2870 cm^{-1} is related to CH_3 vibrational modes due asymmetric stretch of lipids and proteins which are relatively strong in the cytoplasm, while the peak at 2940 cm^{-1} is associated with CH vibrational modes of nucleic acids, predictably relatively strong in the nucleus and nucleolus [8, 9]. Notably, the two nuclear regions are not as identifiable, compared to the cytoplasm, and more detailed analysis is required to differentiate the spectra [3].
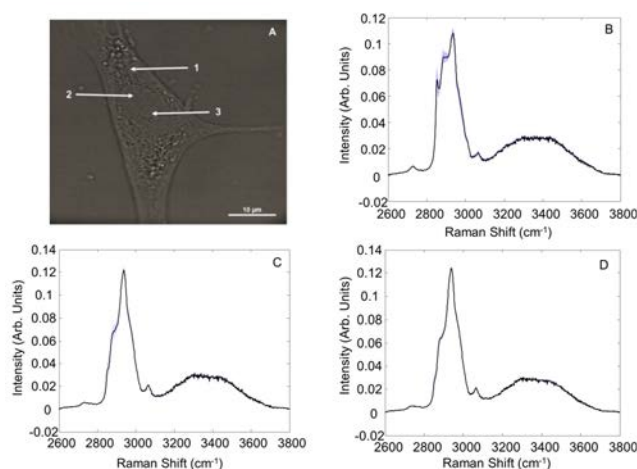

Figure 1: (A) Optical Micrograph of a primary cell cultured from human mucoperiosteal tissue. Mean Raman spectra after water subtraction from (B) Cytoplasm (C) Nucleus (D) Nucleolus of primary cell line with the standard deviation denoted by the shaded region.

Table 1: Raman spectroscopic band assignments of cellular components [7, 8, 23]

Peak (cm^{-1})	Assignment
2851	CH_2 symmetric stretch of lipids
2870	CH_3 asymmetric stretch of lipids and proteins
2890	CH_2 deformation
2936	CH_2 asymmetric stretch of proteins, related to nucleus and nucleoli content.
	CH vibrations in lipids & proteins, ν_{as} CH_2, lipids, fatty acids

3015	=CH, lipids, fatty acids
3180	OH bound
3210	OH (tightly bound)
3285	NH_2 asymmetric stretch (Amide A) which overlaps with OH
3345	OH (partially bound)
3470	OH symmetric stretch (unbound)
3550	OH stretching vibration

Principal Components Analysis can be employed to differentiate the subcellular regions, based on their biochemical differences, as represented by the loadings of the respective discriminating principal components. Figure 2 shows the example for the subcellular regions of the primary cells. The nuclear regions (nucleus and nucleoli) are differentiated from the cytoplasmic region according to PC1, while PC2 differentiates the two regions of the nucleus, albeit poorly.

PC1 is dominated by the negative peaks at 2851 cm^{-1}, assigned to CH_2 symmetric stretching of lipids and 2890 cm^{-1}, assigned to CH_2 deformation [9]. The positive loading is dominated by the 2936 cm^{-1} peak, assigned to the CH_2 asymmetric stretching of proteins, related to nucleus and nucleoli content.
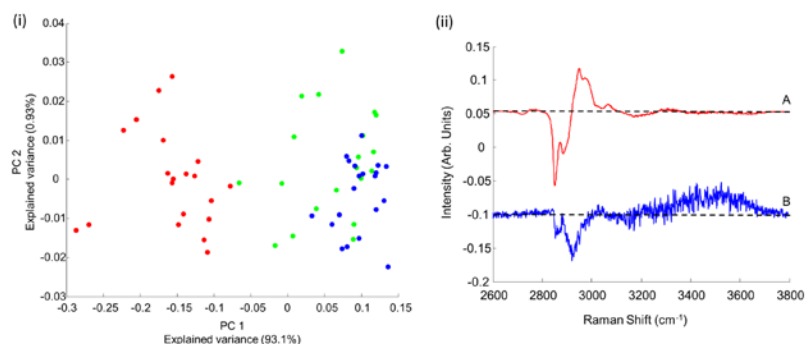

Figure 2: (i) PCA Scatter (Red dots – Cytoplasm, Green Dots – Nucleus, Blue dots – Nucleolus) and (ii) loading plot of PC 1 and PC2 for the spectra of subcellular regions of primary cells, after water subtraction (A:Red = Loading of PC1 and B:Blue = Loading of PC 2). The loadings are off-set for clarity, the dotted line indicating the zero level.

Comparing only the nuclear and nucleolar regions, PCA largely differentiates the two datasets according to PC1 (Figure 3A) which, as shown in Figure 3B, shows a similar profile (inverted) to that which differentiates the cytoplasmic and nuclear regions. The negative distribution of the nucleolar spectra associates it with a higher density of nucleic acid content.
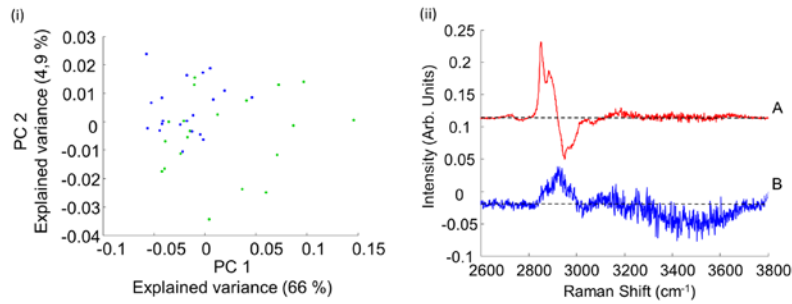
Figure 3: (i) PCA scatter plot of the spectra of the nuclear regions of primary cells after water subtraction (Nucleus: green Dots; Nucleolus: Blue dots) and (ii) Loadings plots corresponding to respectively Loading of PC1 (A, red) and loading of PC 2 (B, blue). The loadings are off-set for clarity, the dotted line indicates the zero level.

3.2. Comparison of Cells lines: Squamous Cell Carcinoma vs Dysplastic Cells vs Normal Cells

3.2.1 Discrimination of cell lines

The differentiation of the subcellular regions of the single cell line, based on their distinct biochemical compositions, serves as a demonstration of the analytical potential of Raman microspectroscopy. The same methodology was applied to compare the respective subcellular regions of the three cell lines, representative of different oral disease pathologies. Figures 4, 5 and 6 show the mean spectra and PCA of the different subcellular regions (cytoplasm, nucleus and nucleolus, respectively) of the three different cell types (NC, DOK and SCC), after subtraction of the water background.

In the case of the cytoplasm (Figure 4), it is possible to distinguish three different clusters, PC1 largely differentiating the normal and DOK clusters from the SCC cluster. The loading of PC1 shows positive features at 2936cm^{-1} indicating increased proteinic content in SCC cells, whereas the normal and DOK cells show increased lipidic content, indicated by the negative loadings of the features at 2851cm^{-1}. The normal and abnormal cell lines (DOK and SCC) are also partially differentiated by PC2, which, as well as features in the spectral region <3000 cm^{-1}, shows differentiating features in the region >3000 cm^{-1}. In addition to aqueous water, this region can also include contributions of other –OH groups, N-H stretching vibrations, Amide A bands of proteins and bound or semi-bound water (Table 1) [23, 24]. In the loading of PC2, positive features at 3180 cm^{-1}, 3210 cm^{-1}, 3285 cm^{-1}, 3345 cm^{-1}, related to Amide A-water and bound water are associated with normal and DOK cells, whereas negative peaks of 3470 cm^{-1} and 3550 cm^{-1}, related to unbound water, are associated with cancer cells.
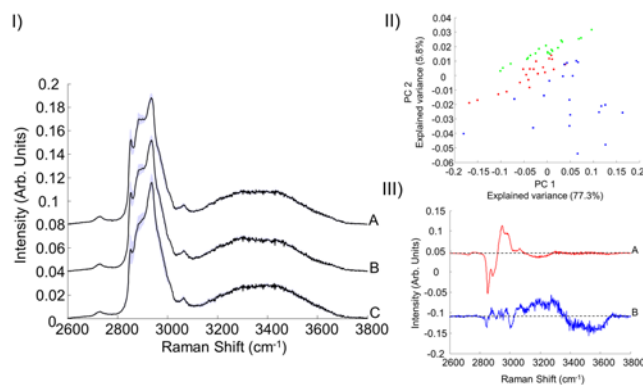

Figure 4: (I) Mean and SD spectra of the cytoplasm A – normal B – DOK C – SCC4 (II) PCA scatter plot differentiating the Raman spectra of the cytoplasm of the primary (normal) cell line (red), from those of the abnormal cell lines (SCC-4 (blue) and DOK (green)). (III) A: Red =Loading of PC1 and B:Blue = Loading of PC 2. Loadings are off set for clarity.

In the case of the nuclear regions, shown in Figure 5, PC2 differentiates the normal from the abnormal cell lines, DOK and SCC, while PC3 differentiates DOK and SCC. The loading of PC2 has positive features at 2851 cm^{-1}, 2936 cm^{-1}, associated with lipids and proteins, while the broader, negative features can be associated with unbound water in the normal cells. The profile of PC3 is very similar to that of PC2 in Figure 4III, which partially discriminates the cytoplasm of normal and DOK cells. While the normal cell spectra are distributed about the zero line, the DOK and SCC are distributed positively and negatively, respectively, indicating that the SCC has relatively more unbound or partially bound water, whereas the DOK are richer in bound water species.
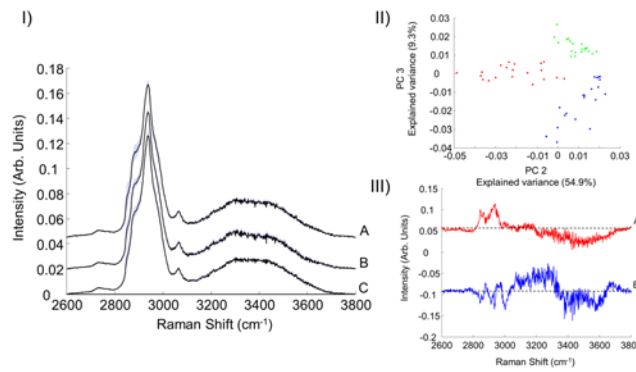

Figure 5: I- (I) Mean and SD spectra of the nucleus A – normal B – DOK C – SCC4 (II) PCA scatter plot differentiating the Raman spectra of the nucleus of the primary (normal) cell line (red), from those of the abnormal cell lines (SCC-4 (blue) and DOK (green)). (III) A:Red =Loading of PC2 and B:Blue = Loading of PC 3. Loadings are off set for clarity.

For the nucleolar data of Figure 6, a very similar patter to the nuclear region is seen, the loadings again indicating that the normal cells are relatively sparse, in terms of lipid and protein content, but rich in unbound water (PC2), whereas PC3 again indicates that the SCC are associated with more free water characteristics.
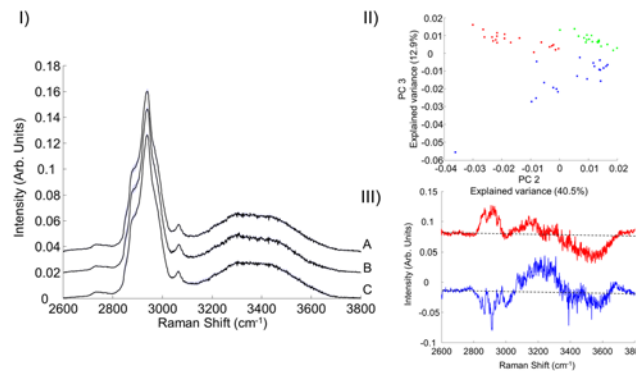

Figure 6: (I) Mean and SD spectra of the nucleolus A – normal B – DOK C – SCC4 (II) PCA scatter plot differentiating the Raman spectra of the nucleolus of the primary (normal) cell line (red), from those of the abnormal cell lines (SCC-4 (blue) and DOK (green)). (III) A: Red =Loading of PC2 and B:Blue = Loading of PC 3. Loadings are off set for clarity.

3.3. Classification

While PCA is a useful tool in terms of verifying the biochemical differences between the respective subcellular regions of the three cell types, PCA-FDA is a more powerful tool to demonstrate the classification/diagnostic potential of the technique. The results of the PCA-FDA analysis of the three way comparison of the different cell types, after the removal of the background of the water are shown in Table 3. For cytoplasm, 99.9% sensitivity and 97.4% specificity for the classification of the cell type is achieved, for the nucleus, 100% sensitivity and 99.1% specificity, and for the nucleoli 100% sensitivity and 95.4% specificity. The most common misclassification, for all subcellular regions, is of normal cells as DOK, indicating that the premalignant pathology is closest in biochemical profile to the normal state.

Table 3: PCA-FDA analysis of the different cell lines after water subtraction

A

		Cytoplasm		
		Predicted groups		
		NC	DOK	SCC
Observation (700 spectra / cell type)	NC	682	16	2
	DOK	0	700	0
	SCC	1	0	699
			Sensitivity	99.9%
			specificity	97.4%

B

		Nucleus		
		Predicted groups		
		NC	DOK	SCC
Observation (700 spectra / cell type)	NC	694	5	1
	DOK	0	700	0
	SCC	1	0	700
			Sensitivity	100%
			specificity	99.1%

C

		Nucleolus		
		Predicted groups		
		NC	DOK	SCC
Observation (700 spectra / cell type)	NC	668	32	0
	DOK	0	700	0
	SCC	0	0	700
			Sensitivity	100%
			specificity	95.4%

4 Discussion

In the analysis of the cellular regions of the normal cells, the spectra of the cytoplasmic regions are clearly distinguishable from the spectra of the combined nuclear regions based on their differing biochemical make up, lipidic features being more prominent in the cytoplasm, proteinic in the nuclear region. A direct comparison of the nucleolar and nuclear regions shows them to be differentiable based on a similar difference in lipidic/proteinic content. Similar to the study of the fingerprint region of the same cell type by Carvalho et al., [3] the study shows the potential of Raman microspectroscopic analysis to biochemically profile the cells at subcellular level.

The true potential of the technique is in differentiating between cell types, and in this context the ability to compare at a subcellular level has potential advantages. Using the fingerprint region, PCA of Raman microspectroscopic was able to differentiate abnormal (DOK and SCC) from the normal cells, based on the nucleolar region, but only partially differentiated the DOK and SCC cells. Differentiation based on the spectroscopic data from the nuclear and cytoplasmic regions was less apparent, consistent with other studies demonstrating that the nucleolar region has the highest diagnostic potential, according to the fingerprint region [3].

Using the high wavenumber region, a clearer degree of differentiation is seen in all cellular regions, although the differentiating profiles are less rich in biomolecular information. Notably, however, significant contributions to the discriminating spectral profiles are in the spectral region >3000cm^{-1}, which can derive from non-aqueous OH groups, N-H stretching vibrations, Amide A, bands of proteins and bound or semi-bound water vibrations.

Although samples have been formalin fixed, and measured in water immersion, it is possible that residual water peaks after subtraction, still contributing to the discrimination between cell lines, may be due to membrane bound or intracellular water [9, 24, 25] and thus may have profiles characteristic of the pathological state. It has been reported that small changes in cell water content, mediated by hormones or by the osmotic effect of ions or organic osmolytes, can play an important role in the regulation of cell metabolism and gene expression, and in the pathology of various diseases. There is also considerable evidence that cell water content may be a major factor in the mechanism of carcinogenesis [26].

Despite remarkable advances in understanding the peculiar properties of water, relatively little is known about the fate of the surface and intracellular water under the actual crowded conditions of biological tissue [24]. Carvalho et al., [9] have stated that the question concerning confined water is quite complex in terms of disease and non disease samples. In terms of the nucleus and nucleolus, water plays an important role in protein folding/misfolding, protein binding to specific DNA, and many other fundamental biological processes, where the balance between the flexibility of a given protein and DNA sequences and the amount of water released from the interface is essential. The internal molecular flexibility in the proteins necessary for biological activity depends on the level of hydration. Furthermore, in this study they showed that the biochemical information imprinted in the IR and Raman spectra of the biological tissue is not limited to lipids, proteins, DNA and RNA as potential biomarkers, OH stretching modes of water in the cellular environment of cancerous human breast tissue are sensitive indicators of the pathological changes in the tissue. To explain the vibrational properties of the interfacial water in normal and cancerous tissue, a characterization of the biological 'surface' to which the water molecules are associated is required, but this characterization is extremely difficult.

In a model for hydrated biological tissues, Sato et al., [25] concluded that protein-water interactions play a critical role in the function of several biomacromolecular systems including collagen tissues. Small changes in structure and dynamical behaviour of water molecules at the peptide/water interface can effectively change both the structure and dynamics of the protein function. Thus, it was realised that the presence of different water clusters plays an important role in collagen properties.

Thus, it is maybe not surprising that the water content in cells can contribute to differentiating pathological states in cells and tissue. Barroso et al. [10] have demonstrated discrimination between oral cancer and healthy tissue based on water content of freshly excised tongue specimens, determined by Raman spectroscopy in the high wavenumber region. McIntyre, [26] discussed the hypotheses that increased cell hydration causes cancer not only by promoting cell division and oncogene expression, but also by inactivating genes inducing cell differentiation, and by preventing apoptosis. Conversely, factors that decrease cell hydration prevent cancer by inhibiting cell division and oncogene expression, while activating genes inducing cell differentiation, and by promoting apoptosis, also the degree of malignancy increasing with the degree of cell hydration. The current study suggests that it is not only the water content, but also the structural characteristics of the water at a cellular level which can contribute to the differentiation between pathological states.

5 CONCLUSION

In conclusion, the results of the present work demonstrate the capacity of the high-wavenumber region (2,800– 3,600 cm^{-1}) to discriminate between the three subcellular regions (cytoplasm, nucleus and nucleolus). Significantly, it was also possible to discriminate the cell types representative of normal, pre-malignant and malignant pathologies with high levels of specificity and sensitivity by analysis the spectral profiles of each of the subcellular regions. Misclassification was largely of normal as pre-malignant, consistent with a similarity of biochemical profiles of these cell types. Notably, for all subcellular regions, features in the region 3000-3600cm-1 featured strongly in the differentiation of the cell types. Although the cells had been formalin fixed and were measured in immersion, and the data was subsequently preprocessed by (bulk) water subtraction, residual OH features which may be attributable to intracellular or membrane bound water seem to contribute significantly to the diagnostic potential.

REFERENCES

1. Holmstrup P, Vedtofte P, Reibel J, Stoltze K. Oral premalignant lesions: is a biopsy reliable? J Oral Pathol Med. 36(5):262-6 (2007).

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

3. Carvalho LF, Bonnier F, O'Callaghan K, O'Sullivan J, Flint S, Byrne HJ et al. Raman micro-spectroscopy for rapid screening of oral squamous cell carcinoma. Exp Mol Pathol. 21(3), 502-509 (2015).

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

5. Nawaz H, Bonnier F, Meade AD, Lyng FM, Byrne HJ. 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 (2011).

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

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

8. Movasaghi Z, Rehman S, Rehman IU. Raman spectroscopy of biological tissues. Appl Spec Rev. 42, 493 (2007).

9. Carvalho LFCS, Sato ET, Almeida JD, Martinho HS. Diagnosis of inflammatory lesions by high-wavenumber FT-Raman spectroscopy. Theo Chem Acc. 130 (6) 1221-1229 (2011).

10. Barroso EM, Smits RW, Bakker Schut TC, ten Hove I, Hardillo JA, Wolvius EB, Baatenburg de Jong RJ, Koljenović S, Puppels GJ. Discrimination between oral cancer and healthy tissue based on water content determined by Raman spectroscopy. Anal Chem. Feb 17;87(4):2419-26 (2015).

11. Barroso EM, Smits RW, van Lanschot CG, Caspers PJ, Ten Hove I, Mast H, Sewnaik A, Hardillo JA, Meeuwis CA, Verdijk R, Noordhoek Hegt V, Baatenburg de Jong RJ, Wolvius EB, Bakker Schut TC, Koljenović S, Puppels GJ. Water Concentration Analysis by Raman Spectroscopy to Determine the Location of the Tumor Border in Oral Cancer Surgery. Cancer Res. Oct 15;76(20):5945-5953 (2016).

12. Santos IP, Caspers PJ, Bakker Schut TC, van Doorn R, Noordhoek Hegt V, Koljenović S, Puppels GJ. Raman Spectroscopic Characterization of Melanoma and Benign Melanocytic Lesions Suspected of Melanoma Using High-Wavenumber Raman Spectroscopy. Anal Chem. Aug 2;88(15):7683-8 (2016).

13. 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

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

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

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

17. das Chagas E Silva de Carvalho LF, Pereira TM, Magrini TD, Cavalcante AS, da Silva Martinho H, Almeida JD Optical diagnosis of actinic cheilitis by infrared spectroscopy. Photodiagnosis Photodyn Ther. Dec;16:27-34 (2016).

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

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

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

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

22. Bertrand D, Courcoux P, Autran JC, Meritan R, Robert P. Stepwise canonical discriminant analysis of continuous digitalized signals: Application to chromatograms of wheat proteins. J Chemometrics 1990.

23. Vyumvuhore R, Tfayli A, Duplan H, Delalleau A, Manfait M, Baillet-Guffroy A. Effects of atmospheric relative humidity on Stratum Corneum structure at the molecular level: ex vivo Raman spectroscopy analysis. Analyst, 138, 4103, (2013).

24. Abramczyk H, Brozek-Pluska B, Krzesniak M, Kopec M, Morawiec-Sztandera A. The cellular environment of cancerous human tissue. Interfacial and dangling water as a ''hydration fingerprint''.Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 1386-1425, (2014).

25. Sato ET, Rocha RA, Carvalho LFCS, Almeida JD, Martinho HS. Molecular model for hydrated biological tissues. Phys Rev E 91, 066610 (2015).

26. Mclntyre GI. Cell hydration as the primary factor in carcinogenesis: a unifying concept. Medical Hypotheses. 66, 518-526 (2006).