---

**Article:**
Mian, S.A., Yorucu, C., Ullah, M.S. et al. (2 more authors) (2017) Raman spectroscopy can discriminate between normal, dysplastic and cancerous oral mucosa: a tissue-engineering approach. Journal of Tissue Engineering and Regenerative Medicine , 11 (11). pp. 3253-3262. ISSN 1932-6254

https://doi.org/10.1002/term.2234

---

Tissue Engineering and Regenerative Medicine

# Raman spectroscopy can discriminate between normal, dysplastic and cancerous oral mucosa: a tissue-engineering approach.



| | |
|---|---|
| Journal: | *Journal of Tissue Engineering and Regenerative Medicine* |
| Manuscript ID: | Draft |
| Wiley - Manuscript type: | Research Article |
| Date Submitted by the Author: | n/a |
| Complete List of Authors: | Mian, Salman; Univeristy of Sheffield, Department of Materials Science and Engineering<br>Ullah, Muhammad; Univeristy of Sheffield, Department of Materials Science and Engineering<br>Yorucu, Ceyla; Univeristy of Sheffield, Department of Materials Science and Engineering<br>Rehman, Ihtesham; Univeristy of Sheffield, Department of Materials Science and Engineering<br>Colley, Helen; University of Sheffield, School of Clinical Dentitsry |
| Keywords: | Tissue engineering, Oral mucosa, Raman spectroscopy, Squamous cell carcinoma, Three-dimensional, Diagnostics |
| | |

# Raman spectroscopy can discriminate between normal, dysplastic and cancerous oral mucosa: a tissue-engineering approach.

Running title: Raman Spectroscopy discrimination of tissue-engineered oral lesions

Salman A. Mian [1,2], Muhammad Saad Ullah[1,2], Ceyla Yorucu[1], Ihtesham U. Rehman[1] and Helen E. Colley[2*]

[1]Department of Materials Science and Engineering, University of Sheffield, Sheffield, S3 7HQ, UK.

[2]School of Clinical Dentistry, University of Sheffield, Sheffield, S10 2TA, UK.

[*] **Corresponding author:**

Dr. Helen Colley

School of Clinical Dentistry

University of Sheffield

Claremont Crescent

Sheffield S10 2TA

United Kingdom

Tel: +44 (0) 114 2717966

Fax: +44 (0) 114 271 7894

Email: h.colley@sheffield.ac.uk

**Abstract**

Head and neck cancer (HNC) is the sixth most common malignancy worldwide. Squamous cell carcinoma, the primary cause of HNC, evolves from normal epithelium through dysplasia before invading the connective tissue to form a carcinoma. However, less than 18% of suspicious oral lesions progress to cancer with diagnosis currently relying on histopathological evaluation, which is invasive and time consuming. A non-invasive, real-time, point-of-care method could overcome these problems and facilitate regular screening. Raman spectroscopy can non-invasively provide information regarding the biochemical composition of tissues. In this study, Raman Spectroscopy was assessed for its ability to discriminate between normal, pre-malignant and head and neck squamous cell carcinoma (HNSCC). Tissue engineered models of normal, dysplastic and HNSCC were constructed using normal oral keratinocytes, dysplastic and HNSCC cell lines and their biochemical content predicted by interpretation of spectral characteristics. Spectral features of normal models were mainly attributed to lipids, whereas, malignant models were observed to be protein dominant. Visible differences between the spectra of normal, dysplastic and cancerous models, specifically in the bands of amide I and III were observed. Normal mucosal models displayed a sharp and weak lipid peak at 1667cm$^{-1}$ whereas HNSCC spectra showed a broad and strong amide I peak at this wavenumber. A shift at 2937cm$^{-1}$ was only observed in DOK, differentiating them from the other tissue types. Multivariate data analysis algorithms successfully identified subtypes of dysplasia and cancer, suggesting that Raman spectroscopy can be used to discriminate between normal and malignant tissues.

**Key words:** Tissue engineering, oral mucosa, Raman spectroscopy, squamous cell carcinoma, diagnostics.

## 1. Introduction

Head and Neck cancer (HNC) is the sixth most common malignancy worldwide with approximately 500 000 new cases per year. Prognosis is poor with only a 50% 5 year survive rate (Johnson et al., 2011, Argiris et al., 2008). Squamous cell carcinoma, the primary cause of HNC, is typically preceded by pre-malignant, cellular abnormalities in the epithelium, called dysplasia. A dysplastic epithelium evolves through mild, moderate and severe stages before invading the connective tissue to form a carcinoma (Napier and Speight, 2008). The low survival rate can in part be attributed to the difficulty in diagnosing the disease at an early stage (Epstein et al., 2002). Currently, visual inspection of the oral cavity is the first line of diagnostic screening, followed by histopathological evaluation of biopsies that are surgically removed from any suspicious lesion. Although histopathological evaluation is at present the most accurate and reliable method for diagnosis it has several limitations. Surgical biopsies are invasive and take time to analyse, causing anxiety for patients as well as being time-consuming and causing delays in treatment (Stefanuto et al., 2014). The need for a scalpel biopsy also reduces the rate at which histopathological evaluation is performed and reduces the frequently of monitoring and screening of suspected lesions. Furthermore, histopathological analysis is associated with inter-observer variability (Reibel, 2003). For all these reasons it is clear that there is the need for a

non-invasive, real-time, point-of-care method to accurately detect and diagnose oral cancerous changes at an early stage.

During cancer progression biochemical changes occur within cancer cells altering the levels of nucleic acids, proteins, lipids and carbohydrates and that may serve as potential markers for disease monitoring (Stone et al., 2002, Stone et al., 2004). Recently optical techniques such as elastic light scattering, fluorescence and Raman spectroscopy have been studied for their ability to characterize the biochemical makeup of biological tissues and therefore their potential as a non-invasive, diagnostic tool for cancer detection (Movasaghi et al., 2012). Among these Raman spectroscopy in particular has shown encouraging results for both the characterisation and detection of disease in different biological tissues including the gastrointestinal tract (Teh et al., 2010), breast tissue (Rehman et al., 2007), lungs (Short et al., 2011), larynx (Stone et al., 2000), testicular (Movasaghi et al., 2012) and head and neck including the oral cavity (Devpura et al., 2012, Valdés et al., 2014, Malini et al., 2006). Raman spectroscopy utilizes a technique of inelastic light scattering, which is characterised by a shift in the wavelength of incident light that occurs due to the specific molecular vibrations of the biological tissues. The Raman spectrum obtained can provide information regarding biochemical configuration and conformations related to the key biological components of tissues i.e. lipids, proteins and nucleic acid and help to distinguish between different tissue types (IU. Rehman et al., 2012). In this study, we have analysed 3D tissue engineered models of normal, dysplastic and cancerous oral mucosa using Raman spectroscopy to determine the potential of the technique to distinguish normal oral mucosa from pre-malignant and cancerous oral mucosa.

## 2. Materials and Methods

All materials used in this study were purchased from Sigma-Aldrich (Poole, Dorset, UK) unless otherwise stated and used as of the manufacturers' instructions.

### 2.1 Cell Culture

The following cells were used in this study: normal oral fibroblasts (NOF) and keratinocytes (NOK) (ethical approval number 09/H1308/66) isolated as previously described (Colley et al., 2011). DOK (ECACC, Health Protection Agency Culture Collections, Salisbury, UK) derived from a dysplastic oral lesion isolated from the dorsal tongue of a 57-year-old male (Chang et al., 1992). D19 and D20 (generously provided by Dr. Keith Hunter, School of Clinical Dentistry, University of Sheffield) were derived from lateral tongue dysplasia (McGregor et al., 2002, McGregor et al., 1997). The head and neck squamous cell carcinoma (HNSCC) cell lines Cal27 and SCC4 (American Tissue Culture Collection, Manassas, VA, USA) were both isolated from SCC of the tongue (Jiang et al., 2009). FaDu (LGC Promochem, Middlesex, UK) were derived from a hypopharyngeal SCC (Rangan, 1972). NOK, D19 and D20 were cultured in Green's media as previously described (Colley et al., 2011). Cal27 and DOK were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS) (v/v). Cell culture medium for DOK was supplemented with 0.5 µg/ml hydrocortisone. FaDu cells were cultured in RPMI-1640 supplemented with 10% FCS (v/v) and 100 IU /ml penicillin and 100 µg/ml streptomycin. SCC4 cells were cultured in DMEM:Ham's F12 (1:1), 10% FCS (v/v), 100 IU /ml penicillin, 100 µg/ml streptomycin and 0.5 µg/ml hydrocortisone. Media was changed every 2-3 days.

### 2.2. 3D Tissue Engineered Models of the Oral Mucosa

Tissue engineered models of the oral mucosa were generated as previously described (Colley et al., 2011). Briefly, de-epidermised dermis (DED) was prepared from cadaveric skin (EuroSkin Bank, Beverwijk, Netherlands). The skin was decelluarised by incubating in 1M NaCl at 37°C for 24 hours before carefully scraping the dermis to remove the epithelium. Processed DED was stored in DMEM at 4°C until use. To produce the tissue engineered models processed DED was cut into 2 $cm^2$ squares and placed into a well of a 6 well plate. A chamfered surgical stainless steel ring (8 mm internal diameter) was placed onto the DED and gentle pressure applied to create a liquid tight seal. To generate cancer or dysplastic models, HNSCC (Cal27, SCC4, FaDu) or dysplastic cells (DOK, D19 and D20) were seeded within the stainless steel rings ($2.5x10^5$) in co-culture with NOF ($5x10^5$). To generate normal oral mucosa models, NOF ($5x10^5$) and NOK ($1x10^6$) were seeded within the stainless steel rings. After three days all models were raised to an air to liquid interface (ALI) before culturing for a further 14 days. Media was changed 2-3 times per week.

### 2.3. Histological analysis

At 14 days mucosal models were fixed in 10 % buffered formalin for at least 24 hours before histological processing using a Leica TP1020 bench top processor (Leica microsystems, Milton Keynes, UK). Samples were paraffin wax embedded and two sequential sections of 4 µm and 20 µm thicknesses cut using a manual rotary microtome (Leica microsystems, Milton Keynes UK) and then mounted onto SuperFrost® Plus glass slides (Thermo Fisher Scientific, MA, USA). Four µm sections

were stained with hematoxylin and eosin (H&E) using a Shandon linear staining machine (Thermo Fisher Scientific, MA, USA). Twenty µm thick sections were cut and de-waxed according to the method of Mian *et al* (2014). Briefly, samples were placed in xylene for 30 minutes followed by immersion in 50%, 70% and 100% ethanol for 5 minutes each before Raman analysis (Mian et al., 2014).

**2.4. Spectroscopic instrumentation and measurements**

Raman spectra were recorded using a DXR Raman microscopic system (Thermo Fisher Scientific, MA, USA). The system was equipped with an Olympus BX51 optical microscope and a 532 nm diode-pumped solid-state laser excitation source. A 50X long working distance objective was used to focus the excitation laser beam to a spot size of 1.1 µm on the sample tissues with a laser power of 10 mW. The allowed spectral range was adjusted for 600 $cm^{-1}$ to 3400 $cm^{-1}$, with an estimated spectral resolution of 5.5 - 8.3 $cm^{-1}$. Exposure time was set to 30 seconds and 5 sample exposures were accumulated to improve the signal to noise ratio. A new location of the tissue was exposed for each spectra acquisition. The spectra obtained were analysed and processed using OMNIC Atlµs™ software suit (Thermo Fisher Scientific, MA, USA).

**2.5. Raman Data Processing**

Raman data from the region 600 $cm^{-1}$ to 3400 $cm^{-1}$ was analysed. Spectra were collected from three different models of tissue-engineered normal, dysplastic and cancerous oral mucosa. A total of 225 spectra were collected for this study (D20 = 19, DOK = 20, D19 = 30, Cal27 = 55, SCC4 = 51, FaDu = 45 and normal oral mucosa NOM =

35). Additional spectra were collected from thick cancer epithelia in order to acquire maximum biochemical information. Raman data was collected within the epithelia. A mean spectrum for each of the samples was produced for the purpose of evaluation and comparison. Smoothing and baseline correction for the spectra was performed using OMNIC Atlμs™ software suit (Thermo Scientific, Madison, WI, USA).

**2.6. Data Analysis**

Both univariate and multivariate data analysis approaches were employed. Peak height analysis was performed over vital contributions from phenylalanine, tryptophan, nucleic acid, amide I and amide II using OMNIC Atlμs™ software suit. Chemo-metric methods were used to quantify the spectral differences of various groups present in the data. These methods were performed over Unscrambler X 10.2 software, purchased from Camo software (Oslo, Norway). Principal component analysis (PCA) was performed over the complete spectral range (600 $cm^{-1}$ – 3400$cm^{-1}$), amide I (1550 $cm^{-1}$ - 1750 $cm^{-1}$) and amide III (1200 $cm^{-1}$ - 1400 $cm^{-1}$) regions of normal, dysplastic and cancerous data set. The maximum variance between the data was observed within first three principal components (PC's). Cluster analysis (CA) was performed over full spectral range (600 - 3400 $cm^{-1}$) by Ward's method using squared Euclidean distance. Linear Discriminant Analysis (LDA) was also performed over the entire spectral range (600- 3400 $cm^{-1}$). Pre-processing comprised of baseline correction and Unit Vector Normalisation. 4 samples from each group were left out at each pass for prediction until a total number of 20 spectra were predicted.

## 3. Results and discussion

### 3.1. Histological classification

Dysplastic oral lesions and their malignant progression are currently diagnosed by microscopic evaluation of cytological and architectural changes within the tissue yet it is well acknowledged that evaluation of lesions is subjective with considerable inter- and intra-observer variation (Speight et al., 2015). Furthermore currently there are no clinically used molecular biomarkers to assist pathologists in the prediction of malignant transformation in oral dysplastic lesions (Speight, 2007). These limitations have led to an increase in research into spectroscopy techniques as potential diagnostic tools. Raman spectroscopy is capable of providing real-time, non-invasive, biochemical information of tissues and could be used as an adjutant to current histopathological evaluation (Stone et al., 2004). 3D tissue-engineered models of normal, dysplastic and cancerous oral mucosa were generated and subject to routine H&E staining for evaluation. Histological analysis of tissue-engineered normal oral mucosa models revealed a stratified epithelium and a fibroblast-populated dermis closely resembling native oral mucosa (Fig. 1a-c). Models produced using dysplastic cell lines (DOK, D19 and D20) produced an epithelium histologically typical of dysplastic lesions with bulbous rete processes, abnormal cytological changes including the presence of mitotic figures in the upper epithelial layers and abnormal keratinocyte maturation (Fig. 1d-f). Oral cancer models generated using HNSCC cell lines (Cal27, FaDu and SCC4) showed typical features common to the pathology of cancerous transformation with abnormal epithelial maturation, resulting in a disordered and thick

mucosal epithelium resembling malignant transformation typically seen *in vivo* although not yet invasive into the underlying connective tissue (Fig. 1g-i).

### 3.2 Raman spectral analysis

Raman spectra were obtained from 20 µm sections of normal, dysplastic and cancerous oral mucosal models. The averaged spectra are presented in Figure 2. Spectral differences are evident in both the fingerprint (600 $cm^{-1}$ to 1800 $cm^{-1}$) and high wavenumber compartments (2800 $cm^{-1}$ to 3400 $cm^{-1}$) with Raman peaks identified attributed to differences in functional groups and biochemical variations including the abundance of proteins, lipids and nucleic acids. Due to the complex nature of biological tissues Raman peak frequencies often overlap making it difficult to assign a particular molecule or single biochemical entity. However, it has been identified that certain biochemical components can produce relatively sharp and characteristic bands that can be used to highlight dissimilarities between different tissue types (Devpura et al., 2012).

Noticeable features of normal oral mucosa spectrum include a sharp C=C lipid stretch at 1653 $cm^{-1}$ (amide I), no prominent contributions at 1245 $cm^{-1}$ (amide III) and 2881 $cm^{-1}$ (lipids) which are attributed to $CH_2$ wagging and $CH_2$ asymmetric stretch, respectively (Fig. 2a). Raman peaks observed in the spectra from dysplastic and oral cancer models represent amide I at 1667 $cm^{-1}$ (DOK, D19, FaDu and SCC4), $CH_2$ bending at 1447 $cm^{-1}$, broad peaks of amide III between 1300-1200 $cm^{-1}$ and a sharp phenylalanine peak at 1003 $cm^{-1}$, all of which signify a major involvement from proteins (Movasaghi et al., 2007, Malini et al., 2006) (Fig. 2b-c).

In biological tissues broader peaks indicate protein dominance as compared to narrow peaks which indicate lipid dominancy. The characteristic peaks observed in the spectrum from tissue-engineered normal oral mucosa models can largely be assigned to tissue lipids with a minimum contribution from protein (Devpura et al., 2012, Movasaghi et al., 2007), consistent with reports that Raman spectra of normal oral tissues generally arise from the surface lipid bilayer and hence give rise to lipid dominated spectrum (Venkatakrishna et al., 2001, Krishna et al., 2004). In contrast cancerous tissues are characterized by large amounts of surface proteins, which provide a protein dominated spectrum, as observed here, and forms the basis of distinguishing between normal and malignant tissues (IU. Rehman et al., 2012, Malini et al., 2006).

In the amide I region normal, D20 and Cal27 models show a shift from 1667 $cm^{-1}$ to 1653 $cm^{-1}$, which differentiates them from the rest of the dysplastic and cancerous tissues, whilst a moderate broadening of the same peak differentiates these models from tissue-engineered normal oral mucosa models (Fig. 2a-b)**.** Furthermore, spectral variations at the amide I region also separates D20 from D19 and DOK as well as Cal27 from FaDu and SCC4 (Fig. 2c). A sharp but weak peak is observed at 1295 $cm^{-1}$ (amide III) only in the spectra of D19 and SCC4 which can be assigned to $CH_2$ deformation (Movasaghi et al., 2007) (Fig. 2c).

A prominent peak at 1100 $cm^{-1}$ is noticeable in the Cal27 spectrum and associated with nucleic acid $PO_2^-$ or C-N stretch which discriminates Cal27 from normal, dysplastic and the other two cancer models (Fig. 2b-c).

Peak at 2881 $cm^{-1}$ is an assignment of $CH_2$ asymmetric stretch of lipids which is absent in the spectra of normal, D20 and FaDu models (Fig. 2a-b). A unique shift of the peak is also observed at 2937 cm-1 in DOK spectra, attributed to $CH_2$ asymmetric stretch, whereas in the rest of the normal, dysplastic and cancer tissue spectra the $CH_2$ asymmetric stretch was observed at 2932 cm-1. This peak clearly separates DOK from all other models (Fig. 2a and c).

Spectral variations particularly at the amide I, amide III and high wavenumber compartment clearly distinguish normal models from D19 and DOK whilst D20 appears more similar to NOM in these regions. A narrow band was noticed in the amide I region (1653 cm-1) in normal and D20 tissue spectra which was broader and shifted to 1667 $cm^{-1}$ in the spectra of D19 and DOK, hence differentiating between them. A broad peak centered at 1667 cm-1 was evident in the FaDu model spectra, separating it from normal, dysplastic and the other cancer models.

The band at the 642 $cm^{-1}$ is associated with C-C stretching and twisting modes of proteins (tyrosine) (Devpura et al., 2011). The bands at 724 $cm^{-1}$ and 780 $cm^{-1}$ regions are an assignment of C-N nucleotide peak or lipid/DNA (Devpura et al., 2011) and ring breathing mode of DNA/Uracil, respectively. Whilst the band assignment at 828 $cm^{-1}$ can be attributed to tyrosine/phosphodiester (Devpura et al., 2011) or O-P-O backbone stretching of DNA (Su et al., 2012). In addition to altered levels of proteins and lipids, the biochemistry of cancer progression can also be linked with abnormal nucleic acid synthesis and metabolism (Heimann et al., 1991, DeBerardinis et al., 2008). In the spectra presented here at we also observe increased levels of DNA/Uracil (780 $cm^{-1}$) in

the cancer models, and to a lesser extent in the dysplastic models, compared to normal tissue. Specifically Cal27 models show a broad and intense peak at 1100 $cm^{-1}$ associated with nucleic acid $PO_2^-$ or C-N stretch. Changes in the relative intensity as well as a shift of these peaks can be related to altered levels of nucleic acids in malignancy (Manoharan et al., 1996, Rehman et al., 2007, Valdés et al., 2014, Huang et al., 2003).

In the cancer models an increase in the intensity of the peak at the band 1582 cm-1 was observed and assigned to C=C bending of phenylalanine, tryptophan and tyrosine (Fig. 2b) (Su et al., 2012, Movasaghi et al., 2007, IU. Rehman et al., 2012). Previous studies have linked increased levels of tryptophan to malignant tissue when compared to normal tissue (Devpura et al., 2012),(Tankiewicz et al., 2006). Here, we report a similar pattern for tryptophan from the tissue-engineered models with the peak attributed to tryptophan, observed at 1582 $cm^{-1}$ (Devpura et al., 2012, Movasaghi et al., 2007), showing a significant increase in intensity as the tissue changes from normal to malignant. Tryptophan is an essential amino acid required for different metabolic processes and is important for rapidly dividing malignant cells (Tankiewicz et al., 2006) which might indicate the reason why it is more abundant in cancerous tissue.

In this study buccal derived keratinocytes were used to generate normal oral mucosal models whilst all the dysplastic and HNC models were generated from cell lines derived from the tongue with the exception of the FaDu cell line which is derived from a hypopharyngeal squamous cell carcinoma. Previously studies have found that Raman spectroscopy is able to discriminate between oral mucosa derived from different

anatomical locations within the oral cavity (Bergholt et al., 2012) and may suggest why models constructed using the FaDu cell line display unique spectral features in the amide I, amide III and lipid regions, differentiating them from all of the other models (fig 2b-c).

### 3.3. Principal component analysis

The spectral variations between different tissue types can be understood by using chemometric analysis methods that enable interpretation of spectral differences in relation to biochemical changes. PCA was performed over two spectral ranges; amide I (1515 $cm^{-1}$ - 1770 $cm^{-1}$) and amide III (1220 $cm^{-1}$ - 1300 $cm^{-1}$). A comparison was made between normal and dysplastic (fig3a), normal and cancer (fig3b) and dysplastic and cancer (fig.3c).

The amide I band shows favorable separation of normal oral mucosa models from all dysplastic models; PC1 separates normal models from D19 and DOK whilst PC2 splits D20 from the normal models (Fig. 3a). However, D19 and DOK models continue to cluster together up to PC8 accounting for 99.8% of the variance (data not shown). By comparison, differences within the dysplastic groups appear to be much greater in the amide III region and the scores plot for this region shows good separation between all groups (Fig. 3a). Whilst the clusters are more widely spread, loadings plots for PC1 and PC2 (data not shown) suggest the contrast in the 1295 $cm^{-1}$ lipid peak and contribution from intensity of the 1220 - 1270 $cm^{-1}$ band to be a major influence. PCA demonstrates that tissue-engineered normal oral mucosa models can be discriminated from dysplastic models using these regions. Again, D20 showed more similarities with

normal tissues, with their clusters forming close to each other as compared to D19 and DOK. This suggests that biochemical characteristics of D20 are more similar to normal tissue compared to D19 and DOK. Furthermore, the separation of D19 from DOK in the amide III region suggests that each of the dysplastic models have compositional differences. Though all models are derived from dysplastic cell lines the analysis indicates separate and disparate behavior by each of the different cell types and may be attributed to patient variability or different stages of dysplasia (mild, moderate or severe).

Within the normal and HNC models (Fig. 3b), the sample groups can be well discriminated using PCA. In the amide I model, PC1 separates FaDu and SCC4, whereas PC2 suggests these two cancer groups also express some similarities. Normal oral models and Cal27 show similar differences with respect to variations explained by PC2. In the amide III model, PC1 and PC2, accounting for 98% of the variance, discriminates FaDu from normal oral models as well as the other cancer models. The corresponding loadings plot to PC1 suggests positive loadings for the 1220-1270 $cm^{-1}$ band relative to the 1270-1320 $cm^{-1}$ in FaDu compared to other models. Amide III of the normal and cancer PCA plot differentiated FaDu from the rest of cancers as well as from normal models reconfirming that differences in spectra from various intra-oral locations needs to be taken into consideration. Tight cluster formations were also observed for SCC4 and Cal27, which were separated both from each other and from normal tissue.

Differentiation between dysplastic lesions and cancerous tissue and predicting malignant transformation can be most challenging in terms of diagnosis. Analysis of

the amide I and III regions gave encouraging separation of D20 (dysplasia) from the cancer groups whereas specifically in the amide III region DOK was discriminated from cancer models (Fig. 3c). The PCA also highlights the disparity that exists within tissues of the same subtype of D19 (dysplasia) and FaDu (Cancer). The clear separation between and within dysplastic and cancer lines suggests substantial differences in the biochemical composition of the tissue (Kamath and Mahato, 2007).

### 3.4. Cluster analysis

Cluster analysis was performed over the complete spectral range (600-3400 $cm^{-1}$). Between the normal and dysplastic tissue-engineered models two main clusters were formed; one exclusively containing all of the dysplastic models with a separation cluster for the normal models (fig. 4a). Cluster analysis for normal and cancer models revealed that the HNC FaDu is distinct to normal oral mucosa as well as the two other HNC models whilst Cal27 is equally distant to NOM as to SCC4 (fig. 4b).

Figure 4c and 4d show cluster analysis of dysplastic and cancer models and all models respectively. Whilst each of the different tissues form well defined clusters, the partitioning does not create main subgroups for normal, dysplastic and cancer samples. D19 appears closer to SCC4 and FaDu than to the other dysplastic models. Dysplastic models generated using the DOK cell line shows the least inner group variability on the relative distance scale which suggests the tissue is more homogeneous compared to the rest.

Cluster analysis formed on the basis of molecular differences between the data set showed separate branches in the dendogram for normal, dysplastic and cancer tissue models as importantly each subtype was grouped separately in the main branch. Both the dysplastic and cancer revealed discrete subsets representing each of the six cell types used independently and suggests a high sensitivity of the technique for the identification of different subtypes. This partitioning indicates that the biochemical variations present within dysplastic and cancer tissues could be identified and clustered separately (Bigio et al., 2000) (Fig. 4c and d).

**3.5. Linear discriminant analysis**

Tissue sections from normal, dysplastic and cancer models were blinded and subject to linear discriminant analysis (LDA) to test for the specificity and sensitivity of Raman spectroscopy to discriminate and thereby correctly classify tissue samples as normal, dysplastic or cancer. LDA predictions of normal versus cancer classified 14 out of the 20 normal mucosa models correctly with 6 misclassified as cancer; giving an overall specificity of 70%. Within the same comparison all cancerous models were classed correctly with a sensitivity of 100% (fig.5a). Analysis of dysplastic versus cancer classified 54 out of the 60 dysplastic mucosa models correctly, 6 were misclassified as cancer resulting in a sensitivity of 90%. 98% sensitivity was achieved for cancer models (fig.5b).

Raman spectroscopy predictions of blinded samples against known models of normal, dysplastic or cancer gave encouraging results with a high percentage of accuracy in categorising tissues. Comparison between all types of models predicted normal models

with a specificity of 75%. Out of 60 dysplastic mucosa models tested 54 were predicted correctly with 6 misclassified as cancers with a sensitivity of 90%. 59 models were predicted correctly as cancer with only one sample misclassified as dysplastic giving a sensitivity of 98% (fig. 5c).

LDA was also completed to test for sensitivity and specificity of Raman spectroscopic discrimination of blinded samples and its ability to correctly classify original cell type (Fig.6a). 100% sensitivity and specificity for normal and cancerous models was achieved. Furthermore, not only were normal models successfully discriminated from cancer models but also cancer models were successfully separated from one another with 100% sensitivity (fig. 6a).

Analysis between different dysplastic and cancer models predicted 100% sensitivity to D19, Cal27 and SCC4 models and only a small percentage of misclassifications were seen between FaDu, D20 and DOK with a sensitivity of 70%, 85% and 90% respectively (fig. 6b).

These findings are consistent with other reports (Singh et al., 2012, Stone et al., 2004). Malini *et al*. (2006) investigated the ability to discriminate between normal and cancerous tissue but also extended their data set to include inflammatory and premalignant lesions. PCA results concluded that it was possible to separate normal from cancer but poor discrimination was recorded amongst abnormal tissues (Malini et al., 2006). Here, we observed that Raman spectroscopy showed 100% sensitivity and specificity between normal and all HNC models and with a very high level of discrimination in identifying between subpopulations of dysplastic and cancer cells.

These results indicate that Raman spectroscopy might be sensitive enough to discriminate between cancer sub types and when coupled with multivariate data analysis techniques have the potential to become a powerful tool to classify different grades, stages or types of tumours.

**4. Conclusion**

Currently surgical excision and histopathological evaluation of a suspicious lesion is the gold standard for diagnosis of HNC, however it is associated with a number of limitations including it being invasive, creating time delays, causing anxiety to the patients and misdiagnosis due to high inter and intra-observer variability. Raman spectroscopy is an optical technique with the ability to extract molecular level information to help determine the functional groups present in a tissue and the molecular conformations of tissue constituents. Here we have shown that differences in peak intensities can be analysed and attributed to cytological changes caused by variations in lipid, protein and nucleic acid contents in cells as malignant transformation occurs. Results obtained in this study demonstrate that Raman spectroscopy not only has the potential to differentiate between normal, pre-malignant and cancerous tissue models but could also be sensitive enough to detect subtypes of dysplasia or cancer on the basis of their sub-cellular differences. In the future it would be interesting to co-relate these findings with *ex vivo* and *in vivo* results to improve clinical diagnostic outcomes.

**Acknowledgements:**

We thank Dr. Keith Hunter for providing the D19 and D20 dysplastic cell lines for this project.

**Conflict of interest:**

The authors declare no conflict of interest.

## REFERENCES


Argiris, A., Karamouzis, M. V., Raben, D. & Ferris, R. L. 2008. Head and neck cancer. *Lancet,* 371**,** 1695-709.

Bergholt, M. S., Zheng, W. & Huang, Z. 2012. Characterizing variability in in vivo Raman spectroscopic properties of different anatomical sites of normal tissue in the oral cavity. *Journal of Raman Spectroscopy,* 43**,** 255-262.

Bigio, I. J., Bown, S. G., Briggs, G., Kelley, C., Lakhani, S., Pickard, D., Ripley, P. M., Rose, I. G. & Saunders, C. 2000. Diagnosis of breast cancer using elastic-scattering spectroscopy: preliminary clinical results. *Journal of Biomedical Optics,* 5**,** 221-228.

Chang, S. E., Foster, S., Betts, D. & Marnock, W. E. 1992. DOK, a cell line established from human dysplastic oral mucosa, shows a partially transformed non-malignant phenotype. *International Journal of Cancer,* 52**,** 896-902.

Colley, H. E., Hearnden, V., Jones, A. V., Weinreb, P. H., Violette, S. M., MacNeil, S., Thornhill, M. H. & Murdoch, C. 2011. Development of tissue-engineered models of oral dysplasia and early invasive oral squamous cell carcinoma. *Br J Cancer,* 105**,** 1582-1592.

DeBerardinis, R. J., Sayed, N., Ditsworth, D. & Thompson, C. B. 2008. Brick by brick: metabolism and tumor cell growth. *Current Opinion in Genetics & Development,* 18**,** 54-61.

Devpura, S., Thakur, J. S., Sethi, S., Naik, V. M. & Naik, R. 2011. Diagnosis of head and neck squamous cell carcinoma using Raman spectroscopy: tongue tissues. *Journal of Raman Spectroscopy***,** n/a-n/a.

Devpura, S., Thakur, J. S., Sethi, S., Naik, V. M. & Naik, R. 2012. Diagnosis of head and neck squamous cell carcinoma using Raman spectroscopy: tongue tissues. *Journal of Raman Spectroscopy,* 43**,** 490-496.

Epstein, J. B., Zhang, L. & Rosin, M. J. 2002. *Can Dental Assoc,* 68**,** 610-621.

Heimann, T. M., Cohen, R. D., Szporn, A. & Gil, J. 1991. Correlation of nuclear morpometry and DNA ploidy in rectal-cancer. *Diseases of the Colon & Rectum,* 34**,** 449-454.

Huang, Z., McWilliams, A., Lui, H., McLean, D. I., Lam, S. & Zeng, H. 2003. Near-infrared Raman spectroscopy for optical diagnosis of lung cancer. *International Journal of Cancer,* 107**,** 1047-1052.

IU. Rehman, Z**.** Movasaghi & Rehman, S**.** 2012. *Vibrational Spectroscopy for Tissue Analysis,* Boca Raton, Florida, United States., Taylor and Francis Group.

Jiang, L., Ji, N., Zhou, Y., Li, J., Liu, X., Wang, Z., Chen, Q. & Zeng, X. 2009. CAL 27 is an oral adenosquamous carcinoma cell line. *Oral Oncology,* 45**,** e204-e207.

Johnson, N. W., Warnakulasuriya, S., Gupta, P. C., Dimba, E., Chindia, M., Otoh, E. C., Sankaranarayanan, R., Califano, J. & Kowalski, L. 2011. Global oral health inequalities in incidence and outcomes for oral cancer: causes and solutions. *Advances in dental research,* 23**,** 237-46.

Kamath, S. D. & Mahato, K. K. 2007. Optical pathology using oral tissue fluorescence spectra: classification by principal component analysis and k-means nearest neighbor analysis. *Journal of Biomedical Optics,* 12**,** 014028-014028-9.

Krishna, C. M., Sockalingum, G. D., Kurien, J., Rao, L., Venteo, L., Pluot, M., Manfait, M. & Kartha, V. B. 2004. Micro-Raman spectroscopy for optical pathology of oral squamous cell carcinoma. *Applied Spectroscopy,* 58**,** 1128-1135.

Malini, R., Venkatakrishna, K., Kurien, J., M. Pai, K., Rao, L., Kartha, V. B. & Krishna, C. M. 2006. Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: A Raman spectroscopy study. *Biopolymers,* 81**,** 179-193.

Manoharan, R., Wang, Y. & Feld, M. S. 1996. Histochemical analysis of biological tissues using Raman spectroscopy. *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy,* 52**,** 215-249.

McGregor, F., Muntoni, A., Fleming, J., Brown, J., Felix, D. H., MacDonald, D. G., Parkinson, E. K. & Harrison, P. R. 2002. Molecular Changes Associated with Oral Dysplasia Progression and Acquisition of Immortality: Potential for Its Reversal by 5-Azacytidine. *Cancer Research,* 62**,** 4757-4766.

McGregor, F., Wagner, E., Felix, D., Soutar, D., Parkinson, K. & Harrison, P. R. 1997. Inappropriate Retinoic Acid Receptor-β Expression in Oral Dysplasias: Correlation with Acquisition of the Immortal Phenotype. *Cancer Research,* 57**,** 3886-3889.

Mian, S. A., Colley, H. E., Thornhill, M. H. & Rehman, I. u. 2014. Development of a Dewaxing Protocol for Tissue-Engineered Models of the Oral Mucosa Used for Raman Spectroscopic Analysis. *Applied Spectroscopy Reviews,* 49**,** 614-617.

Movasaghi, Z., Rehman, S. & Rehman, I. U. 2007. Raman spectroscopy of biological tissues. *Applied Spectroscopy Reviews,* 42**,** 493-541.

Movasaghi, Z., Rehman, S. & Rehman, I. u. 2012. Raman Spectroscopy Can Detect and Monitor Cancer at Cellular Level: Analysis of Resistant and Sensitive Subtypes of Testicular Cancer Cell Lines. *Applied Spectroscopy Reviews,* 47**,** 571-581.

Napier, S. S. & Speight, P. M. 2008. Natural history of potentially malignant oral lesions and conditions: an overview of the literature. *Journal of Oral Pathology & Medicine,* 37**,** 1-10.

Rangan, S. R. S. 1972. A new human cell line (FaDu) from a hypopharyngeal carcinoma. *Cancer,* 29**,** 117-121.

Rehman, S., Movasaghi, Z., Tucker, A. T., Joel, S. P., Darr, J. A., Ruban, A. V. & Rehman, I. U. 2007. Raman spectroscopic analysis of breast cancer tissues: identifying differences between normal, invasive ductal carcinoma and ductal carcinoma in situ of the breast tissue. *Journal of Raman Spectroscopy,* 38**,** 1345-1351.

Reibel, J. 2003. Prognosis of Oral Pre-malignant Lesions: Significance of Clinical, Histopathological, and Molecular Biological Characteristics. *Critical Reviews in Oral Biology & Medicine,* 14**,** 47-62.

Short, M. A., Lam, S., McWilliams, A. M., Ionescu, D. N. & Zeng, H. 2011. Using Laser Raman Spectroscopy to Reduce False Positives of Autofluorescence Bronchoscopies A Pilot Study. *Journal of Thoracic Oncology,* 6**,** 1206-1214.

Singh, S. P., Deshmukh, A., Chaturvedi, P. & Krishna, C. M. 2012. In vivo Raman spectroscopic identification of premalignant lesions in oral buccal mucosa. *Journal of Biomedical Optics,* 17.

Speight, P. M. 2007. Update on oral epithelial dysplasia and progression to cancer. *Head and neck pathology,* 1**,** 61-6.

Speight, P. M., Abram, T. J., Floriano, P. N., James, R., Vick, J., Thornhill, M. H., Murdoch, C., Freeman, C., Hegarty, A. M., D'Apice, K., Kerr, A. R., Phelan, J., Corby, P., Khouly, I., Vigneswaran, N., Bouquot, J., Demian, N. M., Weinstock, Y. E., Redding, S. W., Rowan, S., Yeh, C.-K., McGuff, H. S., Miller, F. R. & McDevitt, J. T. 2015. Inter-Observer Agreement in Dysplasia Grading: Towards an Enhanced Gold Standard for Clinical Pathology Trials. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*.

Stefanuto, P., Doucet, J.-C. & Robertson, C. 2014. Delays in treatment of oral cancer: a review of the current literature. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology,* 117**,** 424-429.

Stone, N., Kendall, C., Shepherd, N., Crow, P. & Barr, H. 2002. Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers. *Journal of Raman Spectroscopy,* 33.

Stone, N., Kendall, C., Smith, J., Crow, P. & Barr, H. 2004. Raman spectroscopy for identification of epithelial cancers. *Faraday Discussions,* 126**,** 141-157.

Stone, N., Stavroulaki, P., Kendall, C., Birchall, M. & Barr, H. 2000. Raman spectroscopy for early detection of laryngeal malignancy: Preliminary results. *Laryngoscope,* 110**,** 1756-1763.

Su, L., Sun, Y. F., Chen, Y., Chen, P., Shen, A. G., Wang, X. H., Jia, J., Zhao, Y. F., Zhou, X. D. & Hu, J. M. 2012. Raman spectral properties of squamous cell carcinoma of oral tissues and cells. *Laser Physics,* 22**,** 311-316.

Tankiewicz, A., Dziemianczyk, D., Buczko, P., Szarmach, I. J., Grabowska, S. Z. & Pawlak, D. 2006. Tryptophan and its metabolites in patients with oral squamous cell carcinoma: preliminary study. *Advances in medical sciences,* 51 Suppl 1**,** 221-4.

Teh, S. K., Zheng, W., Ho, K. Y., Teh, M., Yeoh, K. G. & Huang, Z. 2010. Near-infrared Raman spectroscopy for early diagnosis and typing of adenocarcinoma in the stomach. *Br J Surg,* 97**,** 550-7.

Valdés, R., Stefanov, S., Chiussi, S., López-Alvarez, M. & González, P. 2014. Pilot research on the evaluation and detection of head and neck squamous cell carcinoma by Raman spectroscopy. *Journal of Raman Spectroscopy,* 45**,** 550-557.

Venkatakrishna, K., Kurien, J., Pai, K. M., Valiathan, M., Kumar, N. N., Krishna, C. M., Ullas, G. & Kartha, V. B. 2001. Optical pathology of oral tissue: A Raman spectroscopy diagnostic method. *Current Science,* 80**,** 665-669.

**Figure Legends:**

**Figure 1: Haematoxylin and eosin stained sections of tissue-engineered oral mucosa.** a-c) Normal oral mucosa, d-f) dysplastic oral mucosa, DOK, D19 and D20, respectively. g-i) Cancerous oral mucosa Cal27, SCC4 and FaDu, respectively. Yellow boxes indicate collection points for spectral data on parallel 20 µm sections. Scale bar = 100 µm.

**Figure 2: Average spectra of tissue engineered mucosa models.** a) Normal oral mucosa (NOM) and dysplastic oral mucosa models, (D20, D19 and DOK), b) NOM and cancerous oral mucosa models (SCC4, FaDu and Cal27) and c) cancerous and dysplastic oral mucosa models (SCC4, FaDu, Cal27, D20, D19 and DOK).

**Figure 3: Principal component analysis plots.** PCA and 3D plots for the amide I and amide III spectra range for a) normal versus dysplastic, b) normal versus cancer and c) normal versus dysplastic.

**Figure 4: Cluster analysis of the full spectral range.** a) Normal and dysplastic, b) normal and cancer, c) dysplastic and cancer and d) normal, dysplastic and cancer.

**Figure 5: Final linear discriminant analysis predictions for all unknown tissue engineered models in the broad tissue state.** a) Normal and cancer, b) dysplastic and cancer and c) normal, dysplastic and cancer.

**Figure 6: Final linear discriminant analysis predictions for all unknown tissue engineered models by cellular subtype.** a) Normal (NOM), and cancer tissue-engineered models (Cal27, SCC4 and FaDu) and b) dysplastic (D19, D20, DOK) and cancer tissue-engineered models (Cal27, SCC4 and FaDu).

Figure 1

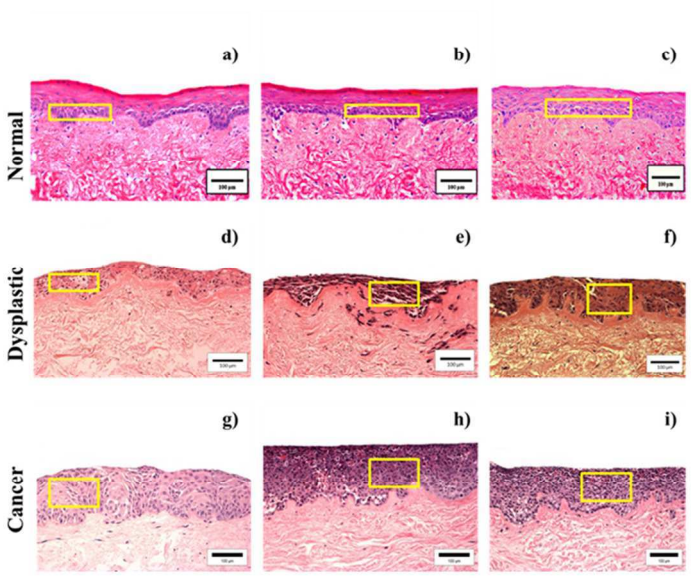


Figure 1: Haematoxylin and eosin stained sections of tissue-engineered oral mucosa. a-c) Normal oral mucosa, d-f) dysplastic oral mucosa, DOK, D19 and D20, respectively. g-i) Cancerous oral mucosa Cal27, SCC4 and FaDu, respectively. Yellow boxes indicate collection points for spectral data on parallel 20 µm sections. Scale bar = 100 µm.
254x190mm (96 x 96 DPI)

Figure 2

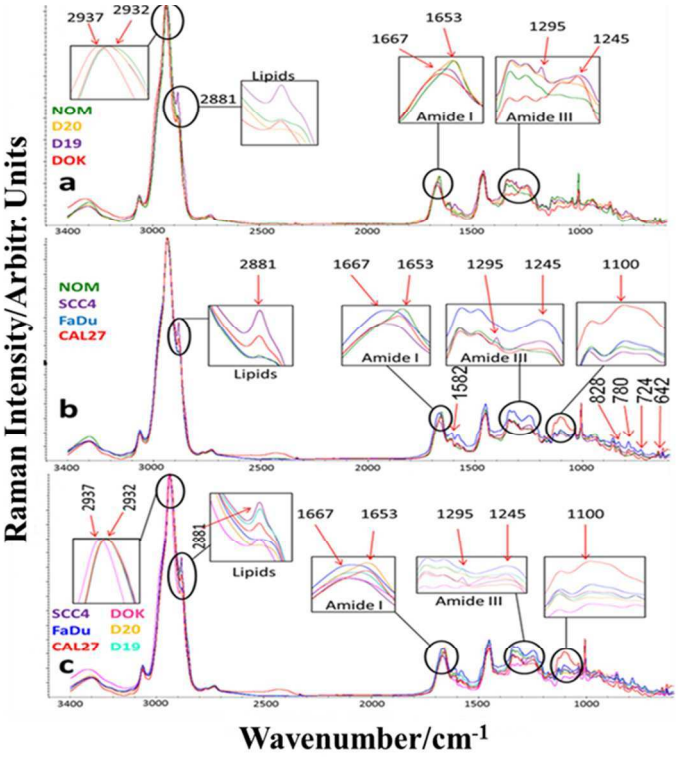


Figure 2: Average spectra of tissue engineered mucosa models. a) Normal oral mucosa (NOM) and dysplastic oral mucosa models, (D20, D19 and DOK), b) NOM and cancerous oral mucosa models (SCC4, FaDu and Cal27) and c) cancerous and dysplastic oral mucosa models (SCC4, FaDu, Cal27, D20, D19 and DOK).
254x190mm (96 x 96 DPI)

Fig: 3

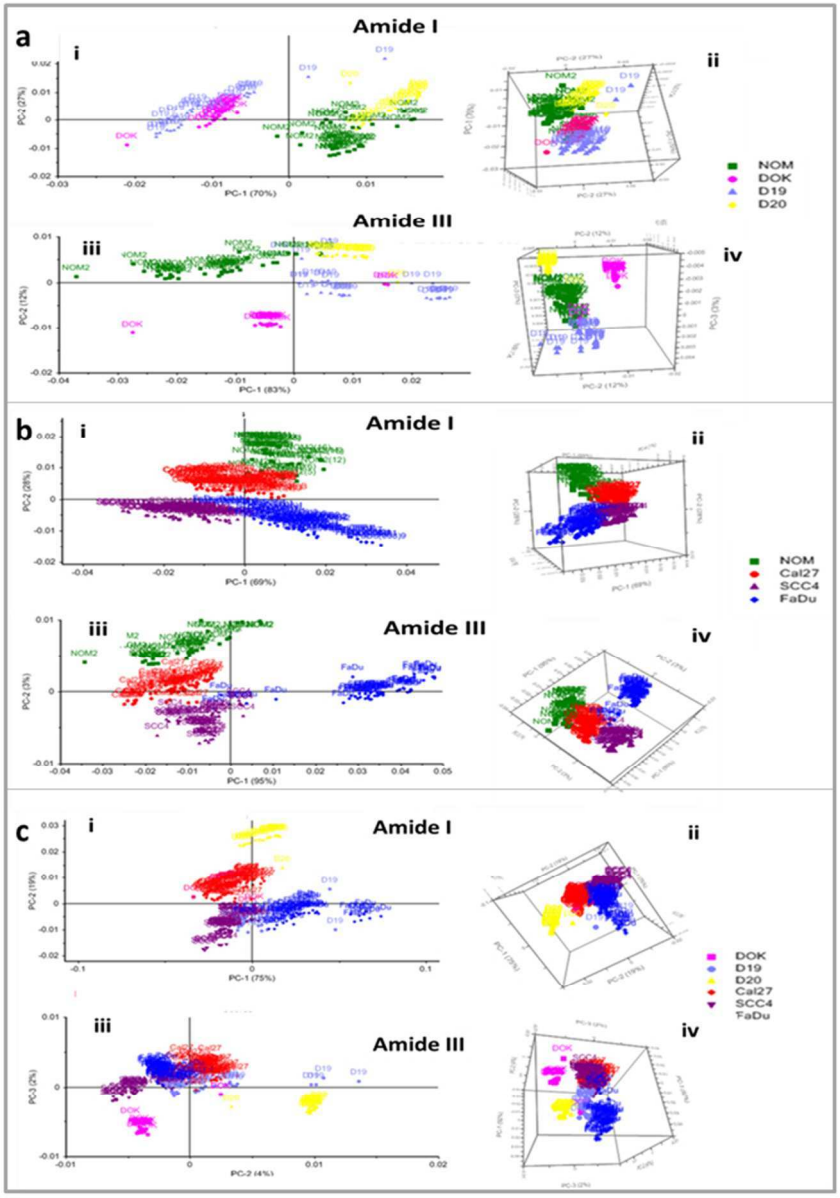


Figure 3: Principal component analysis plots. PCA and 3D plots for the amide I and amide III spectra range for a) normal versus dysplastic, b) normal versus cancer and c) normal versus dysplastic.
190x254mm (96 x 96 DPI)

Figure 4:

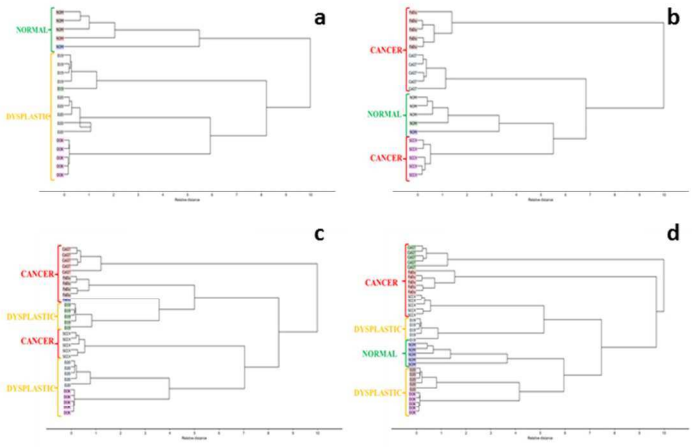


Figure 4: Cluster analysis of the full spectral range. a) Normal and dysplastic, b) normal and cancer, c) dysplastic and cancer and d) normal, dysplastic and cancer.
254x190mm (96 x 96 DPI)

Figure 5:

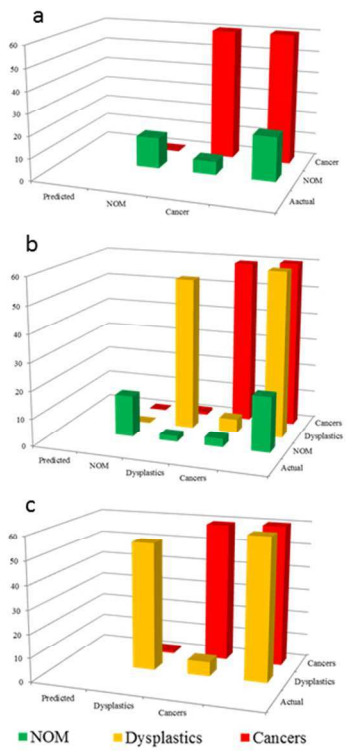


Figure 5: Final linear discriminant analysis predictions for all unknown tissue engineered models in the broad tissue state. a) Normal and cancer, b) dysplastic and cancer and c) normal, dysplastic and cancer.
254x190mm (96 x 96 DPI)

Figure 6:

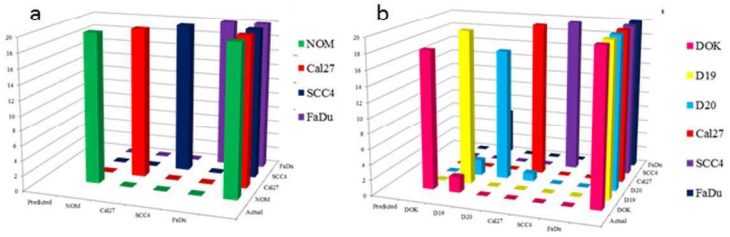


Figure 6: Final linear discriminant analysis predictions for all unknown tissue engineered models by cellular subtype. a) Normal (NOM), and cancer tissue-engineered models (Cal27, SCC4 and FaDu) and b) dysplastic (D19, D20, DOK) and cancer tissue-engineered models (Cal27, SCC4 and FaDu).
254x190mm (96 x 96 DPI)