Raman microspectroscopic study for the detection of oral field cancerisation using brush biopsy samples

Isha Behl
Technological University Dublin, isha.behl@tudublin.ie

Genecy Calado
Technological University Dublin

Anika Vishwakarma
Technological University Dublin

See next page for additional authors

Recommended Citation

Authors

Isha Behl, Genecy Calado, Anika Vishwakarma, Stephen Flint, Sheila Galvin, Claire Healy, Marina Leite Pimentel, Alison Malkin, Hugh Byrne, and Fiona Lyng

FULL ARTICLE

Raman microspectroscopic study for the detection of oral field cancerisation using brush biopsy samples

Isha Behl[1,2]* | Genecy Calado[1,2] | Anika Vishwakarma[1,2] | Stephen Flint[3] |
Sheila Galvin[3] | Claire M. Healy[3] | Marina Leite Pimentel[4] | Alison Malkin[5] |
Hugh J. Byrne[6] | Fiona M. Lyng[1,2]

[1]Centre for Radiation and Environmental Science, FOCAS Research Institute, Technological University Dublin, Dublin, Ireland

[2]School of Physics, Technological University Dublin, Dublin, Ireland

[3]Oral Medicine Unit, Dublin Dental University Hospital, Trinity College Dublin, Dublin, Ireland

[4]Division of Restorative Dentistry and Periodontology, Dublin Dental University Hospital, Trinity College Dublin, Dublin, Ireland

[5]School of Biological Sciences, Technological University Dublin, Dublin, Ireland

[6]FOCAS Research Institute, Technological University Dublin, Dublin, Ireland

*Correspondence
Dr Isha Behl, Centre for Radiation and Environmental Science, FOCAS Research Institute, Technological University Dublin, City Campus, Dublin, Ireland.
Email: behlisha86@gmail.com

Abstract

Field cancerisation (FC) is potentially an underlying cause of poor treatment outcomes of oral squamous cell carcinoma (OSCC). To explore the phenomenon using Raman microspectroscopy, brush biopsies from the buccal mucosa, tongue, gingiva and alveolus of healthy donors (n = 40) and from potentially malignant lesions (PML) of Dysplasia Clinic patients (n = 40) were examined. Contralateral normal samples (n = 38) were also collected from the patients. Raman spectra were acquired from the nucleus and cytoplasm of each cell, and subjected to partial least squares-discriminant analysis (PLS-DA). High discriminatory accuracy for donor and PML samples was achieved for both cytopalmic and nuclear data sets. Notably, contralateral normal (patient) samples were also accurately discriminated from donor samples and contralateral normal samples from patients with multiple lesions showed a similar spectral profile to PML samples, strongly indicating a FC effect. These findings support the potential of Raman microspectroscopy as a screening tool for PML using oral exfoliated cells.

KEYWORDS

contralateral normal, field cancerisation, oral brush biopsy cytological samples, oral potentially malignant lesions, partial least squares-discriminant analysis, Raman microspectroscopy, sensitivity and specificity
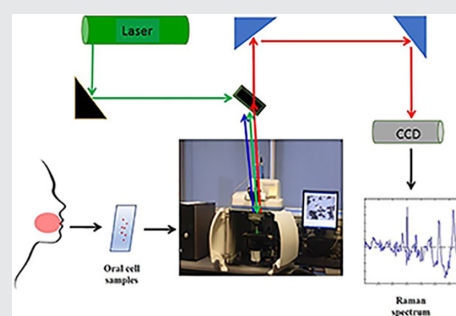
1 | INTRODUCTION

Oral squamous cell carcinoma (OSCC) ranks as the 18th most common cancer in the world and the eighth most frequent cancer in males [1]. OSCC can arise de novo, but is sometimes preceded by potentially malignant

lesions (PML), either leukoplakia (white patch) or erythroplakia (red patch) [2]. It is estimated that ~50% of oral cancers develop from PML [3]. The gold standard method for oral dysplasia or cancer diagnosis is biopsy followed by histopathological examination. The oral mucosa is divided into three basic types of epithelium, depending on the site and function; masticatory, lining and specialised epithelium [4, 5]. Gingiva and hard palate are the sites of the oral cavity which are subject to mechanical forces associated with mastication. These sites are covered by keratinising epithelium resembling that of the epidermis covering the skin. The mucosa is strongly attached to the underlying tissues by a collagenous connective tissue. Since the lips, buccal mucosa, vestibule, alveolar mucosa, soft palate, floor of mouth, and ventral surface of tongue are required to be flexible to assist functions like chewing, speech, or swallowing of a bolus, they are covered with a non-keratinising epithelium. The underlying connective tissue is more elastic and flexible when compared with the connective tissue of the masticatory mucosa. The dorsal surface of the tongue is a specialised epithelium, which can be represented as keratinised and non-keratinised epithelium [4, 5]. This epithelium is attached strongly to the muscle of the tongue.

According to WHO 2017 guidelines, histopathologically, PML are graded based on a three-tiered grading system namely mild, moderate, and severe dysplasia. Carcinoma in situ is considered equivalent to severe dysplasia in this grading system. The stratified epithelial layer of the oral mucosa is comprised of three layers of the epithelium; basal/parabasal, intermediate (spinosum and granulosum) and superficial layer. The superficial layer is keratinised in the case of keratinising epithelium and non-keratinised in the case of non-keratinising epithelium. The basal layer is the only layer of epithelium which has dividing cells, which mature and exfoliate on the surface of the epithelium (superficial layer) [6]. In the case of mild dysplasia, the changes are limited to the lower one-third of the epithelium (basal and parabasal layers), which exhibit cytological and/or architectural alterations described by the WHO 2017 guidelines. Moderate dysplasia shows disordered maturation up to the mid portion of the spinous layer (middle third) of the epithelium. Severe dysplasia/carcinoma in situ comprises of irregular maturation extending to a level above the midpoint of the epithelium (upper third) and may also include the entire thickness of the epithelium [7, 8].

Even though extensive research has been carried out in managing oral PML, the overall 5-year survival rate is ~50% for treated patients [3, 9]. The poor outcomes can be explained by the concept of FC, a term which was coined by Slaughter et al. [10] in 1953. The concept of FC

was demonstrated histologically by showing alterations in the tissue surrounding squamous cell carcinoma for oral cancer [10]. Since then, this concept has been used to explain the development of multiple PML, tumours and recurrent local lesions in various cancers including oropharynx, oesophagus, lungs, stomach, colon, cervix, anus, skin and bladder [11]. It has been observed, however, that oral cancer is most susceptible to this phenomenon, when compared to other cancers [11].

Different theories have been put forward to explain the concept of FC and the phenomenon of secondary primary, or second primary tumours, a different tumour unrelated to the primary tumour, arising in a different oral site [3, 7, 12]. The classical theory suggests that multiple squamous cell lesions occur, independent of each other, due to the fact that the entire oral cavity is exposed to carcinogens such as tobacco at the same time, resulting in the development of multiple independent genetic mutations. Another is a clonal theory, which suggests that a single cell is transformed or mutated that leads to a large extended pre-malignant field by clonal migration of dysplastic and altered cells. Oijen et al. [13] suggested two methods of migration of cells: (a) migration of malignant cells through the saliva (micro metastasis) and (b) intraepithelial migration of the progeny of initially transformed malignant cells.

The concept of FC not only explains the alteration in tissues surrounding squamous cell carcinoma, but also explains the development of multiple potentially malignant changes and the persistence of abnormal tissue even after surgery, which can lead to secondary primary tumours and recurrences [3, 11–13].

Raman spectroscopy is a label-free methodology and has shown good potential for oral cancer diagnosis [14–29]. Previous studies have demonstrated cancer field effects or malignancy-associated changes in skin as well as in cervical organotypic raft cultures, using Raman microspectroscopy [30, 31]. Singh et al. [32] acquired in vivo Raman spectra from 84 subjects (722 spectra) under five categories; healthy controls (no tobacco habit, no cancer), habitual healthy controls (tobacco habit, no cancer), habitual cancer and contralateral normal (opposite side of tumour) and non-habitual contralateral normal (no tobacco habit with cancer). The results suggested that the Raman spectral profile of the mucosa from healthy controls was distinct, but habitual healthy controls had a similar spectral profile to contralateral normal mucosa. They also observed that non-habitual contralateral normal mucosa was different to habitual healthy controls. The study demonstrated the efficacy of Raman microspectroscopy to detect malignancy associated and tobacco related spectral changes. Sahu et al. [15] have also demonstrated the concept of FC in pellets of oral exfoliated

cells using Raman microspectroscopy. They reported that cells from healthy volunteers and healthy volunteers with tobacco chewing habits could be discriminated from cells from patients, both contralateral normal and tumour. They also observed misclassifications between healthy volunteers with tobacco chewing habits and contralateral normal. Again, this demonstrated the concept of FC, and the sensitivity of Raman microspectroscopy to explore the phenomenon.

The development of new diagnostic markers for early detection of FC would be helpful in preventing cancer progression and the development of second primary tumours or PML, even before morphological changes can be seen [12]. Therefore, the present study has been carried out to detect FC based on single cells from oral cytological samples through Raman microspectroscopy.

2 | MATERIALS AND METHODS

2.1 | Sample collection and processing

Oral brush biopsy cytological samples were collected from the buccal mucosa and the ventral surface of the tongue (n = 32) and from the alveolus and gingiva (n = 8) in healthy volunteers. Samples were also collected from identified potential malignant lesions of buccal mucosa and tongue (n = 32) and gingival and alveolar mucosa (n = 8) in the patient group from the Oral Mucosa Dysplasia Clinic of the Dublin Dental University Hospital. In both cases, contralateral normal samples were collected, (n = 30) and (n = 8) respectively, from the opposite side to the clinically evident lesions. There was no evident lesion present on the contralateral side of the lesion and these sites were not tissue biopsied.

Ethical approval was obtained from the St James' Hospital/Adelaide and Meath Hospital Research Ethics Committee (REC ref: 2013/23/05) to collect patient samples, and from Dublin Institute of Technology (now TU Dublin) Research Ethics Committee (REC ref: 15/104) for the collection of healthy donor samples. Written informed consent was obtained from each donor and the study was conducted in accordance with ethical principles founded in the Declaration of Helsinki.

According to WHO 2017 guidelines [7], histopathologically, PML were graded based on a three-tiered grading system, namely mild, moderate, and severe dysplasia, by the expert pathologists at St. James' Hospital, Dublin. Five lesions were excised as part of the patient treatment. The details for the patients are given in Table 1.

The healthy donors were screened before sampling by a trained dentist who also supervised the sample collection to ensure no evident lesion was present inside the

oral cavity while collecting samples. The oral mucosa of healthy donors was not biopsied.

The standardisation of the protocol, including collection tool, sample preparation, Raman data acquisition and processing for Raman microspectroscopic analysis of healthy donors has been reported earlier [33]. The same protocol was applied to patients, to collect samples from the potentially malignant lesion. In order to collect the cells, at first the mouth was rinsed with an alcohol-free mouthwash. Then, samples were collected using an endocervical brush and placed into ThinPrep vials. Sample slides were prepared using the ThinPrep2000 processor with the aim of forming a monolayer of cells for analysis.

2.2 | Raman spectral acquisition

The protocol for spectral acquisition has been previously reported for ThinPrep samples [33, 34]. An XploRA confocal Raman instrument (HORIBA JobinYvon) was used for spectral acquisition. Manual calibration of the grating was done using the 520.7 cm^{-1} Raman band of crystalline silicon. Dark current measurement and recording of the substrate and optics signal was also performed, for data correction. As a source, a 532 nm diode laser was focused by a 100X objective (MPlanN, Olympus, NA = 0.9) onto the sample (~12 mW at the focus) and the resultant Raman signals were detected using a spectrograph with a 1200 g mm^{-1} grating coupled with a CCD. Raman spectra were acquired in the 400 to 1800 cm^{-1} region with an integration time of 30s per spectrum and averaged over two accumulations. Spectra were acquired from the centre of the nucleus and at random from the cytoplasmic regions of the cells. As ~10 to 25 cells were recorded per slide, depending on the quality of the sample, it is expected that any heterogeneity of the cytoplasm should be averaged out.

2.3 | Raman spectral processing and analysis

All the data processing and analyses were carried out using Matlab (Mathworks), PLS-Toolbox (Eigenvector Research Inc.) using in-house algorithms.

The glass spectral interference with the biological Raman signals was removed using the extended multiplicative signal correction (EMSC) described by Kerr et al. [35]. The EMSC algorithm also includes a polynomial term, and order n = 5, was used to correct the baseline of the Raman spectral data set. After glass correction, the data sets for nucleus and cytoplasm were subjected to

TABLE 1 Detailed information about the patients that were recruited to the study

Patient identifier	Histology (WHO 2017)	Sample collection site	Other sites with lesions	Smoking history	Alcohol history
CSP003	Moderate dysplasia	Buccal mucosa	Tongue	Ex smoker	Minimal (units not documented)
CSP006	Moderate dysplasia	Tongue		Never smoker	10 units/week
CSP009	Severe dysplasia	Tongue	Tongue	Never smoker	2 units/week
CSP010	Carcinoma in situ	Tongue		Never smoker	10 units/week
CSP012	Mild dysplasia	Tongue	Buccal mucosa	Ex smoker	12 units/week
CSP013	Moderate dysplasia	Tongue		Never smoker	No details
CSP014	Moderate dysplasia	Buccal mucosa	Tongue	Ex smoker	No details
CSP015	Mild dysplasia	Buccal mucosa	Inner lip	Smoker	15 units/week
CSP016	Mild dysplasia	Tongue	Labial	Ex smoker	14 units/week
CSP017	No dysplasia	Tongue		Smoker	Ex alcohol consumer
CSP020	Mild dysplasia	Tongue	Floor of mouth	Ex smoker	No alcohol
CSP021	No dysplasia	Buccal mucosa	Tongue	Ex smoker	No alcohol
CSP022	No dysplasia	Buccal mucosa		Never smoker	1 units/week
CSP023	Moderate dysplasia	Buccal mucosa		Ex smoker	No alcohol
CSP025	Severe dysplasia	Tongue	Tongue	Smoker	10 units/week
CSP027	No dsyplasia	Buccal mucosa		Smoker	8 units/week
CSP028	Mild Dysplasia	Tongue		Ex smoker	8 units/week
CSP029	Mild Dysplasia	Buccal mucosa		Ex smoker	No details
CSP030	Moderate dysplasia	Tongue		Non-smoker	Ex alcohol consumer
CSP032	Mild Dysplasia	Tongue		Non-smoker	1 unit/week
CSP035	Severe dysplasia	Buccal mucosa	Palate, Tongue	Ex smoker	No alcohol
CSP037	Moderate dysplasia	Tongue		Ex smoker	8 units/week
CSP038	Carcinoma in situ	Buccal mucosa		Non-smoker	Ex alcohol consumer
CSP040	Mild dysplasia	Tongue	Soft palate	Non-smoker	No alcohol
CSP041	Severe dysplasia	Tongue		Smoker	No alcohol
CSP042	Moderate dysplasia	Buccal mucosa	Soft palate, Tongue	Smoker	100 units/week
CSP043	Mild dysplasia	Buccal mucosa	Gingiva	Non-smoker	Ex alcohol consumer
CSP044	Severe Dysplasia	Tongue	Buccal Mucosa	Smoker	80 units/week
CSP046	Carcinoma in situ	Buccal mucosa	Alveolus	Ex smoker	1 units/week
CSP047	Moderate dysplasia	Tongue		Smoker	6 units/week
CSP048	Moderate dysplasia	Tongue	Alveolus	Smoker	10 units/week
CSP050	Moderate dysplasia	Tongue		Ex smoker	20 units/week reduced from 70 units/week at time of biopsy
CSP001	Moderate dysplasia	Alveolus		Smoker	7 units/week
CSP005	Moderate dysplasia	Gingiva	Buccal mucosa	Ex smoker	8 units/week
CSP007	Carcinoma in situ	Gingiva		Never smoker	2 units/week
CSP011	Mild dysplasia	Alveolus	Buccal mucosa	No details	No details
CSP018	Moderate dysplasia	Gingiva	Hard palate, alveolus	Ex smoker	6-7 units/week
CSP019	Moderate dysplasia	Alveolus		Ex smoker	2 units/week
CSP024	Moderate dysplasia	Alveolus	Buccal mucosa	Ex smoker	5-10 units/week
CSP026	Severe dysplasia	Gingiva	Buccal mucosa	Never smoker	1 units/week

TABLE 1 (Continued)

Patient identifier	Histology (WHO 2017)	Sample collection site	Other sites with lesions	Smoking history	Alcohol history
CSP031	Moderate dysplasia	Gingiva		Non-smoker	No alcohol
CSP033	Mild dysplasia	Alveolus	Soft palate	Smoker	No alcohol
CSP034	Mild dysplasia	Alveolus		Smoker	60 units/week

Note: Ex- former user, Never- never used, Non-occasional user.

FIGURE 1 Representative Pap stained samples from, A, healthy volunteer, B, patient contralateral normal sample (black arrows indicate abnormal cells and yellow arrow indicates a micronucleus) and, C, patient potentially malignant lesion sample (red arrow indicates abnormal cell). Scale-bar = 20 μm
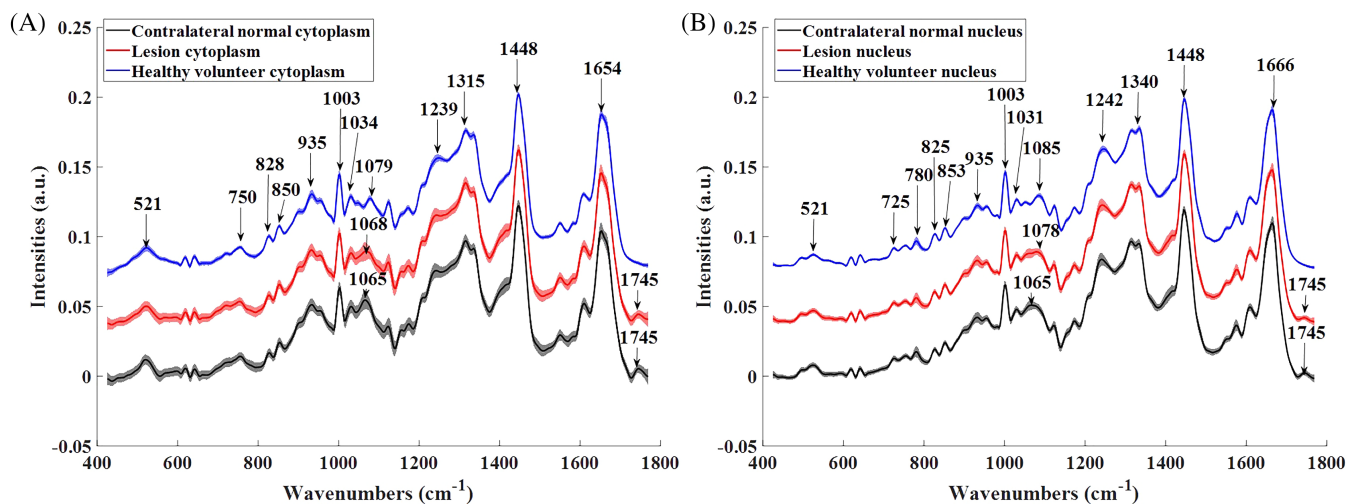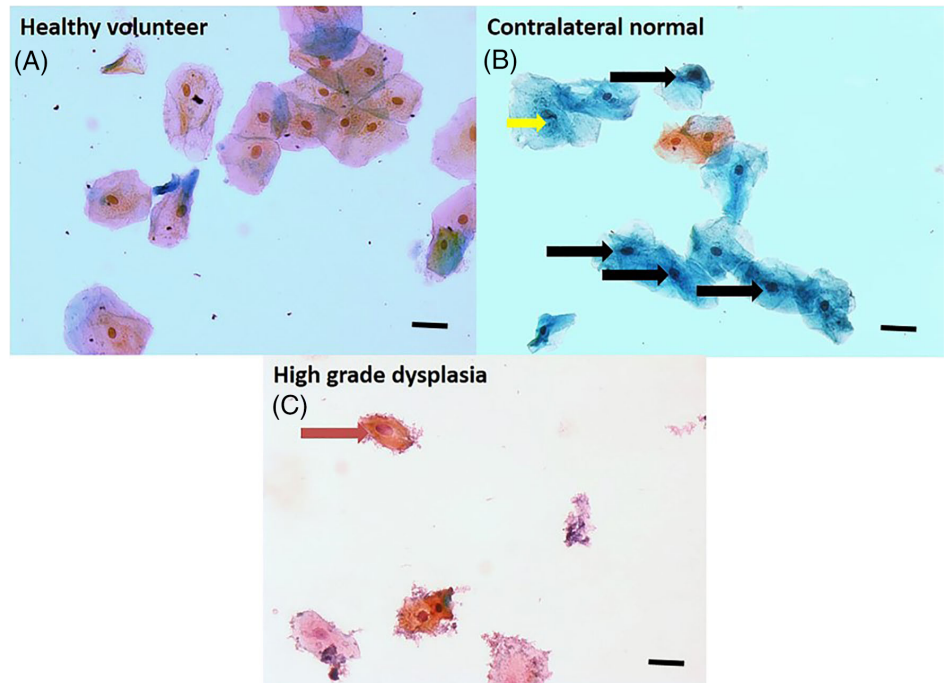

FIGURE 2 Mean spectra for healthy volunteers, potentially malignant lesion (lesion) and contralateral normal from buccal mucosa and tongue, A, Cytoplasm and, B, Nucleus. Shading denotes SD

smoothing (Savitzky-Golay, order = 5; window =13), outlier removal using Grubb's test [36] and vector normalisation.

Subsequently, pre-processed spectra of nucleus and cytoplasm were subjected to partial least squares-discriminant analysis (PLS-DA). It was used to build models in

order to train the data set. PLS-DA is a supervised form of multivariate analysis which works as a linear classifier that aims to maximise the variance between groups and minimise the variance within groups. The loadings of the

TABLE 2 Major Raman band positions [40, 41] and assignments in oral cytology for healthy donors and patients for buccal mucosa and tongue samples

Biomolecule	Raman bands (cm^{-1})
Proteins	571, 625, 645, 650, 656, 752, 760, 828, 831 850, 856, 918, 921, 963, 985, 994, 1000, 1003,1006, 1028, 1034,1040, 1105, 1163, 1177, 1211, 1234, 1239,1242, 1273, 1279, 1324, 1355, 1366, 1406, 1411,1448, 1476, 1546, 1566, 1583, 1600, 1673
Lipids	515, 521, 591, 602, 608, 705, 715, 740, 7 71, 1065, 1079,1085, 1130, 1155, 1300, 1307, 1392, 1417, 1437, 1465, 1639, 1654,1745, 1750
Nucleic acids	725, 783, 915, 1177, 1180,1195, 1248, 1251, 1290, 1296, 1315, 1335, 1340, 1369, 1372, 1485, 1493, 1572, 1662, 1666

discriminate, that is, latent variables (LV) can be plotted to give more information on the source of the variance [37].

Cross validation of PLS-DA classification is carried out in order to avoid over or under-fitting the model due to inappropriate selection of the components used, and secondly to determine the prediction error of the model. In leave one patient out cross validation, (LOPOCV), observations of one patient are excluded from the training set, one at a time, and the resulting model is evaluated by using the left out observations as tests. The procedure is repeated until all patients have been left out once and the average performance across all interactions is considered as the performance of the classification model [38].

3 | RESULTS AND DISCUSSION

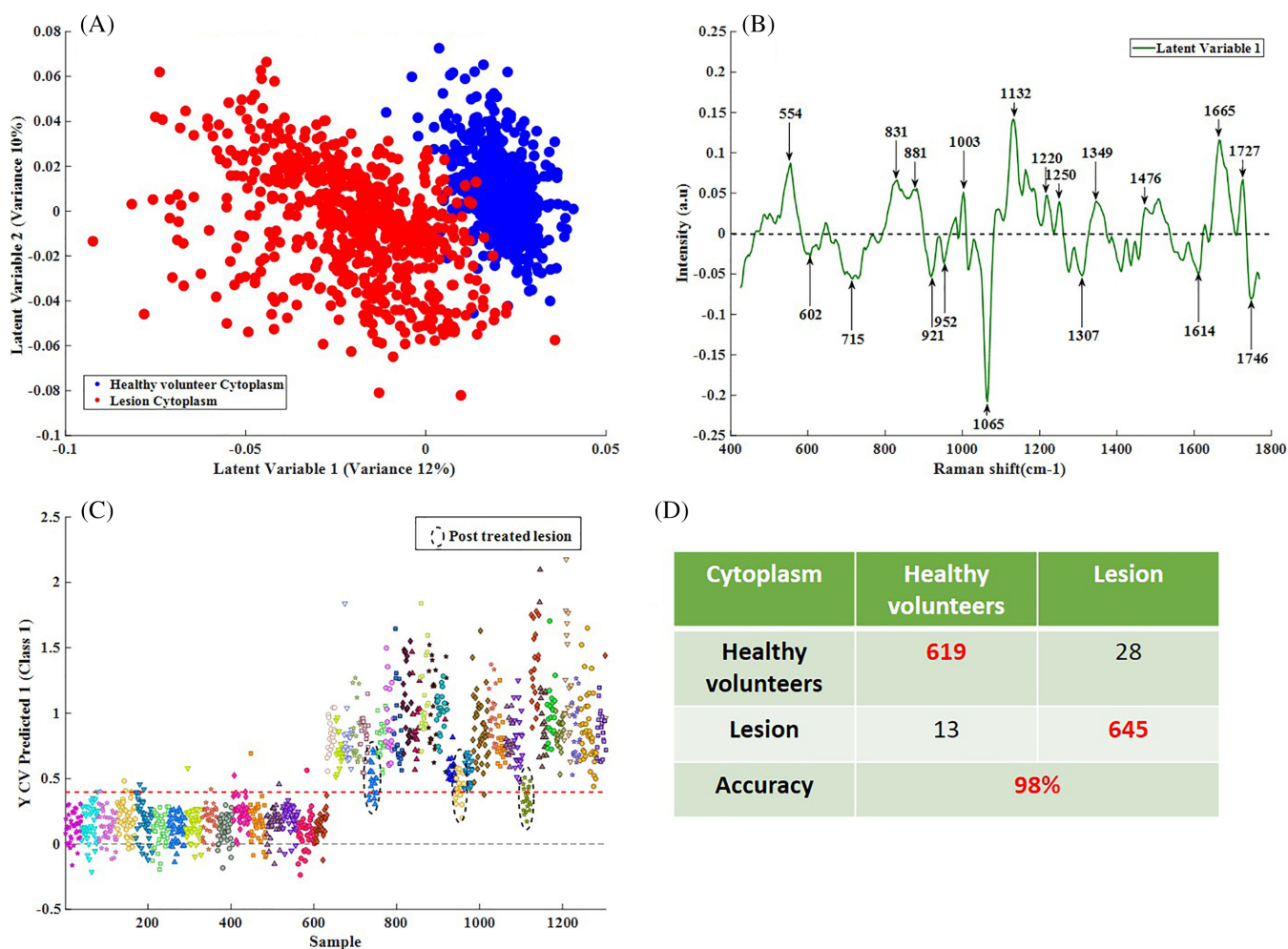It has been demonstrated previously that oral site is a confounding factor when developing classification

FIGURE 3 A, Score plot for healthy volunteer samples, and patient (potentially malignant lesion (lesion)) samples from cytoplasm of buccal mucosa and tongue, B, Latent variable loading, C, Probability predictive plot, D, Confusion matrix

models for PML and healthy donors [39]. Therefore, it has been concluded that classification models should be developed separately for buccal mucosa and tongue samples and for alveolus and gingiva samples.

At first for the present study, brush biopsy samples were collected from buccal mucosa and tongue in healthy donors (n = 32) and from patients (n = 32) with identified leukoplakia, erythroplakia or erythroleukoplakia, showing mild, moderate or severe dysplasia on histological analysis. In the case of contralateral normal samples (n = 30), these samples were collected from the opposite side to the lesion, where no visible lesion could be seen. Selected oral cytological samples were Papanicolaou (Pap) stained and assessed by a cytologist and oral pathologist, and can be seen in Figure 1. A healthy volunteer sample can be seen in Figure 1A. A contralateral normal sample, with dysplastic features, can be in Figure 1B (black arrows indicate abnormal cells and the yellow arrow indicates a micronucleus), whereas a lesion

sample with high grade dysplasia is shown in Figure 1C (red arrow indicates abnormal cell).

The mean spectra of cytoplasmic and nuclear regions from combined buccal mucosa and tongue samples of the cells show slight variations in the 1000 to 1200 cm^{-1} and 1700 to 1800 cm^{-1} regions of the spectrum, which are related to the lipidic content of the samples (Figure 2A, B). It has been demonstrated, in our previous study, that spectra from buccal mucosa and tongue share similar spectral profiles as they have a similar epithelium [33]. Table 2 lists the major Raman band assignments used in the present study for the healthy volunteer samples and patient (potentially malignant lesion and contralateral normal) samples.

At first, to compare the healthy volunteer samples and the potentially malignant lesion samples, a classification model was developed using PLS-DA with LOPOCV for the cytoplasmic and nuclear data sets (Figure 3). In this model, healthy volunteer samples could be
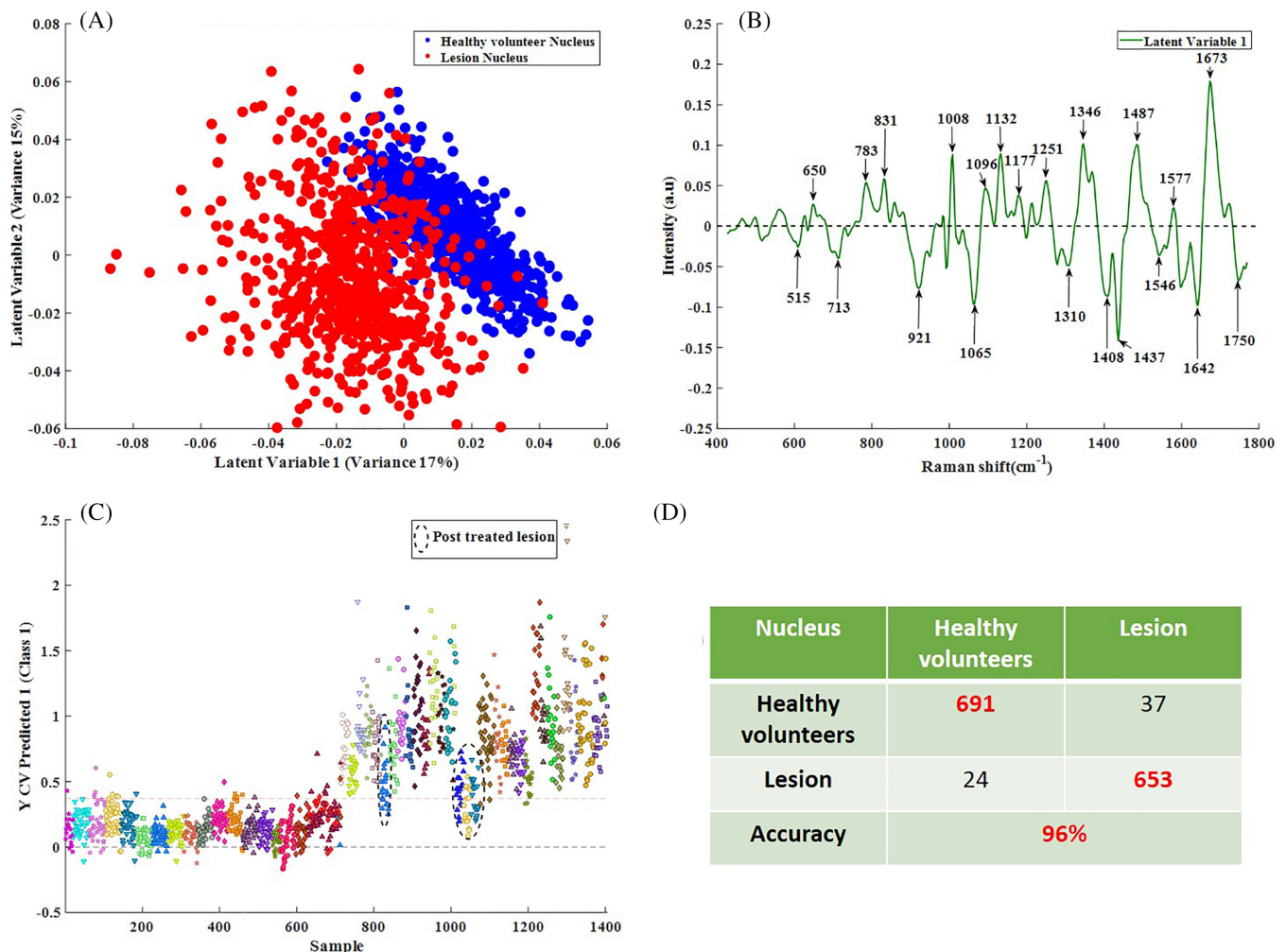

FIGURE 4   A, Score plot for healthy volunteer samples, and patient (potentially malignant lesion (lesion)) samples from nucleus of buccal mucosa and tongue, B, Latent variable loading, C, Probability predictive plot, D, Confusion matrix

discriminated from the potentially malignant lesion samples according to LV1 (Figures 3A and 4A). Positive loadings of LV1, corresponding to the healthy volunteer samples, are rich in proteins, whereas negative LV1 loadings, corresponding to the potentially malignant lesion samples, are rich in lipids for both the cytoplasmic and nuclear data sets (Figures 3B and 4B). Increased lipidic content of the dysplastic cells can be explained by the fact, that during cell division, cells need energy, and this requirement is fulfilled by upregulating lipid metabolism within the cell. Therefore, in dysplasic or cancer cells, the lipidic content increases [42, 43]. A more detailed analysis [44] indicated that the observed differences were largely due to a difference in palmitic acid/ceramide content. Ceramides are sphingolipids containing a sphingosine, or a related base, to which a fatty acid is linked through an amide bond. It is known that if ceramide synthesis is mediated by a de novo pathway, it condenses serine and palmitoyl-CoA (oxidised form of palmitic acid) as precursors in the endoplasmic reticulum. As the

ceramide and palmitic acid spectra are quite similar, many of their bands cancel each other, and the change in relative heights of the bands at 1065 cm^{-1} (ceramide) and 1130 cm^{-1} (palmitic acid), and the emergence of the band at 1740 cm^{-1} were observed in patient samples. The probability of the sample being classified as a healthy volunteer sample or a patient premalignant lesion sample is shown in Figures 3C and 4C. It was observed that a few potentially malignant lesion samples overlap with the healthy volunteer samples and these were identified as treated samples (ie, the lesion had been surgically excised) (Figures 3C and 4C). The confusion matrix of the model indicates 98% and 96% accuracy for the cytoplasmic and nuclear data set, respectively (Figures 3D and 4D).

Notably, healthy volunteer samples could also be discriminated from patient contralateral normal samples using PLS-DA for the cytoplasmic and nuclear data sets (Figures 5A and 6A). The positive loadings of LV1, corresponding to the healthy volunteer samples, are rich
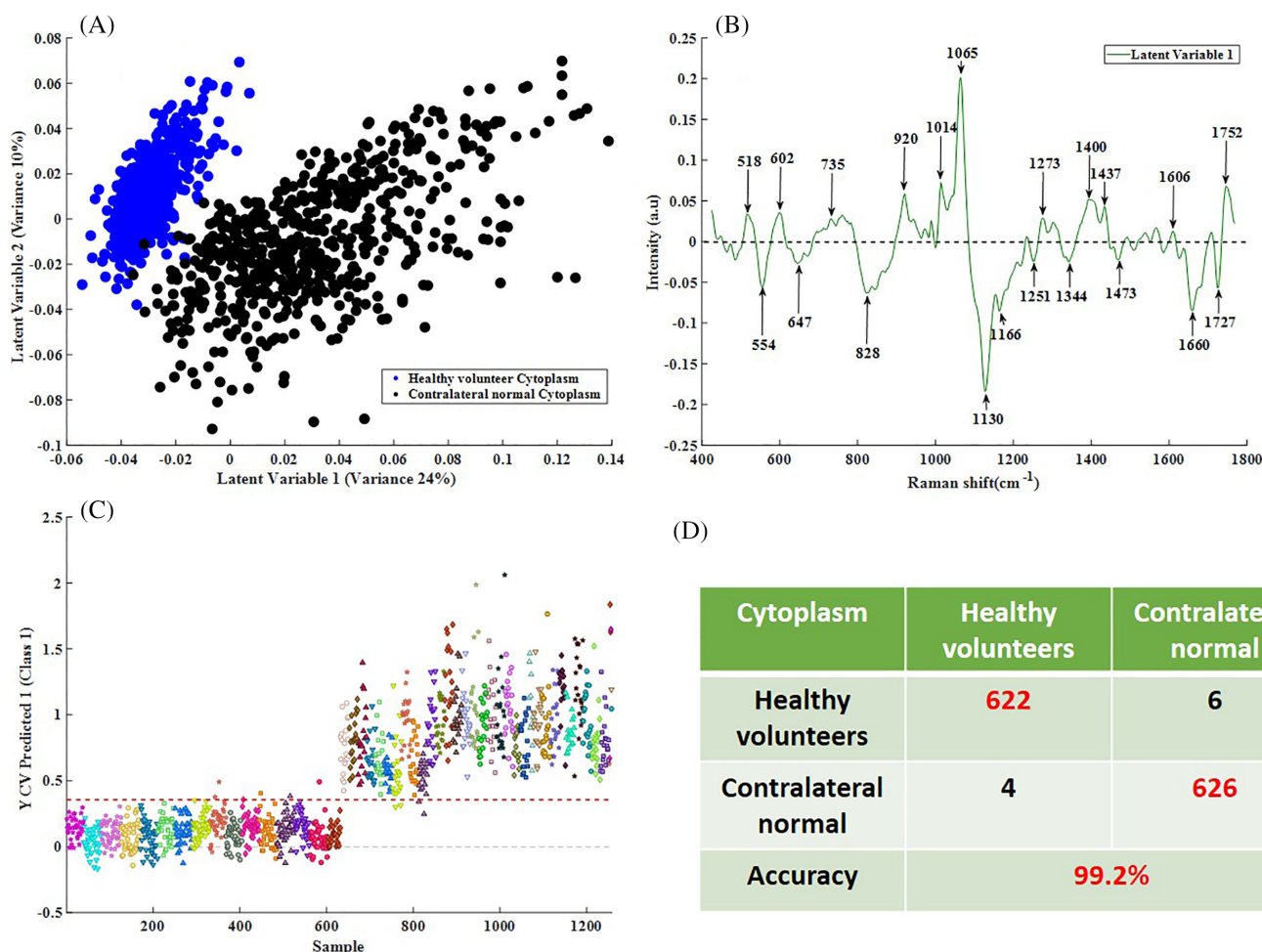

FIGURE 5 A, Score plot for healthy volunteer samples and patient (contralateral normal) samples from cytoplasm of buccal mucosa and tongue, B, Latent variable loading, C, Probability predictive plot, D, Confusion matrix
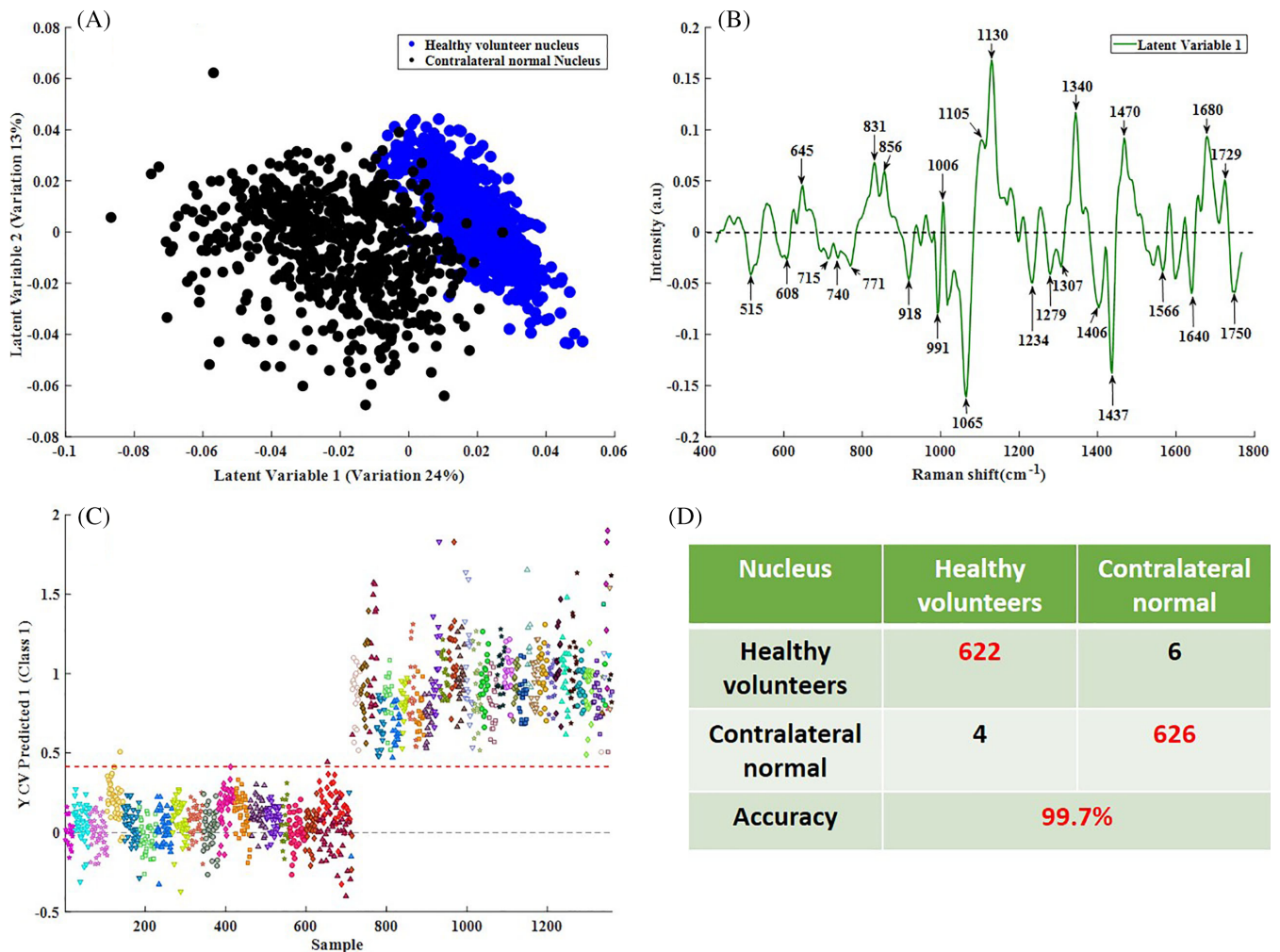
FIGURE 6 A, Score plot for healthy volunteer samples and patient (contralateral normal) samples from nucleus of buccal mucosa and tongue, B, Latent variable loading, C, Probability predictive plot, D, Confusion matrix

in proteins, whereas, negative loadings, corresponding to the contralateral normal samples, are rich in lipids, for both cytoplasm and nucleus (Figures 5B and 6B). The probability of the sample being classified as a healthy volunteer sample or a patient contralateral normal sample is shown in Figures 5C and 6C. The confusion matrix of the model indicates 99.2% and 99.7% accuracy for the cytoplasmic and nuclear data set, respectively (Figures 5D and 6D). No misclassification was observed among contralateral normal samples collected from the patients and healthy volunteer samples.

This observation suggests that, even though the contralateral normal samples were collected from the same patients, with or without treatment (surgical excision of the lesion), they have a distinctively different spectral profile that could be attributed to the FC phenomenon. In patients with PML, normal appearing sites with no apparent lesions (contralateral normal samples) could be discriminated from healthy volunteer samples. The

increase in lipidic content of the contralateral normal cells when compared to healthy volunteers indicates that their spectral profile is similar to cells from PML. Hence, a PLS-DA model was developed for potentially malignant lesion samples and contralateral normal samples from patients.

The potentially malignant lesion samples and contralateral normal samples from patients could be discriminated using PLS-DA for the cytoplasmic and nuclear data sets (Figures 7A and 8A). For both cytoplasm and nucleus, negative loadings of LV1, corresponding to the potentially malignant samples, are rich in proteins, whereas, positive loadings, corresponding to contralateral normal samples, are rich in lipids and proteins (Figures 7B and 8B). The probability of the sample being classified as a potentially malignant lesion sample or a contralateral normal sample is shown in Figure 7C, for the cytoplasmic data set, and in Figure 8C for the nuclear data set. The confusion matrix of the model indicates 77%
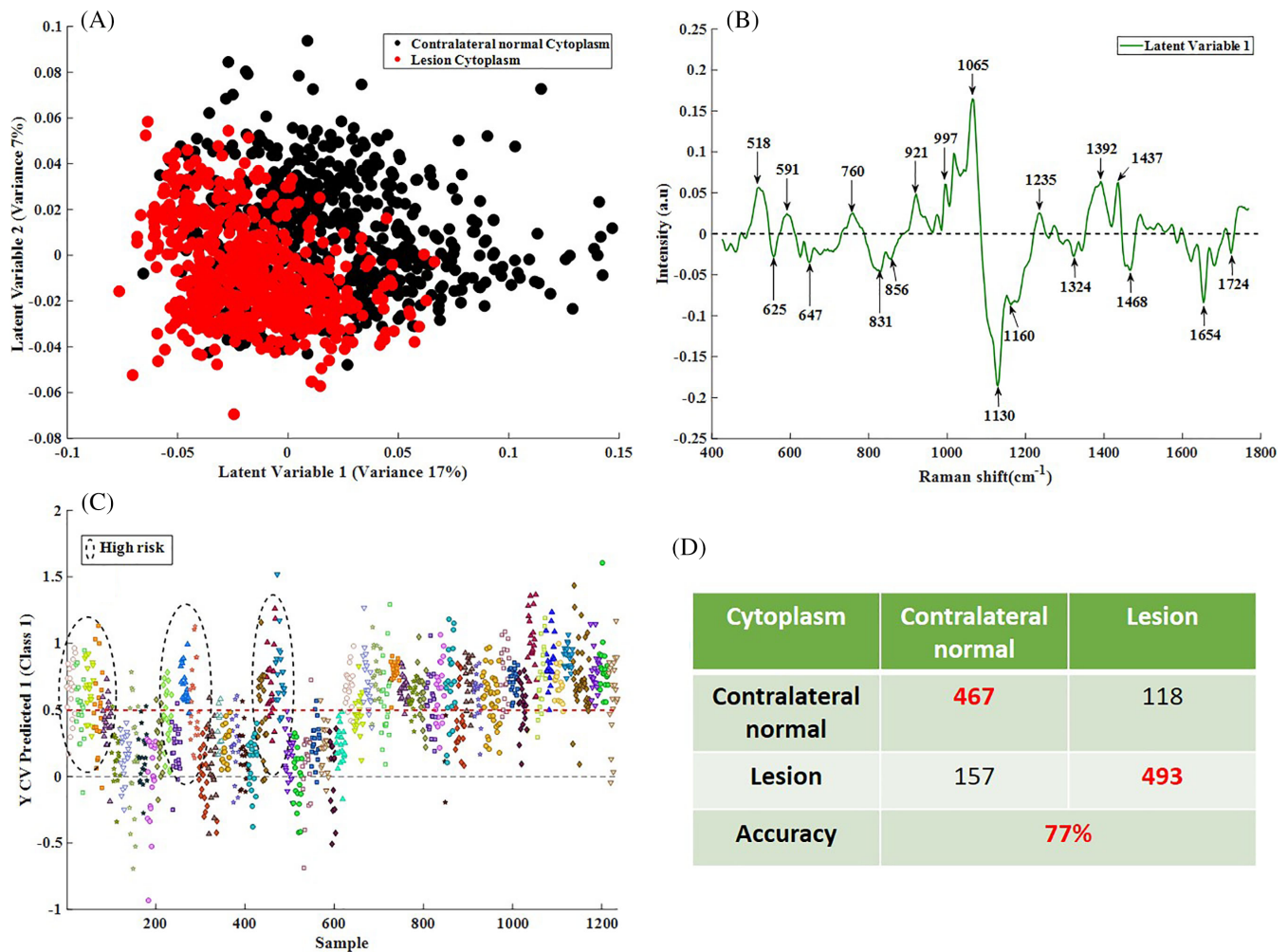
FIGURE 7 A, Score plot for potentially malignant lesion (lesion) samples and contralateral normal samples from cytoplasm, B, Latent variable loading, C, Probability predictive plot, D, Confusion matrix

and 84% accuracy for the cytoplasmic and nuclear data set respectively (Figures 7D and 8D). The increase in lipidic content in contralateral normal cells indicates that due to FC, these cells have undergone molecular changes similar to potentially malignant lesion cells. This explains the significant overlap among contralateral normal samples and potentially malignant samples (Figures 7B and 8B).

It was observed that some patients had developed multiple lesions over the entire oral cavity (ie, high risk patients), including the gingiva and alveolus, and in these patients, the contralateral normal samples exhibited a similar spectral profile to that of the potentially malignant lesion samples in the case of buccal mucosa and tongue sample. It is clinically important to verify the phenomenon of FC for the oral cavity, regardless of the origin of the sampling site in order to improve the treatment outcomes, even before morphological changes become apparent in different oral sites of the oral cavity [12], and

so the phenomenon was investigated for samples from gingiva and alveolus.

Brush biopsy samples were collected from gingiva and alveolus sites in healthy donors (n = 8) and from patients (n = 8) with identified leukoplakia, erythroplakia or erythroleukoplakia, showing mild, moderate or severe dysplasia on histological analysis. In the case of contralateral normal samples (n = 8), these samples were collected from the opposite side to the lesion, where no visible lesion could be seen. It has been demonstrated in our previous study [39][39] that samples of gingiva and alveolus share a similar spectral profile, as they have similar epithelium. Table 2 illustrates the major Raman band assignments used in the present study for the healthy volunteer samples and patient (potentially malignant lesion and contralateral normal) samples.

Similar results were obtained for the gingival and alveolar data set as for the buccal mucosa and tongue
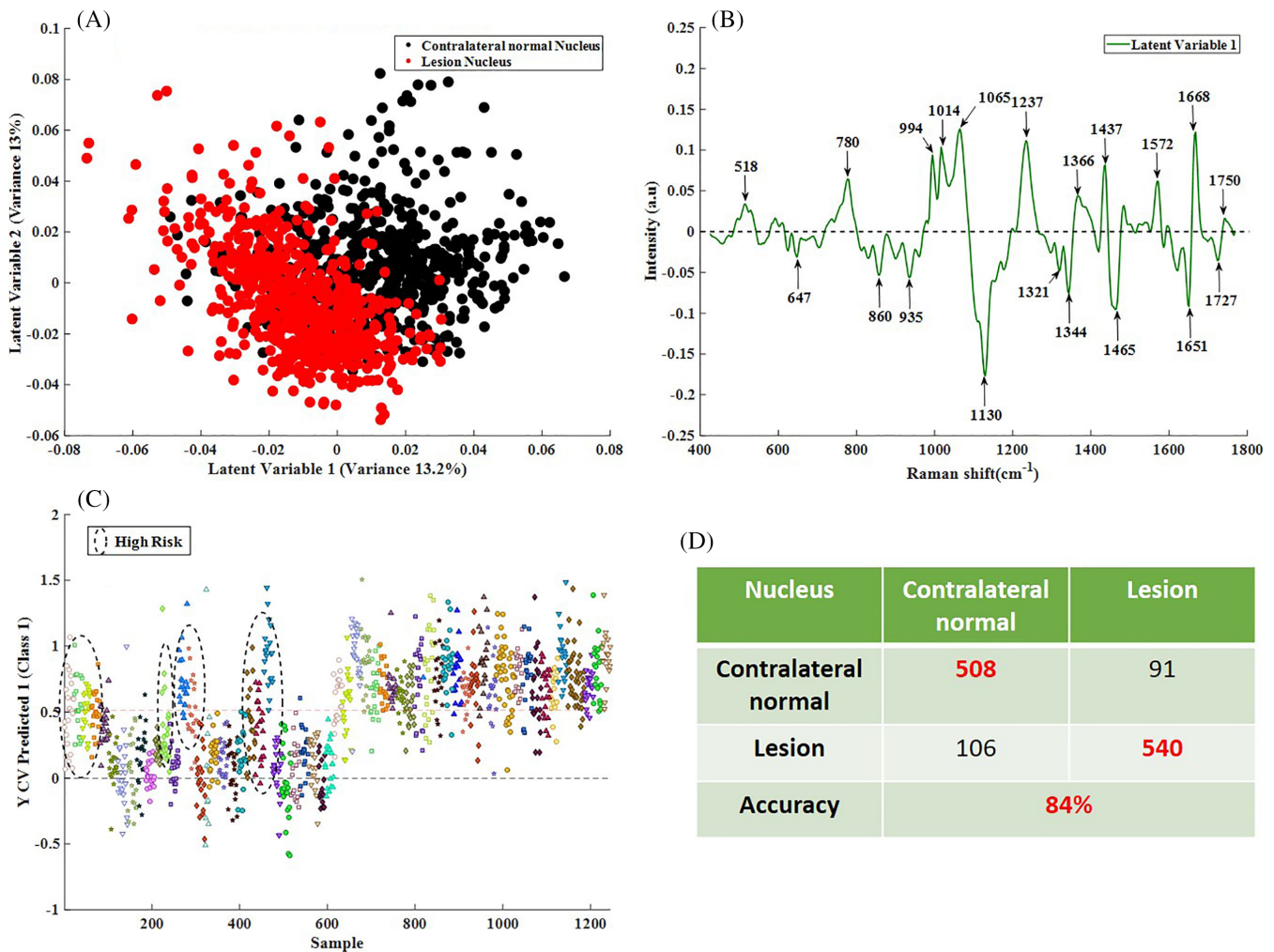
FIGURE 8 A, Score plot for potentially malignant lesion (lesion) samples and contralateral normal samples from nucleus, B, Latent variable loading, C, Probability predictive plot, D, Confusion matrix

TABLE 3 Accuracy for i) Healthy donor vs potentially lesion samples, ii) healthy donor vs patient contralateral normal samples and iii) potentially lesion vs patient contralateral normal samples, based on Raman spectra of cytoplasm (I) and nucleus (II)

Classes	Cytoplasm (I) Accuracy (%)	Nucleus (II) Accuracy (%)
i. Healthy donor vs potentially malignant lesion samples	98	81
ii. Healthy donor vs patient contralateral normal samples	98	80
iii. Potentially malignant lesion vs patient contralateral normal samples	81	87

data set. The PLS-DA models for cytoplasm and nucleus can be seen in Table 3. For healthy donor and potentially malignant lesion samples, accuracies of 98% and 81% were achieved for cytoplasm and nucleus respectively, whereas, for healthy donor and patient contralateral normal samples, accuracies of 98% and 80% were achieved for cytoplasm and nucleus, respectively, as shown in Figures S1–S4. It was observed that, unlike the buccal mucosa/tongue model, the lipidic bands could be seen in healthy donors in the case of gingiva/alveolar model. It is known that gingiva of healthy donors is rich in lipids such as ceramides, saturated fatty acids and cholesterol, which act as an antibacterial barrier against infection [5, 45]. Since the alveolus is an integral part of gingiva, they have a similar lipidic spectral profile. Thus, this explains the high lipidic content for healthy volunteers in the LVs for the cytoplasm and nucleus. Nevertheless, it is also well-known that protein metabolism changes during
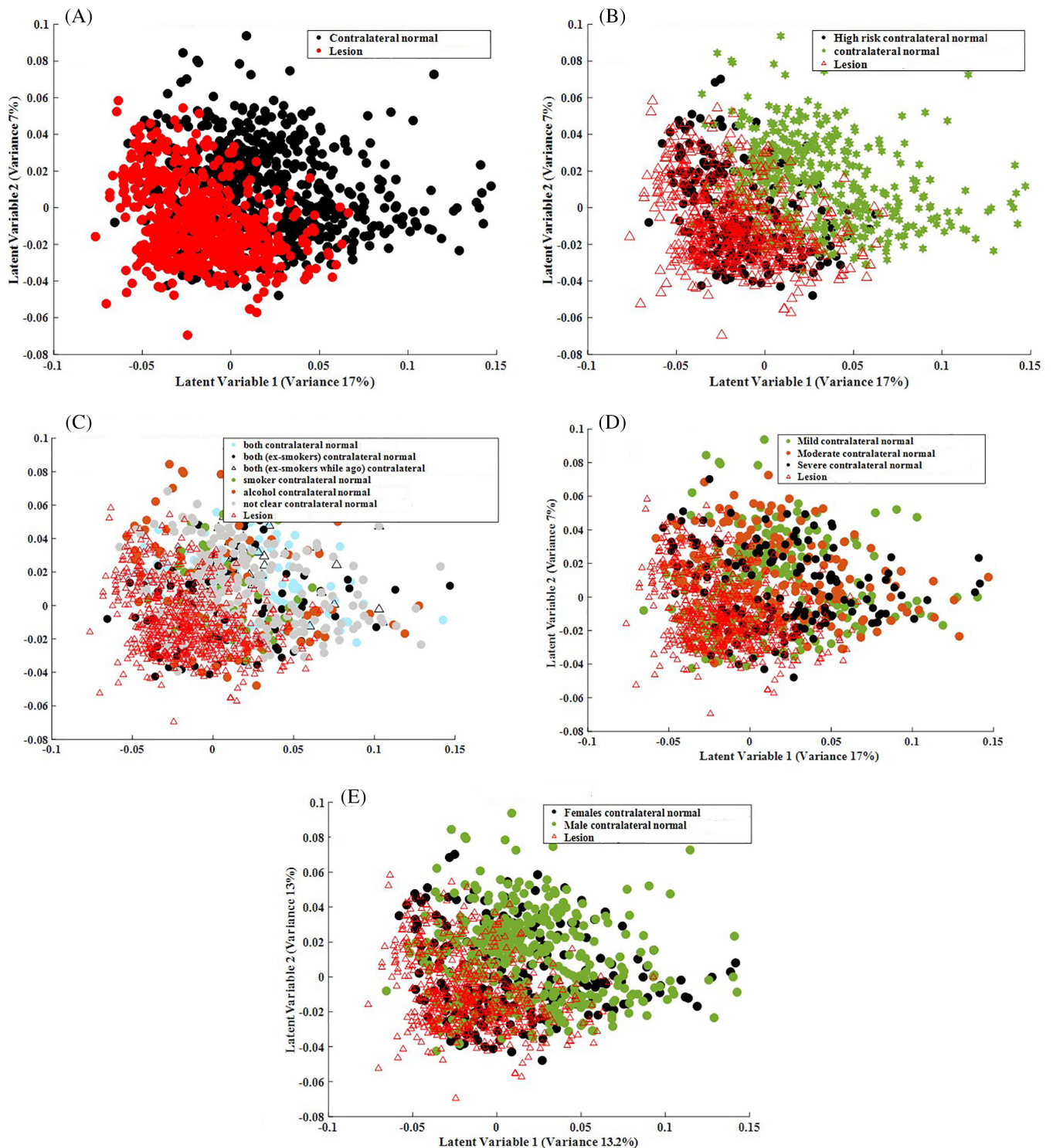
FIGURE 9 Score plot for, A, potentially malignant lesion (lesion) samples and contralateral normal samples for the cytoplasmic data set, B, for potentially malignant lesion (lesion) samples and contralateral normal samples highlighting high risk samples, C, highlighting smoking or alcohol status, D, highlighting degree of dysplasia, and E, highlighting gender

cancer development. To fulfil the growth requirements of the cells, the uptake of amino acids increases from the extracellular matrix leading to increased protein synthesis [46–49]. This explains the prominent protein bands in the patient samples, as the contralateral normal samples have shown increases in protein content similar to the potentially malignant lesion samples.

Similar to potentially malignant lesion samples, the contralateral normal samples have shown increases in protein which indicates that these cells are also
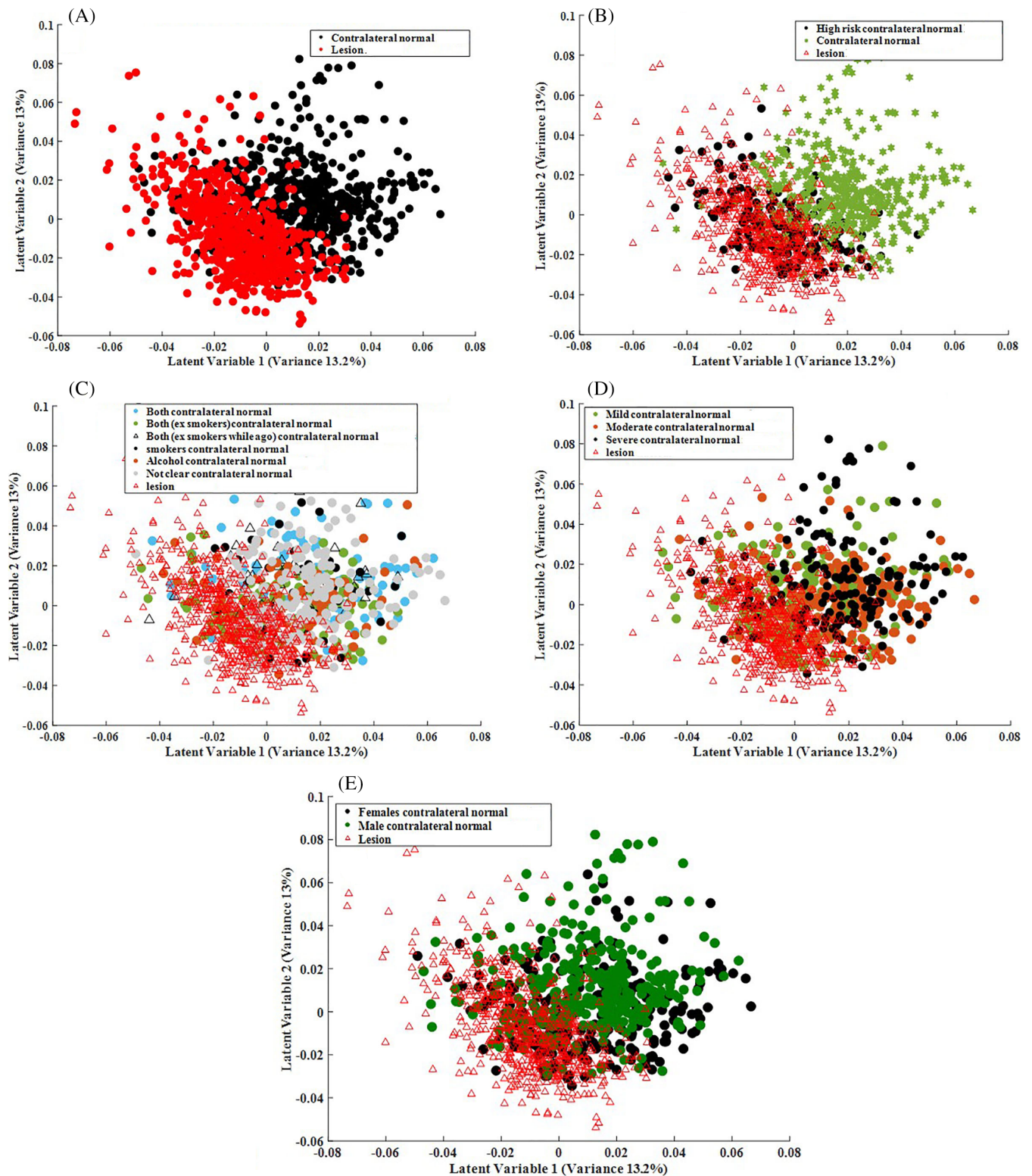
FIGURE 10 Score plot for A, potentially malignant lesion (lesion) samples and contralateral normal samples for the nuclear data set, B, for potentially malignant lesion (lesion) samples and contralateral normal samples highlighting high risk samples, C, highlighting smoking or alcohol status, D, highlighting degree of dysplasia, and E, highlighting gender

undergoing similar changes as the cells from the potentially malignant lesion. Hence, a PLS-DA model was developed for potentially malignant lesion samples and

contralateral normal samples from patients. Accuracies of 81% and 87% were achieved for cytoplasmic and nuclear data sets, as shown in Figures S5 and S6. A

number of contralateral normal samples were found to overlap with the potentially malignant lesion samples (Figures S3a and S4a) and, as observed for the buccal mucosa and tongue samples, these were identified as high risk samples (ie, samples with multiple lesions).

This suggests that these high risk samples may have an elevated probability of developing further lesions when compared to the rest of the patients, as the contralateral normal samples exhibit a similar spectral profile to that of the potentially malignant lesion samples. This can again be explained by the phenomenon of FC. Also, similar to the observations for buccal mucosa and tongue sites, in the model of gingiva and alveolus, the identified high risk samples have multiple lesions in the oral cavity, irrespective of the origin of the sampling site. Therefore, it can concluded that the results obtained from buccal mucosa and tongue samples are reproducible for gingiva and alveolus samples. This establishes the fact that the Raman spectroscopic technique can be used to detect the phenomenon of FC for the entire oral cavity.

Additionally, the effect of smoking and alcohol consumption habits, histopathology and gender were explored in contralateral normal samples. This was done to understand the factors contributing to the spectral overlap between lesion samples and contralateral normal samples in Figures 9A and 10A. No set pattern was observed, related to the history of smoking or alcohol consumption that could be an indicator of high risk samples for the cytoplasmic and nuclear data sets, respectively (Figures 9C and 10C). Similarly, the degree of dysplasia (mild, moderate and severe) of the potentially malignant samples had no influence on the high risk contralateral normal samples (Figures 9D and 10D). In addition, gender was not observed to be a confounding factor for high risk samples (Figures 9E and 10E). This suggested that FC might not be related to gender, smoking, alcohol consumption or degree of dysplasia of the potentially malignant lesion. It was expected that FC could arise due to the consumption of carcinogens [3, 11–13] or may be due to gender [11, 50], but through Raman spectroscopy these were not found to be a confounding factor.

The present explorative study suggests that the contralateral normal samples can be classified as spectrally different from the healthy volunteer samples. The results also suggest that the contralateral normal samples show spectral features similar to the lesion samples in patients who are at high risk of developing multiple lesions. This can be attributed to the phenomenon of FC and can be detected by Raman microspectroscopy. Thus, Raman spectroscopy has shown its efficacy in the detection of the FC effect in patients with oral PML.

4 | CONCLUSION

The phenomenon of FC has been known for some decades, and if it is present, the patient can be at higher risk of developing multiple lesions. This affects the overall treatment outcome of the patient as it is difficult to detect visually, and in general, biopsies would not be conducted from healthy sites of the oral cavity. Although high specification, research grade Raman microscopes can cost hundreds of thousands of Euro, increasingly, more compact systems of adequately high specification, are becoming available, reducing the initial capital outlay to some tens of thousands of Euro. Furthermore, the results of this study have paved the way for a possible route towards rapid and convenient detection of FC without the need of invasive biopsy. In future, it may lead to better prognosis and management of patients with PML, which in turn might increase the 5-year survival rate.

CONFLICT OF INTEREST
The authors declare no potential conflict of interest.

ORCID
Isha Behl https://orcid.org/0000-0003-0676-8988
Hugh J. Byrne https://orcid.org/0000-0002-1735-8610
Fiona M. Lyng https://orcid.org/0000-0002-9876-963X

REFERENCES
[1] GLOBOCAN: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2018. International Agency for Research on Cancer [online] 2018, http://globocan.iarc.fr/Pages/fact_sheets_population.aspx.
[2] B. W. Neville, T. Day, *Ca-A Cancer J. Clin.* **2002**, *52*(4), 195.
[3] A. C. Sathiasekar, D. G. Mathew, M. S. Jaish Lal, A. A. Arul Prakash, K. U. G. Kumar, *J Pharm Bioallied Sci.* **2017**, *9*(Suppl 1), S23.
[4] T. A. Winning, G. C. Townsend, *Clin. Dermatol.* **2000**, *18*(5), 499.
[5] C. A. Squier, M. J. Kremer, *J. Natl. Cancer Inst. Monogr.* **2001**, *52242*(29), 7.
[6] V. Kumar, N. Fausto, A. Abbas, *Robbins and Cotran Pathologic Basis of Disease*, 7th ed., Elsevier, Amsterdam, **2005**.
[7] A. K. El-Naggar, J. K. C. Chan, J. R. Grandis, T. Takata, P. J. Slootweg Eds., *WHO Classification of Head and Neck Tumours*, 4th ed., IARC, France, **2017**.
[8] S. Müller, *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **2018**, *125*(6), 591.
[9] C. F. Poh, C. E. MacAulay, D. M. Laronde, P. M. Williams, L. Zhang, M. P. Rosin, *Periodontol 2000.* **2011**, *57*(1), 73.

[10] D. P. Slaughter, H. W. Southwick, W. Smejkal, *Cancer* **1953**, *6*, 963.

[11] M. Mohan, N. Jagannathan, *Oncol. Rev.* **2014**, *8*(1), 244. Doi: 10.4081/oncol.2014.244.

[12] M. Aparna, P. Shenai, L. K. Chatra, K. M. Veena, P. K. Rao, R. V. Prabhu, K. A. Shahin, *Arch. Med. Heal. Sci.* **2013**, *1*(2), 136.

[13] M. G. C. T. van Oijen, P. J. Slootweg, *Cancer Epidemiol. Biomarkers Prev.* **2000**, *9*(3), 249.

[14] S. P. Singh, A. Deshmukh, P. Chaturvedi, C. M. Krishna, *J. Biomed. Opt.* **2012**, *17*(10), 1050021.

[15] A. Sahu, S. Tawde, V. Pai, P. Gera, P. Chaturvedi, S. Nair, C. M. Krishna, *Anal. Methods* **2015**, *7*, 7548.

[16] A. Sahu, N. Shah, M. Mahimkar, M. Garud, S. Pagare, S. Nair, C. M. Krishna, in *Proc. SPIE 8926, Photon Therap Diagn X*, 89262N, 2014.

[17] A. Sahu, A. Deshmukh, A. D. Ghanate, S. P. Singh, P. Chaturvedi, C. M. Krishna, *Technol. Cancer Res. Treat.* **2012**, *11*, 529.

[18] S. P. Singh, *PhD*, Tata Memorial Centre **2013**.

[19] S. P. Singh, A. Deshmukh, P. Chaturvedi, C. M. Krishna, *J. Cancer Res. Ther.* **2012**, *8*(6), 126.

[20] I. Behl, L. Kukreja, A. Deshmukh, S. P. Singh, H. Mamgain, A. R. Hole, C. M. Krishna, *J. Biomed. Opt.* **2014**, *19*(12), 126005.

[21] H. Krishna, S. K. Majumder, P. Chaturvedi, M. Sidramesh, P. K. Gupta, *J. Biophotonics* **2014**, *7*(9), 690.

[22] M. J. Jeng, M. Sharma, L. Sharma, T. Y. Chao, S. F. Huang, L. B. Chang, S. L. Wu, L. Chow, *J. Clin. Med.* **2019**, *8*(9), 1313.

[23] F. L. J. Cals, T. C. Bakker Schut, P. J. Caspers, R. J. Baatenburg De Jong, S. Koljenović, G. J. Puppels, *Analyst* **2018**, *143*(17), 4090.

[24] C. LFCS, F. Bonnier, C. Tellez, L. Dos Santos, K. O'Callaghan, J. O'Sullivan, S. LES, S. Flint, A. A. Martin, F. M. Lyng, H. J. Byrne, *Exp. Mol. Pathol.* **2017**, *103*(3), 255.

[25] F. L. Cals, T. C. Bakker Schut, J. A. Hardillo, R. J. Baatenburg de Jong, S. Koljenović, G. J. Puppels, *Lab. Invest.* **2015**, *95*(10), 1186.

[26] L. F. C. S. Carvalho, F. Bonnier, K. O. Callaghan, J. O. Sullivan, S. Flint, H. J. Byrne, F. M. Lyng, *Exp. Mol. Pathol.* **2015**, *98*(3), 502.

[27] L. F. C. S. Carvalho, M. S. Nogueira, T. Bhattacharjee, L. P. M. Neto, L. Daun, T. O. Mendes, R. Rajasekaran, M. Chagas, A. A. Martin, L. E. S. Soares, *Clin. Oral Investig.* **2019**, *23*(7), 3021.

[28] L. F. C. S. Carvalho, M5. S. Nogueira, *Analyst* **2018**, *143*(24), 6037.

[29] C. Wu, J. Gleysteen, N. T. Teraphongphom, Y. Li, E. Rosenthal, *Int. J. Oral Sci.* **2018**, *10*(2), 10.

[30] C. A. Lieber, H. E. Nethercott, M. H. Kabeer, *Biomed. Opt. Express* **2010**, *1*(3), 975.

[31] A. Robichaux Viehoever, D. Anderson, D. Jansen, A. Mahadevan-Jansen, *Photochem. Photobiol.* **2004**, *78*(5), 517.

[32] S. P. Singh, A. Sahu, A. Deshmukh, P. Chaturvedi, C. M. Krishna, *Analyst* **2013**, *138*, 4175.

[33] I. Behl, G. Calado, O. Ibrahim, A. Malkin, S. Flint, H. J. Byrne, F. M. Lyng, *Anal. Methods* **2017**, *9*(6), 937.

[34] F. Bonnier, D. Traynor, P. Kearney, C. Clarke, P. Knief, C. Martin, J. J. O'Leary, H. J. Byrne, F. M. Lyng, *Anal. Methods* **2014**, *6*, 7831.

[35] L. T. Kerr, B. M. Hennelly, *Chemom. Intel. Lab. Syst.* **2016**, *158*, 61.

[36] M. Urvoy, F. Autrusseau, in *Proc. 2nd ACM Work. Inf. Hiding Multimed. Secur. - IH&MMSec '14*, 2014, pp. 49–60.

[37] R. G. Brereton, G. R. Lloyd, *J. Chemometr.* **2014**, *28*(4), 213.

[38] R. Gautam, S. Vanga, F. Ariese, S. Umapathy, *EPJ Tech. Instrum.* **2015**, *2*(1), 8.

[39] I. Behl, G. Calado, A. Malkin, S. Flint, S. Galvin, C. M. Healy, M. L. Pimentel, H. J. Byrne, F. M. Lyng, *J. Biophotonics* **2020**, https://doi.org/10.1002/jbio.202000079.

[40] Z. Movasaghi, S. Rehman, I. Rehman, *Appl. Spectrosc.* **2007**, *42*(5), 493.

[41] H. Byrne, G. Sockalingum, N. Stone, Raman Microscopy: Complement or Competitor. in *Biomedical Applications of Synchrotron Infrared Microspectroscopy*, RSC Analytical Spectroscopy Series **2011**, p. 105-143.

[42] S. Beloribi-Djefaflia, S. Vasseur, F. Guillaumond, *Oncogenesis* **2016**, *5*(1), e189.

[43] G. Cascianelli, M. Villani, M. Tosti, F. Marini, E. Bartoccini, M. V. Magni, E. Albi, *Mol. Biol. Cell* **2008**, *19*(12), 5289. https://doi.org/10.1091/mbc.e08-05-0517.

[44] I Behl, *PhD*, TU Dublin, 2019.

[45] D. V. Dawson, D. R. Drake, J. R. Hill, K. A. Brogden, C. L. Fischer, P. W. Wertz, *Int. J. Cosmet. Sci.* **2013**, *35*(3), 220.

[46] R. J. DeBerardinis, A. Mancuso, E. Daikhin, I. Nissim, M. Yudkoff, S. Wehrli, C. B. Thompson, *Proc. Natl. Acad. Sci. USA* **2007**, *104*(49), 19345.

[47] E. L. Lieu, T. Nguyen, S. Rhyne, J. Kim, *Exp. Mol. Med.* **2020**, *52*(1), 15.

[48] R. J. DeBerardinis, S. C. Navdeep, *Oncology* **2016**, *2*, 1.

[49] L. M. Phan, S. C. J. Yeung, M. H. Lee, *Cancer Biol. Med.* **2014**, *11*(1), 1.

[50] J. Giovannoni, M. G. C. T. Van Oijen, P. J. Slootweg, M. Mohan, N. Jagannathan, *Oncol. Rev.* **2014**, *8*(1), 249.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.