Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

Raman spectral signatures of cervical exfoliated cells from liquid-based cytology samples

Padraig Kearney
Damien Traynor
Franck Bonnier
Fiona M. Lyng
John J. O'Leary
Cara M. Martin

SPIE.

Padraig Kearney, Damien Traynor, Franck Bonnier, Fiona M. Lyng, John J. O'Leary, Cara M. Martin, "Raman spectral signatures of cervical exfoliated cells from liquid-based cytology samples," *J. Biomed. Opt.* **22**(10), 105008 (2017), doi: 10.1117/1.JBO.22.10.105008.

Raman spectral signatures of cervical exfoliated cells from liquid-based cytology samples

Padraig Kearney,[a,b] Damien Traynor,[c,d] Franck Bonnier,[c,d,e] Fiona M. Lyng,[c,d,*,†] John J. O'Leary,[a,b,†] and Cara M. Martin[a,b,†]
[a]Coombe Women and Infants University Hospital, Department of Pathology, Dublin, Ireland
[b]Trinity College, School of Medicine, Department of Histopathology and Morbid Anatomy, Dublin, Ireland
[c]Focas Research Institute, Dublin Institute of Technology, DIT Centre for Radiation and Environmental Science, Dublin, Ireland
[d]Dublin Institute of Technology, School of Physics, Dublin, Ireland
[e]Université François-Rabelais de Tours, Faculty of Pharmacy, Tours, France

Abstract. It is widely accepted that cervical screening has significantly reduced the incidence of cervical cancer worldwide. The primary screening test for cervical cancer is the Papanicolaou (Pap) test, which has extremely variable specificity and sensitivity. There is an unmet clinical need for methods to aid clinicians in the early detection of cervical precancer. Raman spectroscopy is a label-free objective method that can provide a biochemical fingerprint of a given sample. Compared with studies on infrared spectroscopy, relatively few Raman spectroscopy studies have been carried out to date on cervical cytology. The aim of this study was to define the Raman spectral signatures of cervical exfoliated cells present in liquid-based cytology Pap test specimens and to compare the signature of high-grade dysplastic cells to each of the normal cell types. Raman spectra were recorded from single exfoliated cells and subjected to multivariate statistical analysis. The study demonstrated that Raman spectroscopy can identify biochemical signatures associated with the most common cell types seen in liquid-based cytology samples; superficial, intermediate, and parabasal cells. In addition, biochemical changes associated with high-grade dysplasia could be identified suggesting that Raman spectroscopy could be used to aid current cervical screening tests. [DOI: 10.1117/1.JBO.22.10.105008]

Keywords: Raman spectroscopy; cytology; ThinPrep; exfoliated cells; superficial cells; intermediate cells; parabasal cells; cervical cancer; cervical dysplasia.

Paper 170445R received Jul. 7, 2017; accepted for publication Oct. 3, 2017; published online Oct. 30, 2017.

1 Introduction

Cervical cancer is the fourth most common malignancy among women worldwide, with an estimated 528,000 cases diagnosed in 2012. Of these, 266,000 women died, with 87% of the deaths occurring in lower-income countries.[1] Cervical intraepithelial neoplasia (CIN) is the premalignant stage of cervical cancer and can be classified as low grade (CIN1) and high grade (CIN2/3).[2]

The gradual progression of cervical cancer allows for the application of screening programs to detect preneoplastic changes before invasive cancer develops. The Papanicolaou (Pap) test is the primary approach to cervical cancer screening. This is a screening test that allows detection of cytological abnormalities of the endocervix and ectocervix. A cervical smear is taken, which is then cytologically examined with the help of a Pap stain.[3] However, the Pap test displays a hugely variable sensitivity and specificity. Arbyn et al.[4] reported a sensitivity of 90.4% and a specificity of 64.6% of Pap cytology for detection of CIN2+, although these results are known to vary between hospitals.

Human papillomavirus (HPV) is the main causative agent of cervical cancer and so HPV DNA testing has recently been incorporated into cervical screening programs.[5] However, while the sensitivity of HPV testing exceeds that of cytology, HPV testing has not been able to match the specificity of cytology.[6]

The development of an alternative to the current approaches is long overdue. Raman spectroscopy is a vibrational spectroscopic technique that has the potential to combine the diagnostic advantages of both cytology and HPV testing. Raman spectroscopy can provide a rapid, label-free, nondestructive measurement of the complete biochemical fingerprint of cells or tissues. If Raman spectroscopy can be shown to be as good as, or better than, the current gold standard, then there is great potential for the technique to be used as an alternative or an adjunct to the Pap smear test, which is available through screening programs in most developed countries. In developing countries, however, because of a lack of resources and infrastructure, screening programs are not available and a "see and treat" approach based on visual inspection using acetic acid may be more appropriate in this setting.[7] With clinical translation to lower cost, automated Raman systems, the advantages would be higher sample throughput and less requirement for highly trained cytologists and pathologists to implement the screening program.

Over the past 15 years, the potential of Raman spectroscopy together with multivariate statistical analysis has been demonstrated for the detection of a variety of cancers, including cervical cancer.[8–10] The majority of Raman spectroscopy studies on cervical cancer has focused on cervical tissue *ex vivo*[11–15]

*Address all correspondence to: Fiona M. Lyng, E-mail fiona.lyng@dit.ie
†Joint senior authors

1083-3668/2017/$25.00 © 2017 SPIE

or *in vivo*,[16–20] with relatively few studies focusing on cervical cytology.[21–24] Extensive studies have, however, been carried out on exfoliated cervical cells using infrared spectroscopy. Early studies by Wong et al.[25] showed significant differences between normal, precancer, and cancer samples but as spectra were recorded from cell pellets rather than from single exfoliated cells, a number of confounding factors, such as polymorphs, endocervical columnar cells, metaplastic cells, cervical mucus, and debris, were subsequently identified.[26–29] More recent infrared studies using liquid-based cytology have shown significant spectral differences between normal, low-grade, and high-grade dysplasia.[30–32] Recent Raman spectroscopy studies showed that HPV-positive and HPV-negative cytology samples could be classified with an accuracy of 98.5%[21] and that normal and cervical cancer cytology samples could be classified with an accuracy of ~80%.[22] Both of these studies used cell pellets rather than recording Raman spectra from single exfoliated cells and this probably resulted in the relatively low classification accuracy in the study by Rubina et al.[22] due to sample heterogeneity. Recent studies from our group have shown that Raman spectra can be successfully recorded from single exfoliated cells from liquid-based cytology Pap test specimens.[23,24] Many of the issues involved in recording Raman spectra from cervical cytology specimens prepared according to the standard clinical protocols as single exfoliated cells on ThinPrep glass slides were addressed by Bonnier et al.[23] and a pretreatment of the slides with hydrogen peroxide to clear blood residue contamination before Raman recording was shown to minimize variability within the data sets resulting in highly reproducible data. Ramos et al.[24] reported sensitivity and specificity values of >90% for classification of negative, CIN1, CIN2, and CIN3 cervical cytology samples. Recently, Traynor et al.[33] investigated hormone-associated changes in the Raman spectra of cytologically negative ThinPrep cervical smear samples related to the menstrual cycle, the use of hormone-based contraceptives, and the onset of menopause, and showed that any spectral changes did not interfere with the ability to discriminate between normal and abnormal cervical cytology samples. This study follows on from this and focuses on defining the spectral signatures of different types of cervical exfoliated cells. Most Pap test specimens contain a variety of different cell types. These include superficial cells, intermediate cells, parabasal cells, white blood cells, endocervical cells, and squamous metaplastic cells. The most abundant of these cell types are superficial and intermediate cells, while parabasal cells are also commonly seen depending on certain factors. The ratio of superficial to intermediate cells is governed by the hormonal status of the patient, i.e., last menstrual period, use of oral contraceptive pill, or intrauterine contraceptive device.[34] As the cervix matures, the number of superficial cells tends to increase, while women taking the oral contraceptive pill tend to produce samples that contain a higher quantity of intermediate cells. In addition, the lack of estrogenic stimulation after menopause results in the loss of cervical cell differentiation, so more parabasal cells are seen in the smear. An increase in parabasal cells is also observed when the patient is exhibiting cervical atrophy.

While the infrared and Raman spectral signatures of superficial, intermediate, and parabasal cell layers have previously been investigated in cervical tissue,[14,15,35,36] to the best of our knowledge, this has not previously been investigated in single exfoliated cervical cells.

In this study, Raman spectral signatures of cervical epithelial cells—superficial, intermediate, and parabasal cells—are presented and it is shown that they can be discriminated from each other. In addition, the spectral signature of high-grade dysplastic cervical cells was compared to each of the normal cervical epithelial cell types.

2 Materials and Methods

2.1 *Sample Collection and Preparation*

Liquid-based cervical cytology samples collected in PreservCyt were obtained from the Cytopathology Department of the Coombe Women and Infants University Hospital. The study population was an anonymized group of women that had given a smear for routine screening. The cervical cells are collected by a clinician using a cervix brush or a combination of a spatula and cytobrush, before transfer of the cells into a ThinPrep vial containing a methanol-based fixative solution called PreservCyt. A total of 70 liquid-based cytology cervical samples, 20 confirmed as cytology negative, and 50 confirmed as high-grade dyskaryosis was used for this study. Ethical approval was obtained from the Research Ethics Committee of the Coombe Women and Infants University Hospital. Once routine cytological diagnosis based on the Pap test was performed, the residual material was used for this study. Cytological diagnoses are reported according to the Bethesda system, such that samples containing low- and high-grade abnormalities are classed, respectively, as low-grade squamous intraepithelial lesion (LSIL) and high-grade squamous intraepithelial lesion (HSIL). LSIL is usually indicative of CIN1, whereas HSIL is indicative of CIN2 or CIN3.[37]

A further set of liquid-based cytology cervical samples, confirmed as cytology negative, was obtained from patients who had an endometrial biopsy taken on the same day as the smear test ($n = 10$). The phase in the menstrual cycle, for each sample, could be confirmed by the corresponding endometrial biopsy. Samples from the proliferative phase (days 5 to 13) ($n = 4$), the secretory phase (days 14 to 28) ($n = 4$), and the interval phase (days 13 to 14) ($n = 2$) of the menstrual cycle were available.

A monolayer of cells was prepared from the ThinPrep vial with the use of the ThinPrep processor 2000. This is an automated system that disperses the cells in the solution, collects them on the TransCyt filter, and deposits them on a ThinPrep slide. The slide was then placed in a vial containing 95% alcohol. The slides also underwent a treatment step to remove any molecular contamination by hemoglobin, which obscures several features of the cellular spectrum.[23] Briefly, slides were treated with a 30% solution of H_2O_2 at room temperature for 3 min, followed by a 70% solution of industrial methylated spirits (IMS) for 3 min followed by multiple dips into 100% IMS to remove any remaining cellular debris and H_2O_2. The slide was then dried in air.

2.2 *Raman Spectroscopy*

A HORIBA Jobin Yvon XploRA system (Villeneuve d'Asq, France), which is coupled to an Olympus microscope BX41, was used for spectral acquisition. As a source, a 532-nm laser of ~12-mW power was focused onto the sample using a 100× objective (MPlanN, Olympus, NA = 0.9) giving a laser spot of 1 to 2 μm. The resultant Raman signals were

detected using a spectrograph with a 1200 g/mm grating coupled to a charge-coupled device. Raman spectra were acquired in the fingerprint region, 400 to 1800 cm^{-1}, with an integration time of 30 s averaged over two accumulations. For each slide, spectra from the nucleus or cytoplasm of at least 30 cells were recorded. The x and y coordinates for each cell were recorded. A Pap stain was then performed once a satisfactory number of cells were recorded. This allowed confirmation of the cytological status of the recorded cells using the previously saved coordinates on the glass slide. Both the glass coverslip and the Pap stain interfere with Raman measurements, and so the confirmatory Pap stain could be carried out only after all Raman measurements were taken.

2.3 *Data Analysis*

All preprocessing and analyses were carried out using MATLAB version 7.9 (2009). Preprocessing steps involved smoothing (Savitzky–Golay $k = 5, w = 13$), baseline correction (rubber band), and vector normalization. The contribution from the glass slide was removed using a nonnegatively constrained least squares method as described in more detail in Ref. 38.

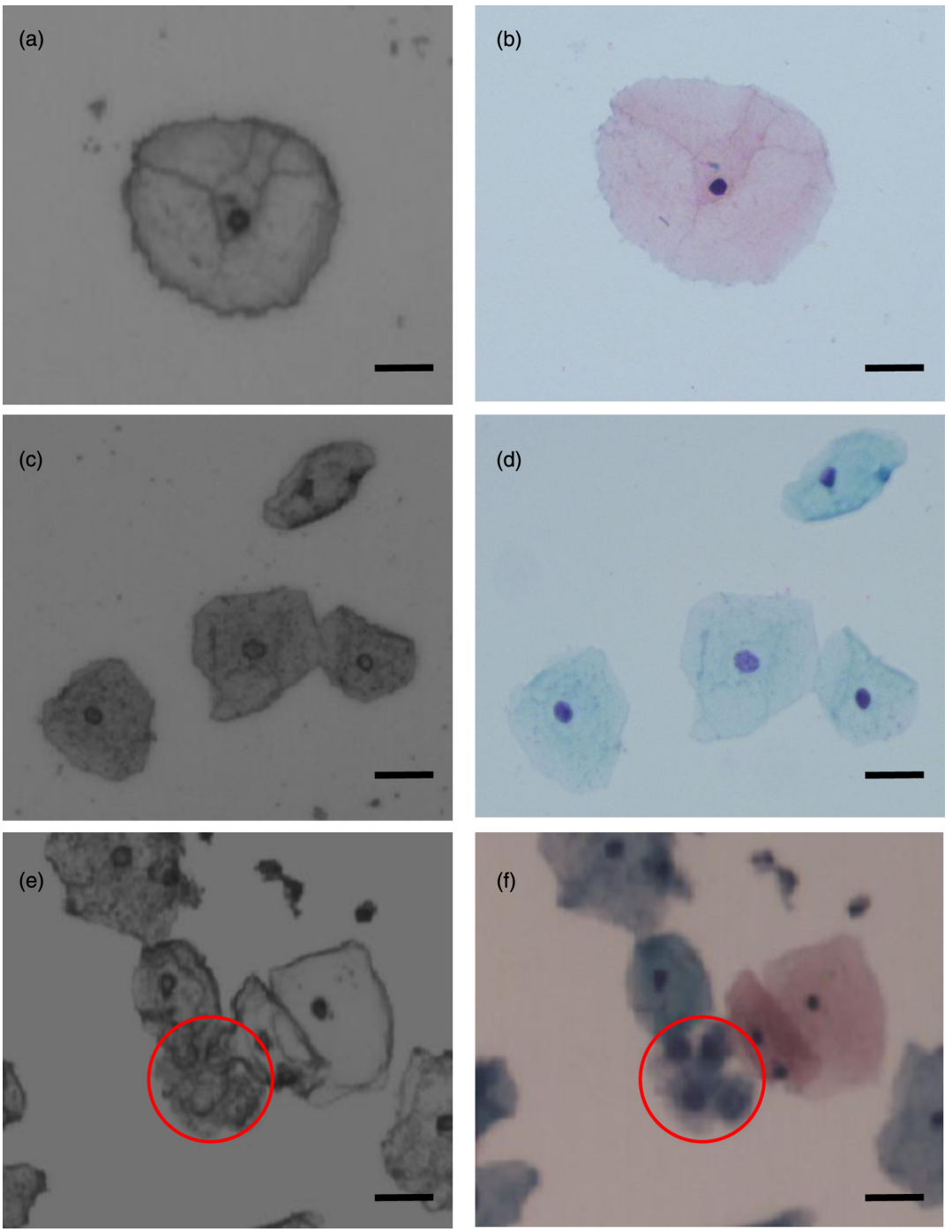

Fig. 1 Unstained and Pap-stained cervical exfoliated cells. (a) and (b) Superficial cells, (c) and (d) intermediate cells, and (e) and (f) parabasal cells (circled), unstained and stained, respectively. Bar = 20 μm.

The data were then subjected to principal component analysis (PCA), an unsupervised multivariate statistical technique allowing evaluation of the variability existing in the data sets.[39]

Moreover, to better appreciate the discrimination achieved by comparison of the Raman spectra, a factorial discriminant analysis (FDA) has been performed on the scores calculated from the PCA.[39] The optimal number of principal components (PCs) to generate the PCA–FDA model was found to be 10 PCs. The output of the analysis can be represented by means of a confusion matrix with all data points attributed to the groups identified by the operator (superficial, intermediate, parabasal, and HSIL cells). The statistical relevancy and robustness of the analysis have been ensured using a 100-fold cross validation, where two-thirds of the spectra have randomly been selected as the calibration set, whereas the remaining third is used as the validation set, thus over 100 independent iterations. The overall classification obtained is represented by the sensitivity and specificity.

3 Results and Discussion

3.1 *Raman Spectral Signatures of Superficial, Intermediate, and Parabasal Cells*

The main aim of the study was to define the spectral signatures of the normal cervical epithelial cell types present on a Pap test specimen. Figure 1 shows typical superficial, intermediate, and parabasal cells on a Pap stained ThinPrep slide. When unstained, superficial and intermediate cell types can be difficult to differentiate [Figs. 1(a) and 1(c)]. The smallest of the three cell types is the parabasal cells [Fig. 1(e)]. As the parabasal cells mature into intermediate and superficial cells, the nucleus shrinks and the cell enlarges. Thus, the superficial and intermediate cells can be differentiated quite well visually from the parabasal cells on the basis of the nucleus to cytoplasm ratio. After Pap staining, superficial cells tend to stain pink due to their metabolic inactivity [Fig. 1(b)], while the metabolically active intermediate and parabasal cells are usually stained blue [Figs. 1(d) and 1(f)].

3.1.1 *Raman spectral signatures of superficial, intermediate, and parabasal cell nuclei*

Figure 2(a) shows mean Raman spectra recorded from the nuclei of cytologically normal cervical exfoliated cells. Cells were confirmed as superficial, intermediate, or parabasal by Pap staining after Raman spectral recording. Noticeable spectral differences between the nuclei of superficial, intermediate, and parabasal cells can be seen at 480, 782, 1339, 1450, and 1670 cm^{-1}. A separation of the groups can be seen in the PCA scatter plot [Fig. 2(b)], with PC1 almost completely separating the superficial cell spectra from the intermediate cell spectra, and PC2 providing reasonable separation of the parabasal cell spectra from the spectra of the superficial and intermediate cells. Some degree of overlap between the different groups might be expected given the progressive maturation of the cells, some cells may be between parabasal and intermediate stages and some between the intermediate and superficial stages. The loadings corresponding to PC1 and PC2, shown in Fig. 2(c), confirm the visual observations, with the majority of the discrimination coming from the peaks mentioned previously. In the PCA scatter plot, the spectra of the superficial cells are on the negative side of PC1 and the spectra of the intermediate cells are on the positive side. Thus, the negative features of the PC1 loading elucidate the discriminating Raman features, which can be attributed to the superficial cells, including glycogen and nucleic acid bands at 480, 1380, and 1490 cm^{-1} and the positive features can be attributed to the intermediate cells, including protein bands at 1450 and 1670 cm^{-1}.[12,13] Despite the Raman spectra being recorded from the nuclei, glycogen features are stronger in the superficial cells as would be expected as these cells are more mature. Signals from glycogen would have been collected from the cytoplasm overlying the nuclei in the superficial cells. In addition, the superficial cell nuclei are denser than the intermediate cell nuclei resulting in compacted chromatin, which produces a stronger Raman signal. The more prominent protein features in the intermediate cells may be due to a greater degree of transcription in these cells in comparison to the superficial cells, as intermediate cells are still metabolically active, in comparison to the mature superficial cells, which are at the end of their life cycle. The parabasal cell spectra are reasonably well discriminated from both cell types by PC2. The PC2 loading

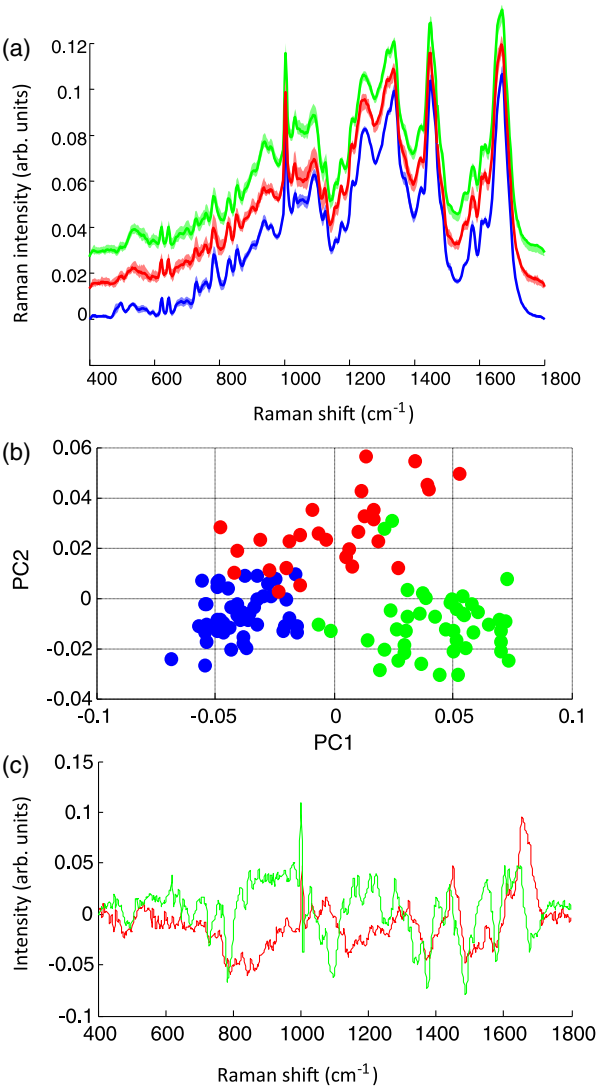

Fig. 2 (a) Mean Raman spectra of superficial (blue), intermediate (green), and parabasal (red) cell nuclei. Shading denotes the standard deviation. Spectra have been offset for clarity. (b) PCA scatter plot and (c) loadings corresponding to the first (red) and second (green) PCs.

mainly consists of negative features, which can be attributed to the superficial and intermediate cells, with contributions coming from peaks at 480 and 1339 cm^{-1} (glycogen), 722, 782, 1096, 1485, and 1576 cm^{-1} (nucleic acids), and 1670 cm^{-1} (proteins). It would be expected that the parabasal cells would have stronger nucleic acid features as they are immature, rapidly proliferating cells and this has previously been shown for cervical tissue *ex vivo*, where spectral maps were recorded across the different cell layers.[14,15] As spectra were recorded from cell nuclei, the finding here most likely reflects the large difference in nuclear size between the parabasal cells and the superficial and intermediate cells resulting in less compacted chromatin and weaker Raman signals.

3.1.2 *Raman spectral signatures of superficial and intermediate cell cytoplasm*

Figure 3(a) shows Raman spectra recorded from the cytoplasm of cytologically normal cervical exfoliated cells. Cells were confirmed as superficial or intermediate by Pap staining after Raman spectral recording. It was not possible to record spectra from the cytoplasm of parabasal cells as the nucleus is very large in these immature cells and there is very little cell cytoplasm. Very distinct spectral differences between the cytoplasm of superficial and intermediate cells can be seen at 480, 850, 938, 1082, and 1336 cm^{-1}. These characteristic glycogen bands[12,13] are more prominent in the spectra from the cytoplasm

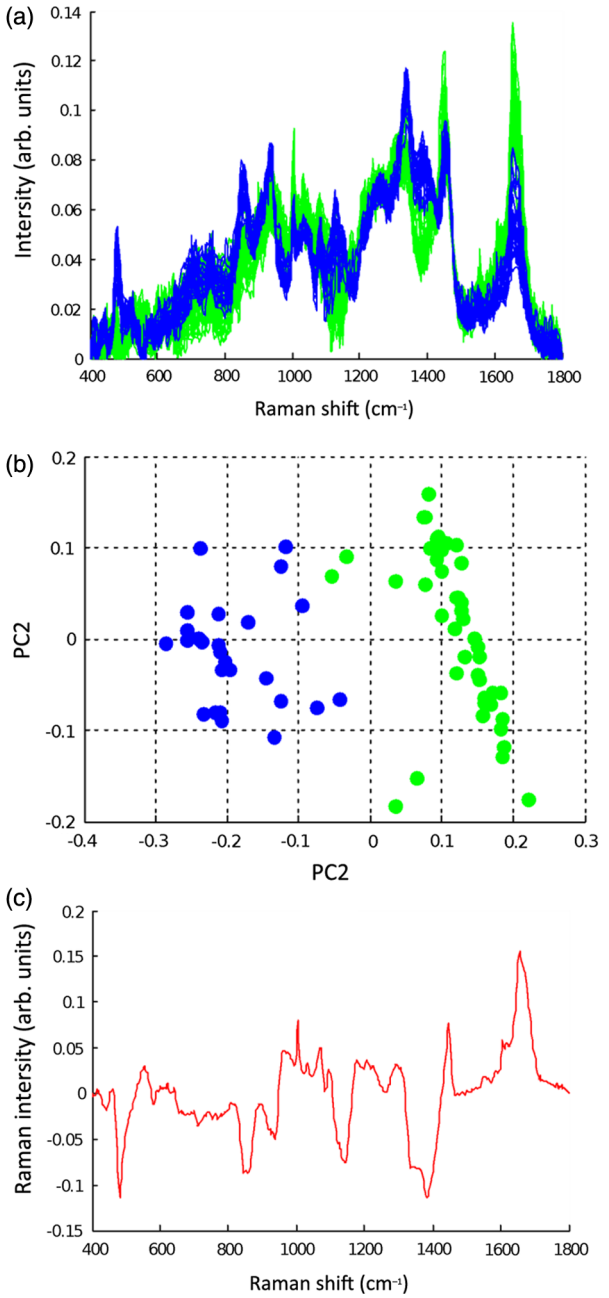

Fig. 3 (a) Raman spectra of superficial (blue) and intermediate (green) cell cytoplasm. (b) PCA scatter plot and (c) loadings corresponding to the first PC.

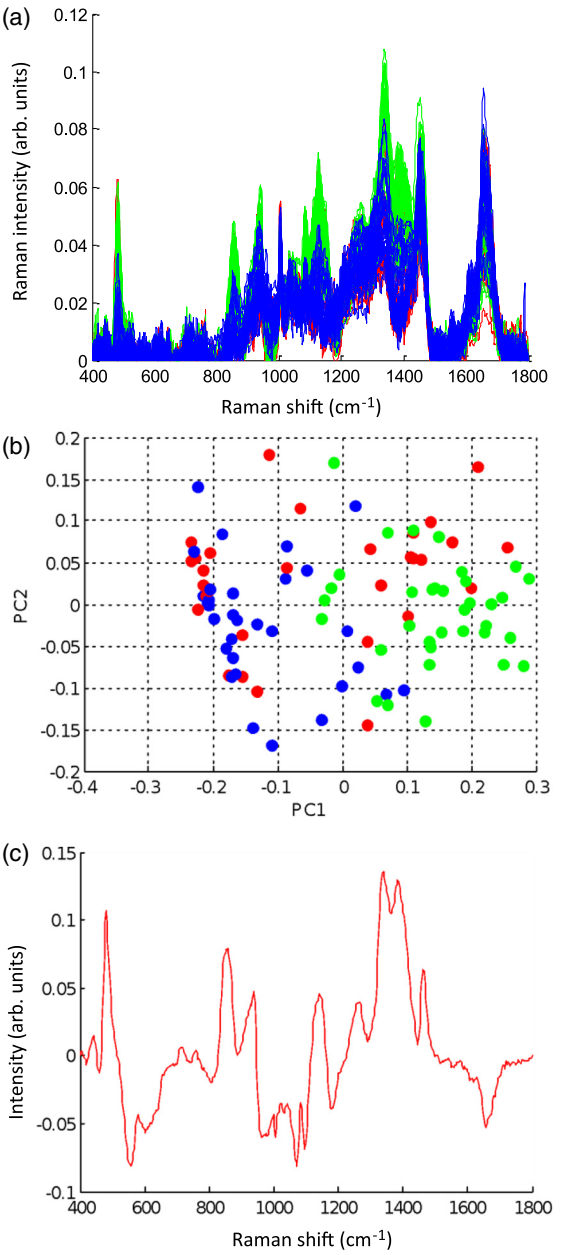

Fig. 4 (a) Raman spectra of the cytoplasm of superficial cells from samples taken at different days of the menstrual cycle, days 5 to 13 (blue), days 13 to 14 (red), and days 14 to 28 (green), (b) PCA scatter plot, and (c) loadings corresponding to the first PC.

of superficial cells. A separation of the two groups can be seen in the PCA scatter plot [Fig. 3(b)], with PC1 almost completely separating the superficial cell spectra from the intermediate cell spectra. The loadings corresponding to PC1 are shown in Fig. 3(c) with negative peaks corresponding to the superficial cells, including glycogen bands at 480, 850, 938, 1082, and 1336 cm^{-1}. Positive peaks correspond to the intermediate cells and are dominated by protein bands at 1450 and 1670 cm^{-1}. As observed previously with the spectra from the cell nuclei, glycogen features are stronger in the superficial cells as would be expected as these cells are more mature and accumulate more glycogen than the intermediate cells.[34] More prominent protein features were again observed in the spectra from the intermediate cells as these cells are metabolically active compared to the superficial cells.

These findings agree well with previous studies on cervical tissue *ex vivo*, where the main spectral difference observed was increased glycogen features in the superficial cell layer.[14,15]

Raman spectra from the cytoplasm of the superficial cells were found to be quite variable, within individual cells and between samples. To investigate this further, an additional set of cytologically normal samples was obtained, where the phase of the menstrual cycle had been confirmed on an endometrial biopsy.

Figure 4(a) shows Raman spectra recorded from the cytoplasm of cytologically normal cervical exfoliated cells from samples taken on different days of the menstrual cycle. Only cells that were confirmed as superficial by Pap staining after Raman spectral recording were used for this analysis.

Raman spectra taken from the cytoplasm of superficial cells from the secretory phase (days 14 to 28) displayed prominent glycogen related peaks at 480, 850, 938, 1125, 1339, and 1380 cm^{-1},[12,13] compared to the superficial cell spectra from the proliferative phase (days 5 to 13).

A reasonable separation of the two groups along PC1 can be seen in the PCA scatter plot [Fig. 4(b)]. The loadings corresponding to PC1 are shown in Fig. 4(c) with positive peaks corresponding to the superficial cells at days 14 to 28, including glycogen bands at 480, 850, 938, 1125, 1339, and 1380 cm^{-1}. The spectra recorded from the superficial cells at days 13 to 14 can be seen across both sides of PC1.

Similar findings were observed in an infrared spectroscopy study by Romeo et al.[40] where spectral changes in the 1200 to 1000 cm^{-1} region due to glycogen were reported throughout the menstrual cycle. However, a recent study from our group showed some changes in glycogen related peaks in the spectra from the nuclei of intermediate cells but no significant difference in glycogen in the spectra from the nuclei of superficial cells across the menstrual cycle.[33] As spectra were recorded from the cell nuclei in the study by Traynor et al.[33] rather than from the cytoplasm as in this study or from cell pellets as in the study by Romeo et al.,[40] no significant variability due to glycogen was observed and Raman spectroscopy was shown to discriminate well between negative cytology samples taken across different days of the menstrual cycle and HSIL cytology samples.

3.2 *Comparison of the Spectral Signature of HSIL Cell Nuclei with that of Superficial, Intermediate, and Parabasal Cell Nuclei*

A second aim of the study was to compare the spectral signature of HSIL cells to each of the normal cervical epithelial cell types, superficial, intermediate, and parabasal cells. Spectra were not

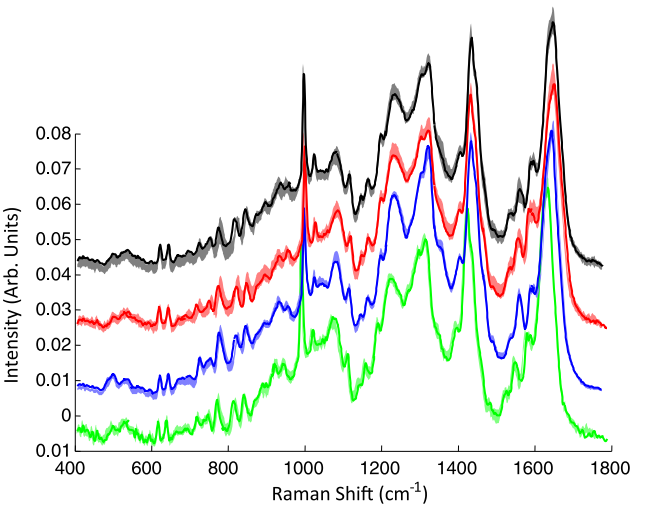

Fig. 6 Mean Raman spectra of superficial (blue), intermediate (green), parabasal (red), and HSIL (black) cell nuclei. Shading denotes the standard deviation. Spectra have been offset for clarity.

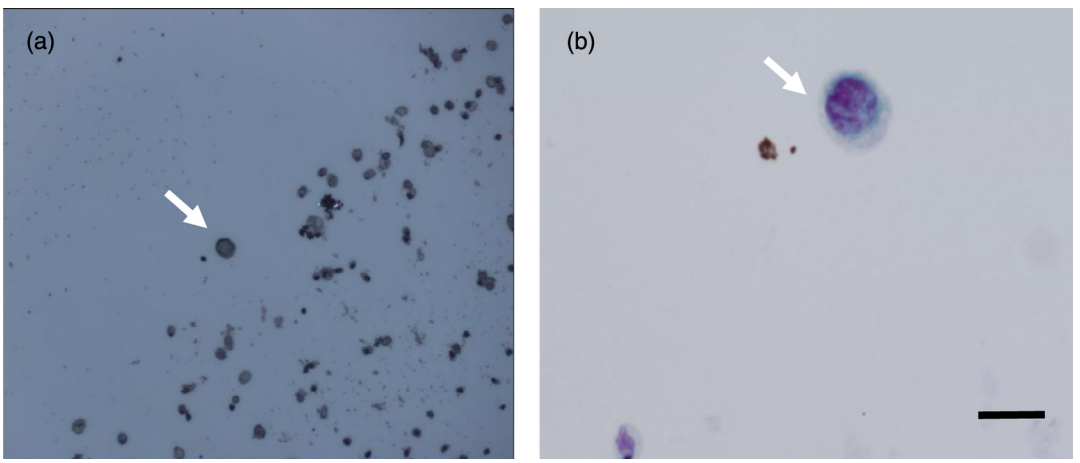

Fig. 5 (a) Unstained and (b) pap-stained cervical exfoliated cell confirmed as HSIL (arrow). Bar = 20 μm.

recorded from the cytoplasm as HSIL cells consist almost entirely of an enlarged nucleus, with very little cytoplasm (Fig. 5).

Figure 6 shows mean Raman spectra recorded from the nuclei of HSIL and normal cervical exfoliated cells. Cells were confirmed as HSIL, superficial, intermediate, or parabasal by Pap staining after Raman spectral recording. Spectral differences between the nuclei of HSIL cells and superficial, intermediate, and parabasal cells were observed at 480, 850, 938, 1125, 1339, 1450, and 1650 cm^{-1}.

PCA was performed, where the spectra from HSIL cell nuclei were compared to the spectra from nuclei of each normal cell type. Each PCA plot and the corresponding loadings are shown in Fig. 7. Figure 7(a) shows good discrimination between superficial cells and HSIL cells along PC1. In the PCA scatter plot, the spectra of the superficial cells are on the negative side of PC1 and the spectra of the HSIL cells are on the positive side. Thus, the negative loadings of PC1 elucidate the discriminating Raman features, which can be attributed to the superficial cells, including glycogen peaks at 480, 850, 938, 1333, and 1380 cm^{-1} and nucleic acid peaks at 780, 1490, and 1580 cm^{-1} [Fig. 7(b)]. Figure 7(c) shows good discrimination between intermediate cells and HSIL cells along PC1. Again, the negative loadings of PC1 display the Raman features, which can be attributed to the intermediate cells, including a glycogen-related peak at 1333 cm^{-1} and protein peaks at 1450 and 1650 cm^{-1}

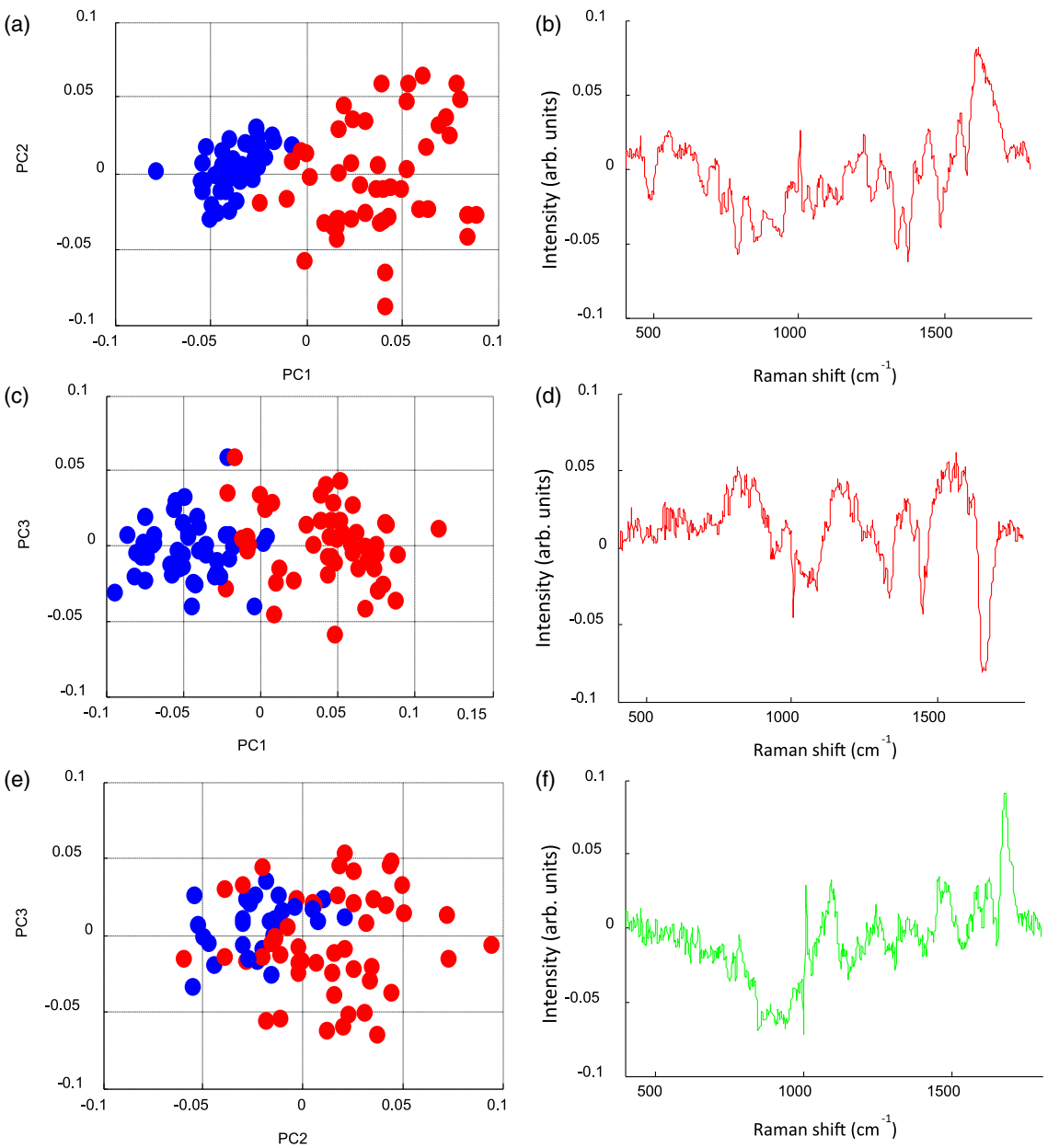

Fig. 7 (a) PCA scatter plot of spectra from the nuclei of superficial cells (blue) and HSIL cells (red) and (b) loadings corresponding to the first PC. (c) PCA scatter plot of spectra from the nuclei of intermediate cells (blue) and HSIL cells (red) and (d) loadings corresponding to the first PC. (e) PCA scatter plot of spectra from the nuclei of parabasal cells (blue) and HSIL cells (red) and (f) loadings corresponding to the second PC.

Table 1 Sensitivity and specificity for PCA–FDA classification of negative cytology versus HSIL cytology using spectra from superficial, intermediate, parabasal cells, or all cell types combined.

	Sensitivity (%)	Specificity (%)
Superficial cells (negative) versus HSIL	99	100
Intermediate cells (negative) versus HSIL	96	99
Parabasal cells (negative) versus HSIL	89	90
Combined negative cells versus HSIL	92	97

[Fig. 7(d)]. Significant overlap was observed in the PCA scatter plot of the spectra from the nuclei of the parabasal cells and the HSIL cells [Fig. 7(e)]. This might be expected given their morphological and potential biochemical similarity as HSIL cells arise from the basal/parabasal layer of cervical tissue. Some discrimination was observed, however, along PC2. The positive loadings of PC2 can be attributed to the HSIL cells with peaks at 1080, 1450, and 1650 cm^{-1} [Fig. 7(f)].

To further quantify the ability of Raman spectroscopy to discriminate between negative and HSIL cytology, PCA–FDA was performed using spectra from superficial, intermediate, parabasal cells, or from all cell types combined (Table 1). Table 1 shows very good classification of both superficial and intermediate cells and HSIL cells (sensitivity of 99% and 96% and specificity of 100% and 99%, respectively). Classification of parabasal and HSIL cells is more challenging because of their morphological and biochemical similarity as noted earlier, but despite this, good sensitivity and specificity were achieved (sensitivity of 89% and specificity of 90%). Interestingly, when all negative cytology cell types were combined (superficial, intermediate, and parabasal cells), good classification of negative cytology and HSIL cytology could be achieved (sensitivity of 92% and specificity of 97%).

In previous studies on cervical exfoliated cells, similar features of glycogen (482, 852, 937, 1339, and 1380 cm^{-1}) and nucleic acids (786, 1485, and 1580 cm^{-1}) were reported to be more prominent in superficial and intermediate cells confirmed as cytology negative compared to HSIL cells.[22,23] The more prominent protein features observed in the HSIL cells compared to the normal parabasal cells in this study may be reflective of a transforming HPV infection controlling the cellular machinery in the HSIL cells.

4 Conclusion

In this study, specific Raman spectral signatures were shown for each of the three most common cell types seen in ThinPrep liquid-based cytology samples, superficial, intermediate, and parabasal cells. Moreover, using PCA, it was shown that the spectra of each of these cell types could be differentiated from one another. An increase in glycogen (480, 850, 938, 1082, and 1336 cm^{-1}) was observed as cells matured from parabasal to intermediate to superficial cells. Intermediate cell spectra exhibited stronger protein related peaks (1450 and 1670 cm^{-1}) suggesting a greater degree of protein production in the intermediate cells as these cells are still metabolically active in comparison to the mature superficial cells. Intermediate and superficial cells could be discriminated from parabasal cells on the basis of more prominent glycogen, nucleic acid, and protein features, most likely because of the larger nucleus to cytoplasm ratio of parabasal cells resulting in weaker Raman signals. Raman spectra from the cytoplasm of superficial cells showed very prominent glycogen peaks, which varied depending on where in the cytoplasm the spectra were recorded but also between samples. This variability between samples was shown to be due to different levels of glycogen in different phases of the menstrual cycle. In the comparisons to the HSIL cells, only Raman spectra from the nuclei of superficial, intermediate, and parabasal cells were used because of the variability observed in the spectra from the cytoplasm and also because the HSIL cells were composed of a large nucleus with very little cytoplasm. Glycogen (482, 852, 937, 1339, and 1380 cm^{-1}) and nucleic acid (786, 1485, and 1580 cm^{-1}) features were observed to be more prominent in superficial and intermediate cells compared to HSIL cells. Interestingly, more prominent protein features were observed in the HSIL cells compared to the parabasal cells potentially as a result of a transforming HPV infection. Sensitivity of >96% and specificity of >99% were achieved for classification of superficial and intermediate cells from HSIL cells and sensitivity of 89% and specificity of 90% were achieved for classification of parabasal cells and HSIL cells despite the morphological and biochemical similarity. Good classification was also achieved when all negative cytology cell types (superficial, intermediate, and parabasal cells) were combined (sensitivity of 92% and specificity of 97%).

Good quality Raman signals with minimal variability could be recorded from the nucleus of the normal and abnormal cells. Raman spectra recorded from the cytoplasm showed a lot more variability due to glycogen accumulations in the superficial cells and this was also found to vary depending on the phase of the menstrual cycle.

Intermediate and superficial cells are the most common cell types present on a liquid-based cytology Pap test specimen. Although it can be difficult to distinguish the two cell types morphologically on the unstained slide, they can be discriminated using Raman spectroscopy based on their biochemical signatures. In addition, both cell types can be discriminated from HSIL cells again based on their biochemical signatures. Parabasal cells are less commonly seen on most Pap test specimens and although they have a similar morphology to HSIL cells, they can be discriminated from HSIL cells based on their biochemical signature. This study has shown the potential of biochemical fingerprinting using Raman spectroscopy for cervical cancer screening using standard liquid-based cytology preparation methods.

Disclosures

There are no conflicts of interest.

Acknowledgments

This research was undertaken as part of CERVIVA, the Irish Cervical Screening Research Consortium and we gratefully acknowledge funding from the Health Research Board Collaborative Applied Research Grant, CARG2012/29, and Enterprise Ireland cofunded by the European Regional

Development Fund and Ireland's EU Structural Funds Programme 2007–2013, CF2011 1045. We thank the cytology staff at the Coombe Women and Infants University Hospital, Dublin, for facilitating the study and technical support staff at the FOCAS Research Institute, Dublin Institute of Technology.

References

1. J. Ferlay et al., "Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer* **136**, E359–E386 (2015).
2. C. M. Martin and J. J. O'Leary, "Histology of cervical intraepithelial neoplasia and the role of biomarkers," *Best Pract. Res. Clin. Obstet. Gynaecol.* **25**, 605–615 (2011).
3. K. Nanda et al., "Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review," *Ann. Intern. Med.* **132**, 810–819 (2000).
4. M. Arbyn et al., "Liquid compared with conventional cervical cytology: a systematic review and meta-analysis," *Obstet. Gynecol.* **111**, 167–177 (2008).
5. G. Ronco et al., "Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials," *Lancet* **383**, 524–532 (2014).
6. M. Arbyn et al., "Human papillomavirus testing versus repeat cytology for triage of minor cytological cervical lesions," *Cochrane Database Syst. Rev.* **28**, CD008054 (2013).
7. N. Thekkek and R. Richards-Kortum, "Optical imaging for cervical cancer detection: solutions for a continuing global problem," *Nat. Rev. Cancer* **8**, 725–731 (2008).
8. M. Diem et al., "Molecular pathology via IR and Raman spectral imaging," *J. Biophotonics* **6**, 855–886 (2013).
9. K. Kong et al., "Raman spectroscopy for medical diagnostics—from in-vitro biofluid assays to in-vivo cancer detection," *Adv. Drug Delivery Rev.* **89**, 121–134 (2015).
10. F. M. Lyng et al., "Raman spectroscopy for screening and diagnosis of cervical cancer," *Anal. Bioanal. Chem.* **407**, 8279–8289 (2015).
11. C. M. Krishna et al., "Vibrational spectroscopy studies of formalin-fixed cervix tissues," *Biopolymers* **85**, 214–221 (2007).
12. F. M. Lyng et al., "Vibrational spectroscopy for cervical cancer pathology, from biochemical analysis to diagnostic tool," *Exp. Mol. Pathol.* **82**, 121–129 (2007).
13. L. E. Kamemoto et al., "Near-infrared micro-Raman spectroscopy for in vitro detection of cervical cancer," *Appl. Spectrosc.* **64**, 255–261 (2010).
14. K. M. Tan, C. S. Herrington, and C. T. A. Brown, "Discrimination of normal from pre-malignant cervical tissue by Raman mapping of de-paraffinized histological tissue sections," *J. Biophotonics* **4**, 40–48 (2011).
15. N. Rashid et al., "Raman microspectroscopy for the early detection of pre-malignant changes in cervical tissue," *Exp. Mol. Pathol.* **97**, 554–564 (2014).
16. A. Mahadevan-Jansen et al., "Development of a fiber optic probe to measure NIR Raman spectra of cervical tissue in vivo," *Photochem. Photobiol.* **68**, 427–431 (1998).
17. E. M. Kanter et al., "Application of Raman spectroscopy for cervical dysplasia diagnosis," *J. Biophotonics* **2**, 81–90 (2009).
18. E. Vargis et al., "Effect of normal variations on disease classification of Raman spectra from cervical tissue," *Analyst* **136**, 2981–2987 (2011).
19. S. Duraipandian et al., "Simultaneous fingerprint and high-wavenumber confocal Raman spectroscopy enhances early detection of cervical precancer in vivo," *Anal. Chem.* **84**, 5913–5919 (2012).
20. S. Duraipandian et al., "Near-infrared-excited confocal Raman spectroscopy advances in vivo diagnosis of cervical precancer," *J. Biomed. Opt.* **18**, 067007 (2013).
21. E. Vargis et al., "Near-infrared Raman microspectroscopy detects high-risk human papillomaviruses," *Transl. Oncol.* **5**, 172–179 (2012).
22. S. Rubina et al., "Raman spectroscopic study on classification of cervical cell specimens," *Vib. Spectrosc.* **68**, 115–121 (2013).
23. F. Bonnier et al., "Processing ThinPrep cervical cytological samples for Raman spectroscopic analysis," *Anal. Methods* **6**, 7831–7841 (2014).
24. I. Ramos et al., "Raman spectroscopy for cytopathology of exfoliated cervical cells," *Faraday Discuss.* **187**, 187–198 (2016).
25. P. T. T. Wong et al., "Infrared-spectroscopy of exfoliated human cervical cells—evidence of extensive structural-changes during carcinogenesis," *Proc. Natl. Acad. Sci. U. S. A.* **88**, 10988–10992 (1991).
26. B. R. Wood et al., "FTIR microspectroscopic study of cell types and potential confounding variables in screening for cervical malignancies," *Biospectroscopy* **4**, 75–91 (1998).
27. M. J. Romeo et al., "Influence of benign cellular changes in diagnosis of cervical cancer using IR microspectroscopy," *Biopolymers* **67**, 362–366 (2002).
28. P. T. T. Wong et al., "Detailed account of confounding factors in interpretation of FTIR spectra of exfoliated cervical cells," *Biopolymers* **67**, 376–386 (2002).
29. L. Chiriboga et al., "Infrared spectroscopy of human tissue. II. A comparative study of spectra of biopsies of cervical squamous epithelium and of exfoliated cervical cells," *Biospectroscopy* **4**, 55–59 (1998).
30. J. M. Schubert et al., "Spectral cytopathology of cervical samples: detecting cellular abnormalities in cytologically normal cells," *Lab. Invest.* **90**, 1068–1077 (2010).
31. N. C. Purandare et al., "Biospectroscopy insights into the multi-stage process of cervical cancer development: probing for spectral biomarkers in cytology to distinguish grades," *Analyst* **138**, 3909–3916 (2013).
32. K. Gajjar et al., "Histology verification demonstrates that biospectroscopy analysis of cervical cytology identifies underlying disease more accurately than conventional screening: removing the confounder of discordance," *PLoS One* **9**, e82416 (2014).
33. D. Traynor et al., "A study of variability due to hormonal effects in cervical smear samples using Raman spectroscopy," *J. Biophotonics* (2017).
34. R. L. Cowell and A. C. Valenciano, *Koss' Diagnostic Cytology and its Histopathologic Bases*, Vol. **1**, Lippincott Williams and Wilkins, Philadelphia (2006).
35. L. Chiriboga et al., "Infrared spectroscopy of human tissue. IV. Detection of dysplastic and neoplastic changes of human cervical tissue via infrared microscopy," *Cell. Mol. Biol.* **44**, 219–229 (1998).
36. B. R. Wood et al., "Fourier transform infrared (FTIR) spectral mapping of the cervical transformation zone, and dysplastic squamous epithelium," *Gynecol. Oncol.* **93**, 59–68 (2004).
37. R. Nayar and D. C. Wilbur, "The Pap test and Bethesda 2014," *Cancer Cytopathol.* **123**, 271–281 (2015).
38. O. Ibrahim et al., "Improved protocols for pre-processing Raman spectra of formalin fixed paraffin preserved tissue sections," *Anal. Methods* **9**, 4709–4717 (2017).
39. F. Bonnier and H. J. Byrne, "Understanding the molecular information contained in principal component analysis of vibrational spectra of biological systems," *Analyst* **137**, 322–332 (2012).
40. M. J. Romeo, B. R. Wood, and D. McNaughton, "Observing the cyclical changes in cervical epithelium using infrared microspectroscopy," *Vib. Spectrosc.* **28**, 167–175 (2002).

Padraig Kearney holds his BSc degree in microbiology from Trinity College Dublin and his MSc degree in biomedical diagnostics from Dublin City University. He completed his PhD in 2016, which involved developing a novel approach to cervical cancer screening, using Raman spectroscopy of individual cervical cells in ThinPrep cervical cytology samples.

Damien Traynor holds his BSc degree in medical and molecular cytology from the Dublin Institute of Technology (DIT). He achieved membership of the Academy of Laboratory Science in cytology in 2012 and worked as a primary screener for two years before joining DIT as a research technician. He is working on a Health Research Board funded clinical evaluation of Raman spectroscopy for the diagnosis of cervical precancer and is working on his PhD part time.

Franck Bonnier holds his MSc and PhD degrees from Université de Reims Champagne-Ardenne. He is a lecturer in analytical chemistry in Université François-Rabelais de Tours. His research is focused on biomedical applications of infrared and Raman spectroscopy supported by advanced data mining protocols. He has published 44 peer-reviewed research papers, and has a h-index of 13 and >1200 citations.

Fiona M. Lyng holds her BSc degree from Trinity College Dublin and her PhD from University College Dublin. She is a head of the DIT Centre for Radiation and Environmental Science. Her research is focused on translational research for cancer diagnosis and cancer treatment. She has published over 100 peer-reviewed research papers, and has a h-index of 33 and >3500 citations.

John J. O'Leary is a professor/chair of pathology at the Trinity College Dublin and a director of pathology at the Coombe Women and Infants University Hospital, Dublin. He heads a multi-investigator team focused on the molecular characterization of several cancer systems including: ovary, cervix, prostate, thyroid, head and neck cancer, and cancer stem cell biology. He has published more than 190 peer-reviewed papers, and has a h-index of 49 and >9600 citations.

Cara M. Martin is an assistant professor in molecular pathology and tumor biology in the Department of Histopathology at the Trinity College Dublin. She leads the Cervical Cancer Research Group based in the Coombe Women and Infant's University Hospital. Her research program consists of translational health services-based research in female gynecological cancers. She has published 57 peer-reviewed articles and 13 book chapters, and has a h-index of 23 and >2500 citations.