2007-01-01

Vibrational Spectroscopy for Pathology from Biochemical Analysis to Diagnostic Tool

Fiona M. Lyng
Dublin Institute of Technology, Fiona.lyng@dit.ie

Eoghan O'Faolain
Dublin Institute of Technology

J. Conroy
Trinity College Dublin

Aidan D. Meade
Dublin Institute of Technology, aidan.meade@dit.ie

Peter Knief
Dublin Institute of Technology, peter.knief@dit.ie

See next page for additional authors

Authors
Fiona M. Lyng, Eoghan O'Faolain, J. Conroy, Aidan D. Meade, Peter Knief, Brendan Duffy, M. Hunter, J. Byrne, P. Kelehan, and Hugh J. Byrne

Vibrational Spectroscopy for Cervical Cancer Pathology, from biochemical analysis to diagnostic tool

F.M. Lyng[1], E. Ó Faoláin[2], J. Conroy[3], A. Meade[2], P. Knief[2], B. Duffy[2], M.B. Hunter[4], J.M. Byrne[4], P. Kelehan[5] and H.J. Byrne[2]

[1]Radiation and Environmental Science Centre, Focas Institute, Dublin Institute of Technology, Kevin St, Dublin 8, IRELAND

[2]Focas Institute, Dublin Institute of Technology, Kevin St, Dublin 8, IRELAND

[3]Department of Clinical Medicine, Trinity College Dublin, Dublin 2, IRELAND

[4]Department of Histology, National Maternity Hospital, Holles St, Dublin 2, IRELAND

[5]Department of Pathology and Laboratory Medicine, National Maternity Hospital, Holles St, Dublin 2, IRELAND

Address for correspondence:

Dr. Fiona M. Lyng

Radiation and Environmental Science Centre

Focas Institute

Dublin Institute of Technology

Kevin St, Dublin 8

IRELAND

Phone +35314027972 fax +35314027904

Email: fiona.lyng@dit.ie

Abstract

Cervical cancer is the second most common cancer in women worldwide with 80% of cases arising in the developing world. The mortality associated with cervical cancer can be reduced if this disease is detected at the early stages of development or at the pre-malignant state (cervical intra-epithelial neoplasia, CIN). The aim of this study was to investigate the potential of Raman spectroscopy as a diagnostic tool to detect biochemical changes accompanying cervical cancer progression. Raman spectra were acquired from proteins, nucleic acids, lipids and carbohydrates in order to gain an insight into the biochemical composition of cells and tissues. Spectra were also obtained from histological samples of normal, CIN and invasive carcinoma tissue from 40 patients. Multivariate analysis of the spectra was carried out to develop a classification model to discriminate normal from abnormal tissue. The results show that Raman spectroscopy displays a high sensitivity to biochemical changes in tissue during disease progression resulting in an exceptional prediction accuracy when discriminating between normal cervical tissue, invasive carcinoma and cervical intra-epithelial neoplasia (CIN). Raman spectroscopy shows enormous clinical potential as a rapid non invasive diagnostic tool for cervical and other cancers.

Keywords

Vibrational spectroscopy, Raman spectroscopy, cervical cancer, diagnosis, biological macromolecules, cervical intraepithelial neoplasia (CIN), principal component analysis, linear discriminant analysis

Introduction

Cervical cancer is the second most common cancer among women worldwide and generally is more common in developing countries. However, substantial decreases in cervical cancer mortality have been observed in Western countries and some developing countries with well developed screening programmes (Parkin et al 2005). A Pap smear is used to screen for CIN and cervical cancer in the general female population. Cervical cytology can have a high specificity of 95-98% but a sensitivity of lower than 50% (Nanda et al 2000). Other methods such as automated cytology and human papilloma virus (HPV) testing (Willis et al 2005, Bolger et al 2006, Koliopoulos et al 2006) have been introduced to reduce the false negative rates. An abnormal Pap smear is followed by colposcopic examination, biopsy and histological confirmation of the clinical diagnosis. This involves the visual examination of histological sections. The grading characteristics are quite subjective and pre-malignancy may not be visually perceptible at all.

Optical technologies such as fluorescence spectroscopy (Chang et al 2005), polarised light scattering spectroscopy (Gurjar et al 2001), optical coherence tomography (Escobar et al 2004, Zuluaga et al 2005) and confocal reflectance microscopy (Carlson et al 2005) have emerged in recent years as promising tools for diagnosis of cervical and other cancers. The potential of vibrational spectroscopy for cervical cancer diagnosis has also been recognised. A number of studies have shown that Fourier Transform Infra Red (FTIR) spectroscopy (Wood et al 2003, Mordechai et al 2004, Romeo et al 2004) or Raman spectroscopy (Mahadevan-Jansen et al 1998, Utzinger et al 2001, Ó Faoláin et al 2005) can be used for detecting spectral changes in malignant and pre-malignant cells. Raman spectroscopy is, however, more suited to *in vivo* studies because of minimal interference from water.

Raman spectroscopy is an optical method based on inelastic light scattering. The sample is illuminated by monochromatic laser light and interactions between the incident photons and molecules in the sample result in scattering of the light. The exact energy required to excite a molecular vibration depends on the masses of the atoms involved in the vibration and the type of chemical bonds between these atoms and may be influenced by molecular structure, molecular interactions and the chemical microenvironment of the molecule. Therefore, the positions, relative intensities and shapes of the bands in a Raman spectrum carry detailed information about the molecular composition of the sample (Socrates 2004).

The aim of the present study was to gain an insight into the biochemical composition of cells and tissue by measuring Raman spectra from proteins, nucleic acids, lipids and carbohydrates and to show the diagnostic potential of Raman spectroscopy for cervical cancer and pre-cancer.

Materials and methods

Sample preparation

The samples examined were amino acids (all 20 amino acids), a dipeptide (arginine-lysine), proteins (albumin (bovine), β-galactosidase, and collagen (calf skin)), purines (adenine and guanine) and pyrimidines (cytosine and thymine), nucleic acid (salmon DNA), carbohydrates (glucose, glycogen), lipids (phosphatidylcholine and phosphatidylinositol) (Sigma-Aldrich, Dublin, Ireland). Spectra were recorded in a polycrystalline form on a glass slide with no preparation required.

Formalin fixed paraffin preserved (FFPP) tissue samples were obtained from both the School of Biological Sciences, DIT, Kevin Street (tonsil, kidney and prostate) and the National Maternity Hospital, Holles St., Dublin (cervical tissue). Two parallel 10μm FFPP sections were cut from each block using a microtome, mounted on glass slides and dried. Samples (with the exception of cervical tissue) were dewaxed in-house in DIT prior to investigation by immersion in baths of Xylene (BDH), Ethanol Absolut (Merck) and Industrial Methylated Spirits 95% (Lennox) and air-dried. Cervical tissue was dewaxed using the same procedure outlined above, but xylene was substituted with hexane, due to the improved level of dewaxing (Ó Faoláin et al. 2005). One section from each sample (the reference section) was stained with Haematoxylin and Eosin and the other kept unstained for spectroscopic examination.

FFPP cervical tissue sections were characterised by consultant pathologist Dr. P. Kelehan, National Maternity Hospital, Holles St., Dublin, and the samples consisted of 20 normal and 20 invasive carcinoma sections from 40 individuals. Of the 20 carcinoma samples, 10 samples were

identified as having various grades of cervical intraepithelial neoplasia (CIN), which were also marked for examination.

Spectroscopy

The Instruments S.A. Labram Raman spectroscopic confocal microscope was used, with an Argon Ion laser operating at an excitation wavelength of 514.5nm. A x50 objective lens was used. The laser power at the sample was measured and found to be 6.5±0.05mW, focused to a spot size of ≈2μm at the tissue surface. A 30 second accumulation time was used for the crystalline samples. Raman spectra of the various tissues were accumulated for 150 seconds. A total of 10 spectra were recorded from different spots on each sample.

Selected spectra were baseline corrected, dark current subtracted and were smoothed using a 10 point moving average. Where multiple spectra are displayed, spectra have been offset to facilitate clarity.

Data Analysis

All spectra (unfiltered) were subjected to spurious peak ("cosmic ray") removal and baseline correction using a common baseline, in Labspec (v. 4.02 Jobin Yvon), before being exported in ASCII format to Microsoft Excel. Spectra were normalised to the spectral maximum, from 0 to 1. Basic matrix manipulations and data reduction was carried out in Microsoft Excel Professional 2003 (v. 11.0), before being exported into Minitab to perform Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Minitab Release 14.1 Statistical Software Analysis Programme was used to produce PCA scores and LDA plots, as well as to carry out leave-one-out cross validation.

Results and Discussion

Biochemical Analysis

Amino acids, peptides and proteins

All 20 common amino acids were examined using Raman spectroscopy and found to display characteristic spectral features. Figure 1a shows the chemical structure and the Raman spectrum of alanine and phenylalanine. The Raman spectra show many narrow bands, as would be expected due to the relatively simple structure of the amino acids. Each of the Raman peaks corresponds to a vibrational energy associated with the bonds within the sample. The most distinctive peak in the spectrum of phenylalanine is the peak at 1004 cm^{-1} which can be attributed to the carbon-carbon stretching mode within the phenyl ring. The fingerprint region of the spectrum (≈ 500 - 1700 cm^{-1}) contains the majority of information and above 1700 cm^{-1} the spectrum contains mainly CH, CH_2 and CH_3 stretching modes around the 3000 cm^{-1} region (table 1). The spectra illustrate the ability of Raman spectroscopy to distinguish these molecules with their own distinctive signature spectrum.

Raman spectra of the amino acids, arginine (Arg) and lysine (Lys) were compared to the Raman spectrum of the dipeptide Arg-Lys (figure 1b). The spectrum of the dipeptide clearly derives from a combination of the two amino acids (indicated by the arrows) and interestingly contains a new peak at 1665 cm^{-1}. This is the Amide I band mainly due to C=O stretching modes of the peptide bond formed between arginine and lysine (table 1).

Three different proteins, a serum protein, an enzyme and a structural protein, were examined; albumin, β-galactosidase and collagen (figure 1c). Compared to the amino acids and peptides, it

is clear that the proteins display a more detailed Raman signature, a function of their structure and crystalline form.

Purines, pyrimidines and nucleic acids

Figure 2a displays the Raman spectra of the purines and pyrimidines present in DNA, as well as a spectrum of DNA. Again contributions from the constituent molecules are present in the spectrum of DNA, but as seen with the protein spectra the peaks broaden as the molecular complexity increases.

Carbohydrates

Figure 2b shows the Raman spectra of carbohydrates; glucose and glycogen. Glucose displays many narrow vibrational modes, similar to the spectra of single amino acids, due to its relatively simple structure. The spectrum of glycogen on the other hand, shows broader spectroscopic features, as a result of its more complex polysaccharide structure.

Lipids

Raman spectra of phosphatidylcholine and phosphatidylinositol are shown in figure 2c. The two phospholipids, which exist mainly in the extracellular membrane, can be clearly differentiated by presence of the peak seen at 714 cm^{-1}, which can be attributed to the C-N vibrational mode in the membrane phospholipid head (Stone et al 2002) present in phosphatidylcholine but absent in phosphatidylinositol.

Tissue

Raman spectroscopy has been clearly shown to identify and discriminate between molecular components of cells and tissue. A number of different tissue types were examined and some of the generalised assignments common to different tissue types are shown in figure 3. Raman bands found in these samples include the amide I band, resulting mainly from C=O stretching

vibrations of proteins, CH_2 bending and CH_2 deformation of lipids and carbohydrates, Amide III vibrations of proteins and CC stretching of carbohydrates.

Diagnostic applications

Cervical cancer

Figure 4 shows the different cell types seen in normal cervical tissue in an unstained FFPP tissue section together with the Raman spectra recorded from basal cells, epithelial cells and connective tissue. Spectra were recorded from a single sample and each spectrum represents a different spot within the sample. The spectra of the three different cell types do have a degree of similarity as seen previously for different tissue types (figure 3). However, there are also many spectral features differentiating the different cell types. The spectra of basal cells show strong bands at 724, 779 and 1578 cm^{-1} which are characteristic of nucleic acids. The same contributions were observed in the spectrum of DNA (figure 2a). The morphology of basal cells consists of a single line of tightly packed cells, with large nuclei in relation to the compacted surrounding cytoplasm. In addition, these cells are constantly dividing, providing cells to the parabasal layer. For both of these reasons a high concentration of DNA would be expected in the basal cells. Spectra of epithelial cells have characteristic glycogen bands at 482, 849, 938, 1082 and 1336 cm^{-1} as observed in the spectrum of glycogen in figure 2b. Collagen contributions can be clearly seen in the spectra of connective tissue at 850, 940 and 1245 cm^{-1} as seen in the spectrum of collagen in figure 1c.

Figure 5a compares the Raman spectra collected from normal epithelial cells and invasive carcinoma from a selection of different patients. Glycogen contributions are clearly visible in the

spectra from the normal epithelial tissue. The most obvious bands arise at 482 cm^{-1}, 849 cm^{-1} and 938 cm^{-1} and are due to glycogen skeletal deformation, CCH aromatic deformation and CCH deformation respectively. However there are also other glycogen contributions not as apparent, including a CC stretching band at 1082 and CH_3CH_2 wagging at 1336 cm^{-1}. These glycogen bands (482, 849 and 938 cm^{-1}) are absent in the spectra from invasive carcinoma, as well as a reduction in the intensity of the CC stretching mode (1082 cm^{-1}). Glycogen, a polysaccharide, is the principal storage form of glucose. Cervical cells are unusual among other epithelial cells in that they accumulate large amounts of glycogen during the maturation process (Chiriboga et al 1998). Glycogen is known to be linked with cellular maturation and disappears with the loss of differentiation during neoplasia. The spectra of invasive carcinoma also show characteristic nucleic acid bands. These include prominent bands at 724, 779 and 1578 cm^{-1}, but also at 829, 852, 1002, 1098, 1240 cm^{-1}. Distinct bands were also seen at 1366 cm^{-1}, a shoulder at 1484 cm^{-1} and a band at 1578 cm^{-1}. An increase in the intensity of the Amide I band (1655cm^{-1}) was observed in the spectra of carcinoma samples compared to the normal tissue samples. The increased nucleic acid and protein bands are a result of the increased proliferation of these tumour cells.

To investigate if pre-malignant changes could be highlighted using Raman spectroscopy, 10 areas of neoplasia (CIN) from 10 different patients were marked by a pathologist and a selection of the resulting Raman spectra are shown in Figure 5b. A number of the spectral features identified in the invasive carcinoma samples were also observed in the CIN samples such as the nucleic acid bands at 724, 779, 852, 1366 and 1578 cm^{-1}. This indicates that early biochemical changes can be identified using Raman spectroscopy.

Principal component analysis was used to reduce the number of parameters needed to represent the variance in the spectral data set. The principal components were then used to generate a linear discriminant model. PCA is a chemometric technique that can resolve a complete spectral data set into a few principal components and can thus identify and isolate important trends within the data set (Jackson 1991). LDA applies a linear discriminant function that maximises the variance in the data between groups and minimises the variance between members of the same group (Otto 1999). All three tissue classes were successfully discriminated as shown in figure 6. The classification model was tested using a leave one out cross validation in which all but one spectrum was used to build the model. This model was then used to predict the remaining spectrum. This was repeated for all 498 spectra and the results are shown in table 2. Of 498 tissue spectra, 492 were correctly classified as normal, invasive carcinoma or CIN. The cross validation mis-classified 6 spectra, 2 of which were normal samples assigned as invasive carcinoma. The other 4 were either invasive carcinoma or CIN mis-classified as either CIN or invasive carcinoma respectively. Importantly, no abnormal samples were classified as normal. Based on the cross validation results, sensitivity and specificity values were calculated as 99.5% and 100% respectively for normal tissue, 99% and 99.2% respectively for CIN and 98.5% and 99% respectively for invasive carcinoma. It is possible that these values may be slightly over optimistic because of the LDA method used with leave one out cross validation. However, similar methods have been used previously to classify bladder and prostate cancer (Crow et al 2005), breast cancer (Haka et al 2002) and basal cell carcinoma (Nijssen et al 2002). Ideally, the use of an unknown test set would provide a more robust analysis of the sensitivity and specificity of the technique and this is planned for future work.

In conclusion, the results show the ability of Raman spectroscopy to classify cervical cancer and pre-cancer with high sensitivity and specificity. These classifications are based on biochemical changes known to accompany cervical cancer such as loss of differentiation and increased proliferation. Such an automated technique measuring biochemical changes with improved diagnostic capability could allow faster, more effective patient management and inevitably would increase survival rates. Raman spectroscopy offers enormous potential as a 'next generation' technology to assist pathologists and cytologists with cervical cancer screening and diagnosis.

Acknowledgements

The authors would like to express sincere thanks to all the staff in the Histology Department in the NMH Holles St. for all their generous co-operation. The Focas Institute is funded under the Programme for Research in Third Level Institutions (PRTLI) administered by the Higher Education Authority (HEA).

References

Bolger N. , Heffron C., Regan I., Sweeney M., Kinsella S., McKeown M., Creighton G., Russell J., O'Leary J., 2006. Implementation and evaluation of a new automated interactive image analysis system Acta Cytol. 50(5), 483-91

Carlson K., Pavlova I., Collier T., Descour M., Follen M., Richards-Kortum R., 2005. Confocal microscopy: Imaging cervical precancerous lesions. Gynecologic Oncology, 99 (3) S84-88

Chang S.K., Pavlova I., Marin N., Follen M., Richards-Kortum R., 2005. Fluorescence spectroscopy as a diagnostic tool for detecting cervical pre-cancer. Gynecologic Oncology, 99 (3) S61-63

Chiriboga L., Xie P., Yee H., Vigorita V., Zarou D., Zakim D., Diem M., 1998. Infrared spectroscopy of human tissue. I. Differentiation and maturation of epithelial cells in the human cervix. Biospectroscopy 4(1):47-53

Crow P., Molckovsky A., Stone N., Uff J., Wilson B., WongKeeSong L.M., 2005. Assessment of fiberoptic near-infrared raman spectroscopy for diagnosis of bladder and prostate cancer. Urology. 65(6), 1126-30.

Escobar P.F., Belinson J.L., White A., Shakhova N.M., Feldchtein F.I., Kareta M.V., Gladkova N.D., 2004. Diagnostic efficacy of optical coherence tomography in the management of preinvasive and invasive cancer of uterine cervix and vulva. Int J Gynecol Cancer. 14(3), 470-4

Gurjar R.S., Backman V., Perelman L.T., Georgakoudi I., Badizadegan K., Itzkan I., Dasari R.R., Feld M.S., 2001. Imaging human epithelial properties with polarized light-scattering spectroscopy, Nat Med. 7(11), 1245-8

Haka A.S., Shafer-Peltier K.E., Fitzmaurice M., Crowe J., Dasari R.R., Feld M.S., 2002. Identifying microcalcifications in benign and malignant breast lesions by probing differences in their chemical composition using Raman spectroscopy. Cancer Res. 62, 5375-5380

Jackson J.E., 1991. A User's Guide to Principal Components. New York, Wiley.

Koliopoulos G., Arbyn M., Martin-Hirsch P., Kyrgiou M., Prendiville W., Paraskevaidis E., 2006. Diagnostic accuracy of human papillomavirus testing in primary cervcal screening: A review and meta-analysis of non-randomised studies. Gynecologic Oncology, doi:10.1016/j.ygyno.2006.08.053

Mahadevan-Jansen A., Ramanujam N., Malpica A., Thomsen S., Utzinger U., Richards-Kortum R., 1998. Near-Infrared Raman Spectroscopy for In Vitro Detection of Cervical Precancers. Photochemistry and Photobiology, 68(1), 123-132.

Mordechai S., Sahu R.K., Hammody Z., Mark S., Kantarovich K., Guterman H., Podshyvalov A., Goldstein J., Argov S., 2004. Possible common biomarkers from FTIR microspectroscopy of cervical cancer and melanoma. Journal of Microscopy-Oxford, 215, 86-91.

Nanda K., McCrory D.C., Myers E.R., Bastian L.A., Hasselblad V., Hickey J.D. Matchar D.B. 2000. Accuracy of the Papanicolaou test in screening for and follow up of cervical cytologic abnormalities: a systematic review. Ann Intern Med, 132, (10), 810-819

Nijssen A., Bakker Schut T.C., Heule F., Caspers P.J., Hayes D.P., Neumann M.H., Puppels G.J., 2002. Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy. J Invest Dermatol. 119, 64-9.

Ó Faoláin E., Hunter M.B., Byrne J.M., Kelehan P., Byrne H.J., Lyng F.M., 2005. The potential of vibrational spectroscopy in the early detection of cervical cancer: an exciting emerging field. Proc. SPIE, Vol. 5826, 25-36.

Ó Faoláin E., Hunter M.B., Byrne J.M., Kelehan P., Lambkin H.A., Byrne H.J., Lyng F.M., 2005. Raman spectroscopic evaluation of efficacy of current paraffin wax section dewaxing agents. J. Histochem. Cytochem., 53, 121-131.

Otto M., 1999. Chemometrics: Statistics and Computer Applications in Analytical Chemistry. New York, Wiley.

Parkin D.M., Bray F., Ferlay J. and Pisani P., 2005. Global Cancer Statistics 2002, CA Cancer J Clin; 55:74-108

Romeo M., Matthaus C., Miljkovic M., Diem M., 2004. Infrared microspectroscopy of individual human cervical cancer (HeLa) cells. Biopolymers, 74 (1-2), 168-171.

Socrates G., 2004 “Infrared and Raman Characteristic Group Frequencies : Tables and Charts - 3 edition” John Wiley & Sons.

Stone N., Kendall C., Shepard N., Crow P., Barr H., 2002. Near-infrared Raman spectroscopy for the classification of epithelial and pre-cancers and cancer. Journal of Raman Spectroscopy, (33), 564-573.

Utzinger U., Mahadevan-Jansen A., Hinzelman D., Follen M., Richards-Kortum R., 2001. Near Infrared Raman Spectroscopy for In Vivo Detection of Cervical Precancers. Applied Spectroscopy, 55 (8), 955.

Willis B.H., Barton P., Pearmain P., Bryan S., Hyde C., 2005. Cervical screening programmes: can automation help? Evidence from systematic reviews, an economic analysis and a simulation modelling exercise applied to the UK. Health Technol Assess, 9 (13), 1-207

Wood B.R., Chiriboga L., Yee H., Quinn M.A., McNaughton D., Diem M., 2003. Fourier transform infrared (FTIR) spectral mapping of the cervical transformation zone and dysplastic squamous epithelium. Gynecologic Oncology, 93 (1), 59-68

Zuluaga AF, Follen M, Boiko I, Malpica A, Richards-Kortum R., 2005. Optical coherence tomography: a pilot study of a new imaging technique for noninvasive examination of cervical tissue Am J Obstet Gynecol. 193(1), 83-88

Legends

Figure 1 Raman spectra of (a) alanine (A) and phenylalanine (B), (b) arginine (Arg) (A), lysine (Lys) (B) and dipeptide Arg-Lys (C) and (c) three proteins in ascending chain length; albumin (A), β-galactosidase (B) and collagen (C). Assignments of the main Raman vibrational modes are detailed in table 1.

Figure 2 Raman spectra of a) cytosine (A), thymine (B) (pyrimidines), adenine (C), guanine (D) (purines) and DNA (E), b) carbohydrates, glucose (A), glycogen (B) and (c) phospholipids, phosphatidylinositol (A) and phosphatidylcholine (B). Raman peak assignments are detailed in table 1.

Figure 3 Raman spectral features of various tissues, A) Tonsil B) Kidney and C) Prostate, with generalised assignments, where **ν** = stretch and **δ** = deformation.

Figure 4 (a) Photomicrograph of unstained tissue section, with different cell types identified and (b) Raman spectra recorded from basal cells (A), epithelial cells (B) and connective tissue (C) in cervical tissue sections. The main spectral features associated with each cell type are highlighted.

Figure 5 Raman spectra of a) normal cervical epithelial cells (A) and invasive carcinoma cells (B) and (b) Raman spectra of CIN tissue. Assignments of the main Raman vibrational modes are detailed in table 1.

Figure 6 Linear discriminant analysis of the principal components of the first derivative, normalised, 10 point averaged spectra, over the entire spectral range, **C** = CIN, **N** = normal and **T** = invasive carcinoma

Table 1

Table 1 Peak position and assignments of main Raman vibrational modes

Peak Position (cm^{-1})	**Assignment**
622	C-C twisting
724	CH_2 deformation
746	CH_2 rocking
754	Symmetric ring breathing
779	Ring vibration
832	CCH deformation aliphatic
853	CCH deformation aromatic
873	CC stretch
922	C-C stretching
1004	CC aromatic ring breathing
1034	C-C stretching
1065	C-N stretch
1096	C-C chain stretching
1098	CC stretch
1102	CC stretch
1124	CC skeletal stretch *trans*
1214	CC stretch backbone carbon phenyl ring
1236	CN stretch, NH bending Amide III band
1240	CN stretch, NH bending Amide III band
1314	CH deformation
1337	CH_2 deformation
1335	CH_2 deformation
1366	CH_2 bending
1440	CH_2 scissoring
1484	CH_2 deformation
1548	NH deformation; CN stretch Amide II band
1578	C=C olefinic stretch
1585	C=C stretching
1602	CO stretching
1660 - 1665	C=O stretch Amide I α-helix
2930	CH_2 stretching (2930 cm^{-1})
2932	CH3 symmetric stretch (2932 cm^{-1})

Table 2

Table 2 Summary of cross-validation classification results and sensitivity and specificity values

	Diagnosis (Histopathology)		
Predicted (Raman)	*Normal*	*CIN*	*Invasive carcinoma*
Normal	198	0	0
CIN	0	99	3
Invasive carcinoma	2	1	195
Total	200	100	198
Number correct	198	99	195
Proportion	0.990	0.990	0.985
Sensitivity	99.5%	99.0%	98.5%
Specificity	100%	99.2%	99.0%

Figure 1

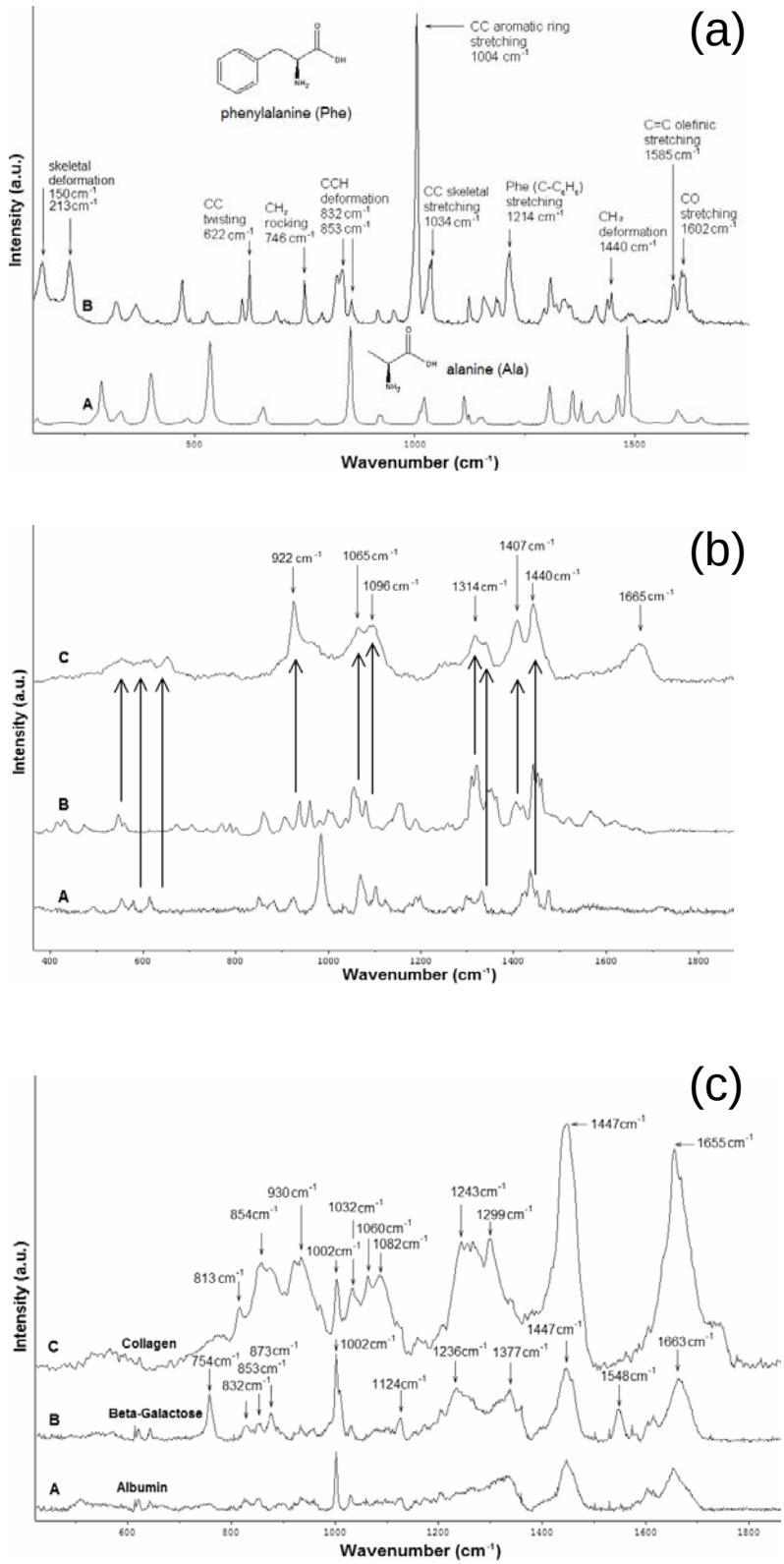

(a)
phenylalanine (Phe)
CC aromatic ring stretching 1004 cm-1
skeletal deformation 150cm-1 213cm-1
CC twisting 622 cm-1
CH2 rocking 746 cm-1
CCH deformation 832 cm-1 853 cm-1
CC skeletal stretching 1034 cm-1
Phe (C-C6H5) stretching 1214 cm-1
CH2 deformation 1440 cm-1
C=C olefinic stretching 1585 cm-1
CO stretching 1602 cm-1
alanine (Ala)
B
A
Intensity (a.u.)
Wavenumber (cm-1)
(b)
922 cm-1
1065 cm-1
1096 cm-1
1314 cm-1
1407 cm-1
1440 cm-1
1665 cm-1
C
B
A
Intensity (a.u.)
Wavenumber (cm-1)
(c)
1447cm-1
1655 cm-1
930 cm-1
854cm-1
1032cm-1
1060cm-1
1082cm-1
1002cm-1
1243cm-1
1299 cm-1
813 cm-1
1447cm-1
Collagen
754cm-1
873cm-1
853cm-1
1002cm-1
1236cm-1
1377cm-1
1663cm-1
1548cm-1
832cm-1
1124cm-1
Beta-Galactose
Albumin
C
B
A
Intensity (a.u.)
Wavenumber (cm-1)

Figure 2

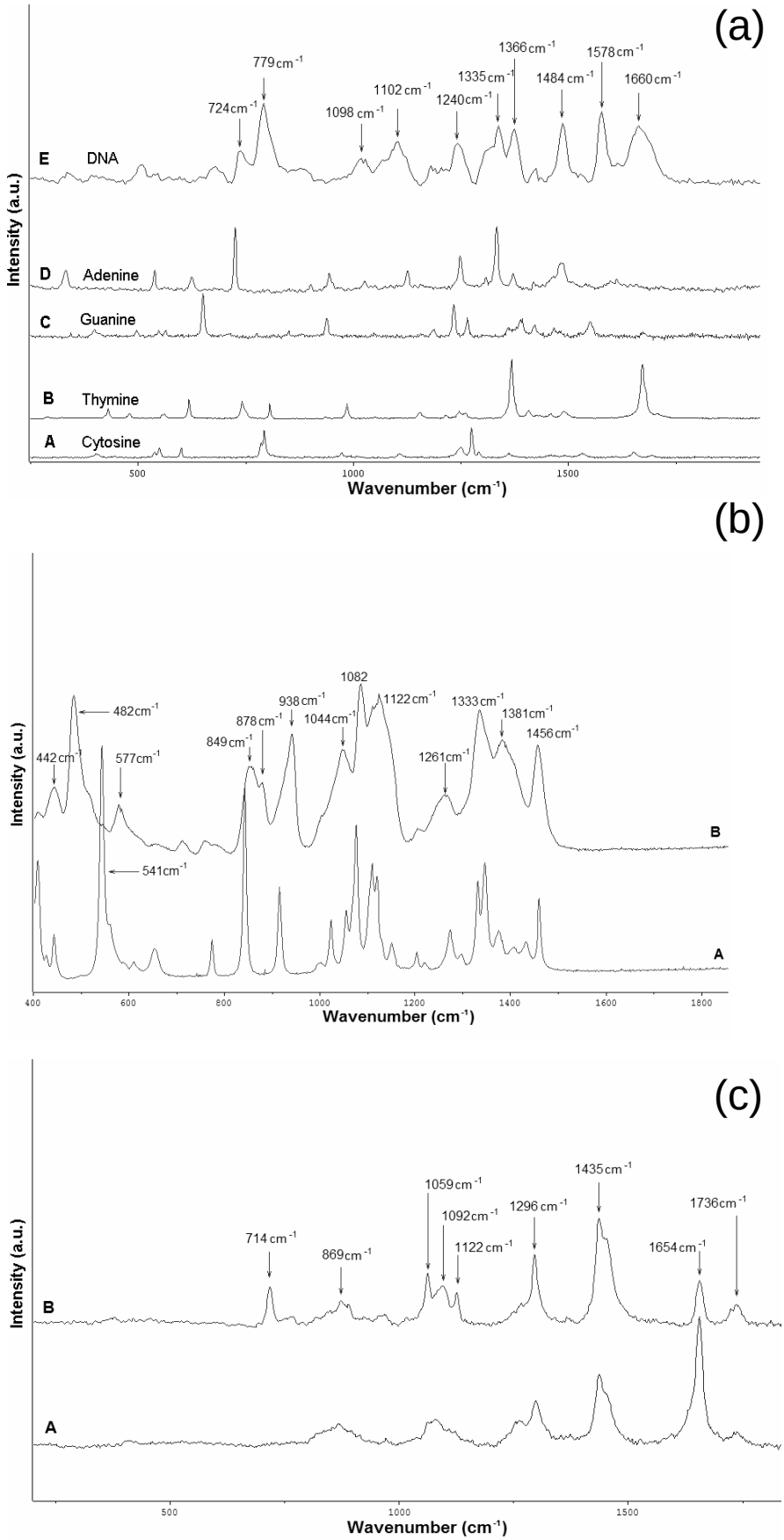

(a)
E DNA
D Adenine
C Guanine
B Thymine
A Cytosine
724 cm-1
779 cm-1
1098 cm-1
1102 cm-1
1240 cm-1
1335 cm-1
1366 cm-1
1484 cm-1
1578 cm-1
1660 cm-1
Intensity (a.u.)
Wavenumber (cm-1)
(b)
442 cm-1
482 cm-1
577 cm-1
541 cm-1
849 cm-1
878 cm-1
938 cm-1
1044 cm-1
1082
1122 cm-1
1261 cm-1
1333 cm-1
1381 cm-1
1456 cm-1
B
A
Intensity (a.u.)
Wavenumber (cm-1)
(c)
714 cm-1
869 cm-1
1059 cm-1
1092 cm-1
1122 cm-1
1296 cm-1
1435 cm-1
1654 cm-1
1736 cm-1
B
A
Intensity (a.u.)
Wavenumber (cm-1)

Figure 3

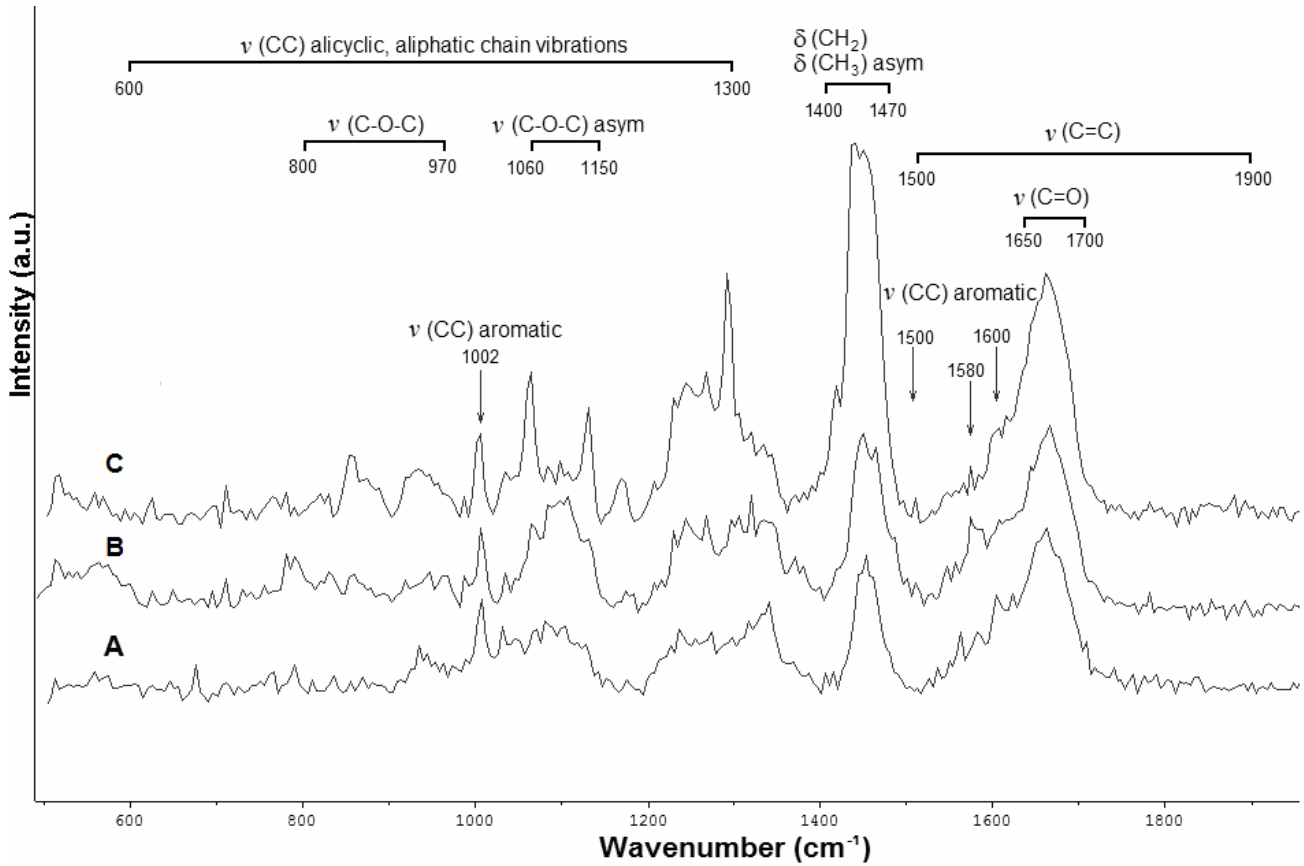
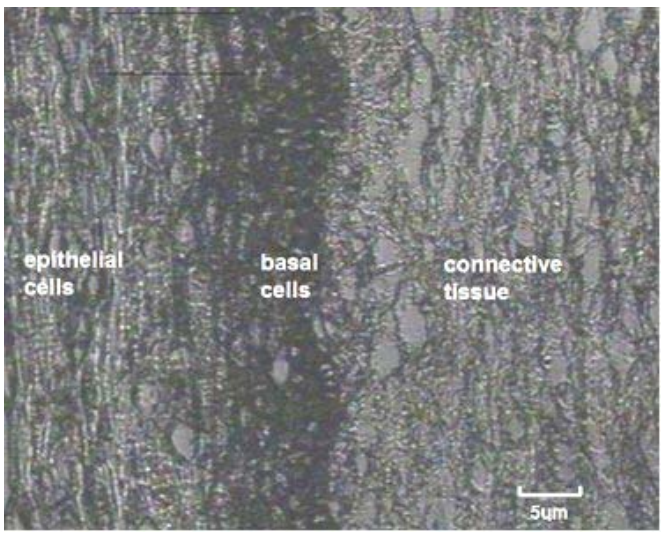

ν (CC) alicyclic, aliphatic chain vibrations
600
1300
δ (CH2)
δ (CH3) asym
1400
1470
ν (C-O-C)
800
970
ν (C-O-C) asym
1060
1150
ν (C=C)
1500
1900
ν (C=O)
1650
1700
ν (CC) aromatic
1002
ν (CC) aromatic
1500
1580
1600
C
B
A
Intensity (a.u.)
600
800
1000
1200
1400
1600
1800
Wavenumber (cm-1)

Figure 4


epithelial cells
basal cells
connective tissue
5um

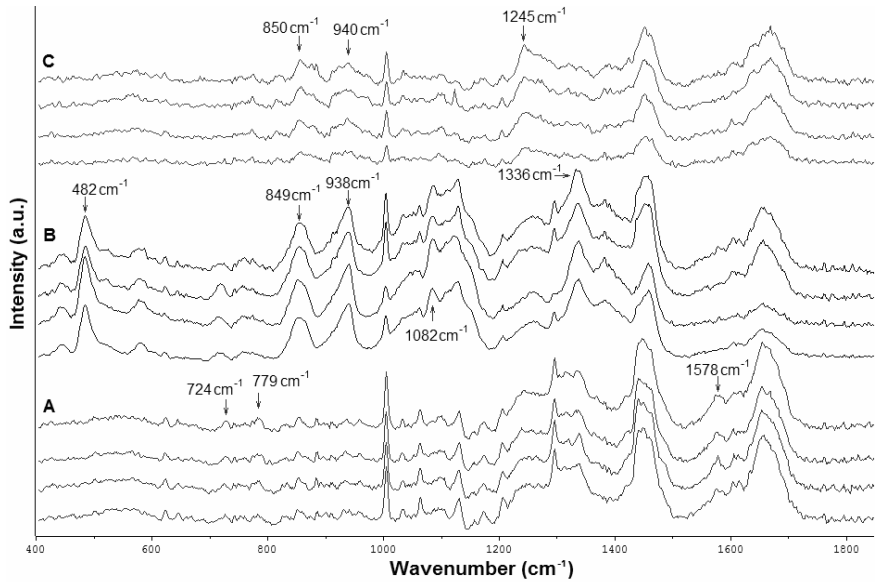
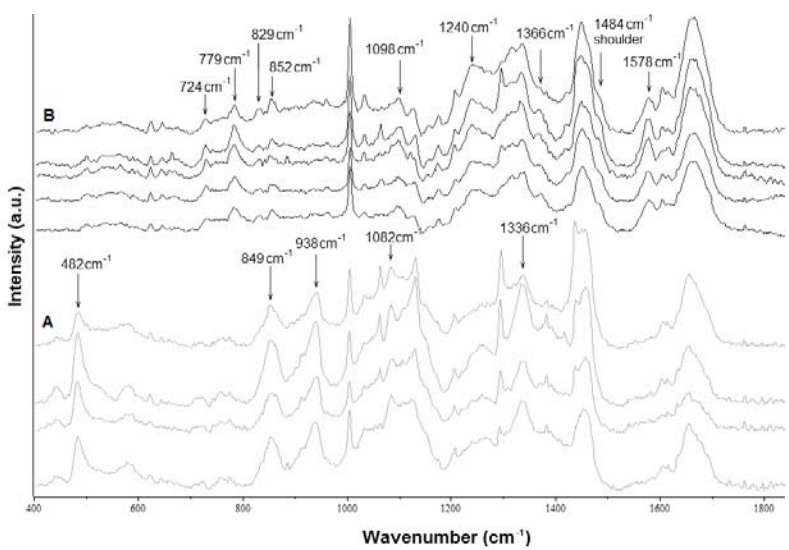

C
B
A
850 cm-1
940 cm-1
1245 cm-1
482 cm-1
849 cm-1
938 cm-1
1336 cm-1
1082 cm-1
724 cm-1
779 cm-1
1578 cm-1
Intensity (a.u.)
400
600
800
1000
1200
1400
1600
1800
Wavenumber (cm-1)

Figure 5


724 cm-1
779 cm-1
829 cm-1
852 cm-1
1098 cm-1
1240 cm-1
1366 cm-1
1484 cm-1
shoulder
1578 cm-1
B
462 cm-1
849 cm-1
938 cm-1
1082 cm-1
1336 cm-1
A
Intensity (a.u.)
400
600
800
1000
1200
1400
1600
1800
Wavenumber (cm-1)

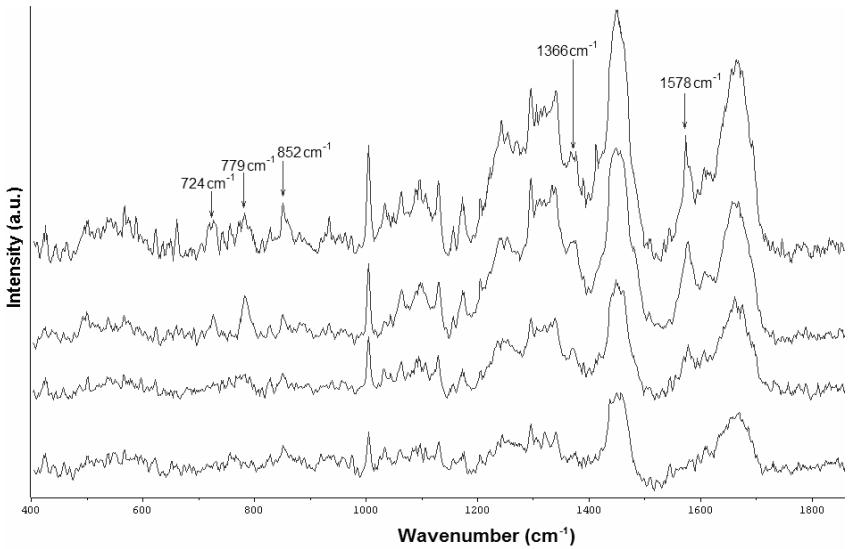
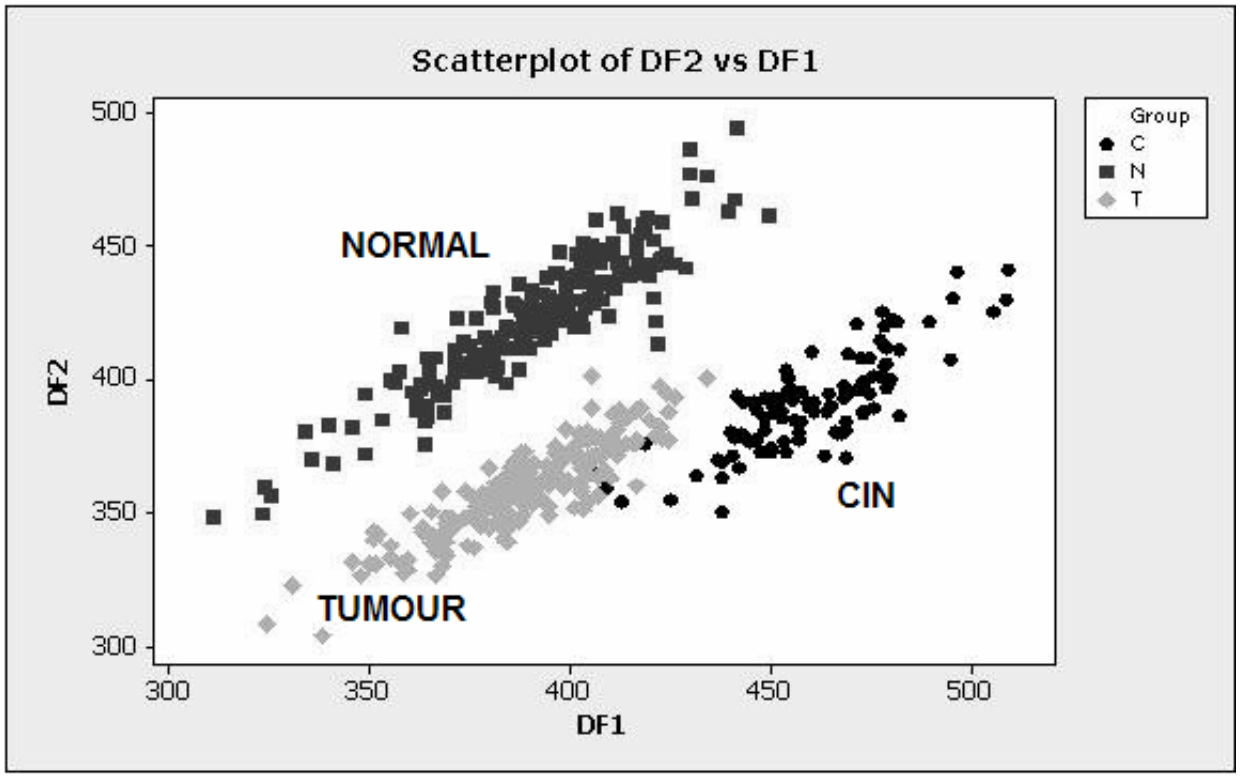

724 cm-1
779 cm-1
852 cm-1
1366 cm-1
1578 cm-1
Intensity (a.u.)
400
600
800
1000
1200
1400
1600
1800
Wavenumber (cm-1)

Figure 6


Scatterplot of DF2 vs DF1
Group
C
N
T
NORMAL
CIN
TUMOUR
DF2
DF1
500
450
400
350
300
300
350
400
450
500