2019

Raman Spectroscopy of Blood Plasma Samples from Breast Cancer Patients at Different Stages.

H. F. Nargis
University of Agriculture, Faisalabad, Pakistan

H. Nawaz
University of Agriculture, Faisalabad, Pakistan

A. Ditta
University of Agriculture, Faisalabad, Pakistan

See next page for additional authors

Recommended Citation
Nargis, H.F. et al (2019). Raman spectroscopy of blood plasma samples from breast cancer patients at different stages.*Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 222, 117210. doi:10.1016/j.saa.2019.117210

Title: Raman spectroscopy of blood plasma samples frombreast cancer patients at different stages.

Authors:

H. F. Nargis[a], H. Nawaz[a]*, A. Ditta[a], T. Mahmood[a], M. I. Majeed[a], N. Rashid[b], H. N. Bhatti[a], M. Saleem[c], K. Jilani[d], F. Bonnier[e]and H.J. Byrne[f]

[a]Department of Chemistry, University of Agriculture, Faisalabad, Pakistan.

[b]University of Central Punjab, Faisalabad campus, Faisalabad, Pakistan.

[c]National Institute of Lasers and Optronics (NILOP), Islamabad, Pakistan.

[d]Department of Biochemistry, University of Agriculture, Faisalabad, Pakistan.

[e] EA 6295 Nano-médicaments and nano-sondes, Université de Tours, Tours, France.

[f] FOCAS Research Institute, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland.

*Corresponding Author: DrHaq Nawaz

E-mail: haqchemist@yahoo.com

Abstract

Raman spectroscopy was employed for the characterisation of blood plasma samples from patients at different stages of breast cancer. Blood plasma samples taken from clinically diagnosed breast cancer patients were compared with healthy controls using multivariate data analysis techniques (principal components analysis – PCA) to establishRaman spectral features which can be considered spectral markers of breast cancer development.All the stages of the disease can be differentiated from normal samples. It is also found that stage 2 and 3 are biochemically similar, but can be differentiated from each other by PCA. The Raman spectral data of the stage 4 is found to be biochemically distinct, but very variable between patients.Raman spectral features associated with DNA and proteins were identified, which are exclusive to patient plasma samples. Moreover, there are several other spectral features which are strikingly different in the blood plasma samples of different stages of breast cancer. In order to further explore the potential of Raman spectroscopy as the basis of a minimally invasive screening technique for breast cancer diagnosis and staging, PCA-Factorial Discriminant Analysis (FDA) was employed to classify the Raman spectral datasetsof the blood plasma samples of the breast cancer patients, according to different stages of the disease, yielding promisingly high values of sensitivity and specificity for all stages.

Keywords: Breast cancer, blood plasma, Raman spectroscopy, Multivariate data analysis, cancer staging,

Introduction

Breast cancer is a major threat to the health of women, primarily, all over the world. According to the International Agency for Research on Cancer(IARC), of all diagnosed cancers, about 11.9% (1.7 million cases)are of the breast[1]. Its incidence is spreading globally, except in the United States, where it is comparatively stable [2], and, in 2017, and estimated 250,000 diagnosed cases of female breast cancer and 40,000 associated deaths were reported[3].According to the U.S. National Cancer Institute, the associated cost of the care of breast cancer will reach approximately $20.5 billion in 2020 [4]. However, the survival is more favorable, although the mortality rateis lower in developed countries than developing regions [5]. The contribution of genetic factors in breast cancer is reported to be only 5-10% [6-7], posing a challenge for potential prevention. The variable risk factors may include late childbearing, lactation failure, weight gain after menopause, nulliparity, lack of exercise, and alcohol consumption [8]. Breast cancer cells develop in the tissue of the mammary gland and have an evolution period of 7 years, as 100-300 days are required to double the number of cancer cells and the emergence of breast neoplasm requires 30 doublings[9]. This highlights the potential importance of early detection and the developments of screening techniques which are capable of detecting the early stage abnormalities at a micro level. The different stages of cancer are an important element ofthediagnosis protocol, particularly for the survival prospects. The classification depends on the size of tumour (T), involvement of lymph nodes (N) and metastasis/no metastasis to other organs (M) [5]. In stage 0, caner has not developed further than the instigation point, while in stage 1, it has spread to some fatty breast tissues, it has become a bit bigger in size at stage 2 and has spread up to 3 lymph nodes, while in stage 3, it has spread to the chest wall. Critically at stage 4, also referred to as the metastatic phase, it has spread from the breast or lymph nodes to bones, liver, lungs and brains, beyond which there is no further stage classification.

In the field of medicine/ clinical practice, the range of techniques which are currently used for the diagnosis and subsequent treatment of breast cancer include fluorescence, optical bioluminescence, optical imaging, MRI (magnetic resonance imaging), X-ray imaging and ultrasound [10-12]. These techniques have low sensitivity and resolution, and X-ray employs potentially harmful radiation which can be very risky for patients. Moreover, when visible on x-

ray or MRI, the cancer is already at an advanced stage. There is a clinical need for a high sensitivity, real time, label free and less invasive method to screen patients for early stage symptoms, and also to establish the stage of development of the disease.

Raman spectroscopy has been demonstrated to provide specific molecular signatures of a variety of biological samples and other materials [13-15]. It can detect changes of molecular composition and structure which take place during the formation of a tumour, in the constituents of cells and tissue such as nucleic acids, carbohydrates, proteins and lipids. Blood carries the metabolic products of a tumour which can be monitored during circulation. As these biochemical changes are manifest prior to the appearance of clinical symptoms commonly detected by medical imaging, Raman spectroscopy has the potential to provide early diagnosis of the cancer [16-22].

Regarding number of studies have already explored Raman spectroscopy for the diagnosis of breast cancer based on blood plasma/serum samples from clinically diagnosed patients. Bilal et al. applied Partial least squares(PLS) regression analysis to the Raman spectral data acquired from blood serum of breast cancer and healthy individuals and coefficients of regression were determined which are associated with the changes at the bio-molecular level, hence indicating the development of the disease[23]. In another study, Pichardo-Molina et al. applied Raman spectroscopy along with principal components analysis (PCA) and PCA-Linear Discriminate Analysis (LDA) to breast cancer and control blood samples, identifying specific bands which distinguish breast cancer blood serum samples from the healthy ones[24]. To date, however, to the best of our knowledge, there have been no reports in which the technique has been applied to differentiate the different stages of breast cancer by using blood plasma samples.

In this study, Raman micro spectroscopy, supported by multivariate analyses, is employed to analyse the characteristic signatures of blood plasma samples from patients diagnosed with breast cancer of either stage 2, 3 or 4, and to differentiate them from both normal healthy controls and from each other. In doing so, the underlying biochemical origin of the differentiating features is explored, and the potentialof the technique for minimally invasive diagnosis and staging of the disease is demonstrated.

Materials and methods

Sample preparation and protocol

Two groups were selected for study, which included 18 female patients with a confirmed clinical and histopathological diagnosis of breast cancer and 8 healthy/control subjects. Notably, the available breast cancer samples belong to stages 2, 3 and 4 with six, eight and four blood plasma samples respectively.Samples of stage-1 were not available, as patients rarely present to the clinic at this stage of the disease. Blood plasma samples were collected from breast cancer patients that were admitted to two hospitals; Allied and PINUM Hospital, Faisalabad and were approved for study by the Ethical Committee. Patients from both hospitals had similar socioeconomic and ethnic backgrounds. Blood samples were collected in Ethylene Diamine Tetra Acetic Acid (EDTA) vials and centrifuged at 3400 rpm for 5 min to yield the plasma samples.

Raman spectral acquisition

A 20 µl drop of each plasma sample was placed on an aluminium slide at room temperature and Raman spectra were acquired before the plasma sample had dried.The process of placing the 20µl drop on the aluminium slide was repeated 3 times, employing a clean aluminium slide each time, acquiring 20-25 spectra in total from each blood plasma sample.

Raman spectral acquisition from all of the 19 female patients with a confirmed diagnosis of breast cancer and 8 healthy/control blood plasma samples was performed using a Raman micro spectrometer(Peak Seeker Pro-785; Agiltron, USA). The Peak Seeker Pro-785 utilises a 785nmdiode laser as the source, deliveringa laser power of ~60 mW at the sample. The laser was delivered tothe sample through a 10× objective. The acquisition of the Raman spectra for all thesamples was executed from 600 to 1800 $cm^{-1}$and 20–25 Ramanspectra per sample were acquired with an acquisition time of 30 second.

The assignments of the Raman spectralfeatures used in interpretation of the results were taken from theliterature[25-29] and are described in **Table 1**.

Data preprocessing

All data processing of the Raman spectra was performed using Matlab 7.2 and established protocols [30]. Data pre-processing included smoothing, baseline correction, vector normalisation and substrate removal. All spectra, including substrate backgrounds, were vector normalised and smoothed using a Savitzky-Golay smoothing method (order 5,13 point window). A rubber band

correction for baseline removal for all the spectra was carried out and the substrate spectra were subtracted from each spectrum.

Data analysis

Principal components analysis (PCA) is a mathematical procedure involvingthe transformation of possibly correlated variables into a smallernumber of uncorrelated variables, known as principal components(PC), basically to reduce the dimensionality of the data whilst maintainingtheir variability. The first principal component accounts forthe dominant source of variability in the data, and each successive principal component accounts for the next highest source of the remaining variability. PCA-Factorial Discriminant Analysis (FDA) has been employed to differentiate the Raman spectra based on the stage of the breast cancer, demonstrating the diagnostic potential of the technique. The discriminant analyses help to classify the unknown samples using prior knowledge of the category from training sets. To predict the association of the spectral data to a defined group (stage of the disease), the coupling of PCA with FDA makes use of the calculated PC scores to better evaluate the rate of discrimination acquired based on a data set [31]. The classification values obtained for each iterationare gathered to producea confusion matrix which is representative of the discrimination of the data. For this type of strategy, the model is restricted to a specified number of iterations, dictated by the number of patients. In present study, there are 27 groups of Raman spectra (8 healthy and patients with a confirmed clinical and histopathological diagnosis of breast cancer).

Wavenumbers (cm^{-1})	**Peak assignments**	**References**
625	Nucleotide conformation	[32]
689	Nucleotide conformation	[33]
700	n (C-S) trans (aminoacid methionine)	[34]
714-6	C-N (membrane phospholipids head)/adenine $CN_2(CH_3)_3$ (lipids)	[35]

761	Tryptophan, d (ring)	[34]
770	Phosphate	[36]
778	Phosphatidylinositol	[37]
788	phosphodiester bands in DNA	[38]
798	CH out of plane deformation	[39]
828	tyrosine/protein	[40]
848	Most probably due to single bond stretching vibrations for the amino acids and valine. Tyrosine (Fermi resonance of ring fundamental and overtone)	[41] [34]
857	(collagen type I)	[40]
885	Disaccharide (cellobiose), (C-O-C) skeletal mode	[34]
911	Glucose	[37]
917	Proline, hydroxyproline	[40]
1100	C-C vibration mode of the gauche-bonded chain	[42]
1138	n(C-C)-lipids, fatty acids	[34]
1145	n(C-C)-lipids, fatty acids	[34]
1173	Cytosine, guanine	[43]
1185	Anti-symmetric phosphate vibrations	[44]
1268	d (=C-H) (phospholipids)	[45]
1285	Typical phospholipids	[46]
1307	CH_3/CH_2twisting, wagging &/or bending mode of collagens & lipids	[40]
1319	Guanine (B,Z-marker)	[43]

1410	ns COO_2 (IgG)	[45]
1440	CH_2 deformation in normal breast tissue	[47]
1609	Cytosine (NH_2)	[43]
1660-63	DNA	[48]

Table 1: Assignments of the Raman spectral features, derived from literature.

Results and discussion

Figure 1 shows the mean Raman spectra with standard deviation of control and breast cancer blood plasma samples (mean Raman spectra of breast cancer stage-2, stage-3 and stage-4), and potentially differentiating features are indicated by vertical lines (Spectra have been offset for clarity). Raman spectral features, including those at 689 (nucleotide conformation), 770 (phosphate), 788 (phosphodiester bands in DNA),828 (tyrosine/protein), 848 (single bond stretching vibrations for the amino acids and valine and polysaccharides), 885(disaccharide (cellobiose), (C-O-C) skeletal mode), 1138 (n(C-C)-lipids, fatty acids),1173(cytosine, guanine), 1185 (anti-symmetric phosphate vibrations), 1268 (d (=C-H) (phospholipids)), 1285 (typical phospholipids), 1307 (CH_3/CH_2 twisting, wagging &/or bending modeof collagens & lipids) and 1319cm^{-1} (guanine (B,Z-marker)) have higher intensities in the mean Raman spectra of patient samples. In contrast, the Raman spectral features at 700 (n (C-S) trans (aminoacid methionine)), 761 (Tryptophan, d (ring)), 798 (CH out of plane deformation) and 1410 cm^{-1} (ns COO_2) have higher intensities in the mean Raman spectra of control/healthy volunteers than in the patient samples.

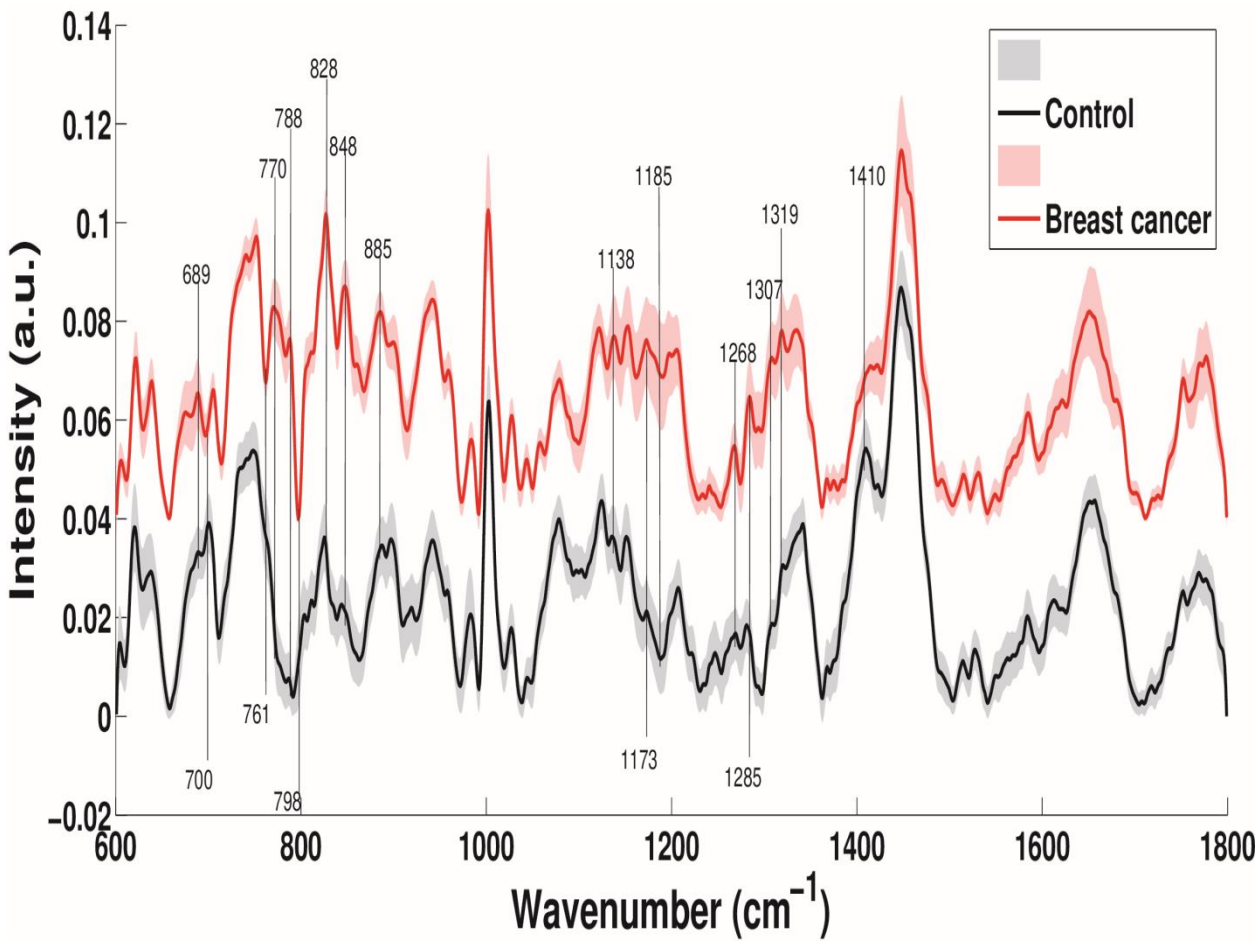

Figure 1: Mean Raman spectra of control/healthy (black) and breast cancer (red) blood plasma samples. Spectra have been offset for clarity.

Figure 2shows the PCA scatter plot of the Raman spectral data of the blood plasma samples of breast cancer patients of all different stages versus healthy ones. PC1 explains 79% of the variance, and shows reasonably good differentiation of the Raman spectral data of healthy and breast cancer patients. Stage 2 and Stage 3 samples are well differentiated from normal controls, but are largely overlapped, although it may be considered that the Stage 3 samples are more differentiated according to PC1. The overlap of the Stage 2 and 3 clusters indicates that, although they are clinically distinguished by the extent of the disease progression, they are spectrally, and therefore biochemically similar, in terms of their manifestation in the blood stream. Notably, the outlying clusters of the spectra have been identifiedas associated with the samples of stage-4,which are clustered according to individual patients. They are not significantly more differentiatedthan the other stages according to PC1, but are further differentiated according to PC2, indicating a distinct biochemical manifestation in the blood stream.

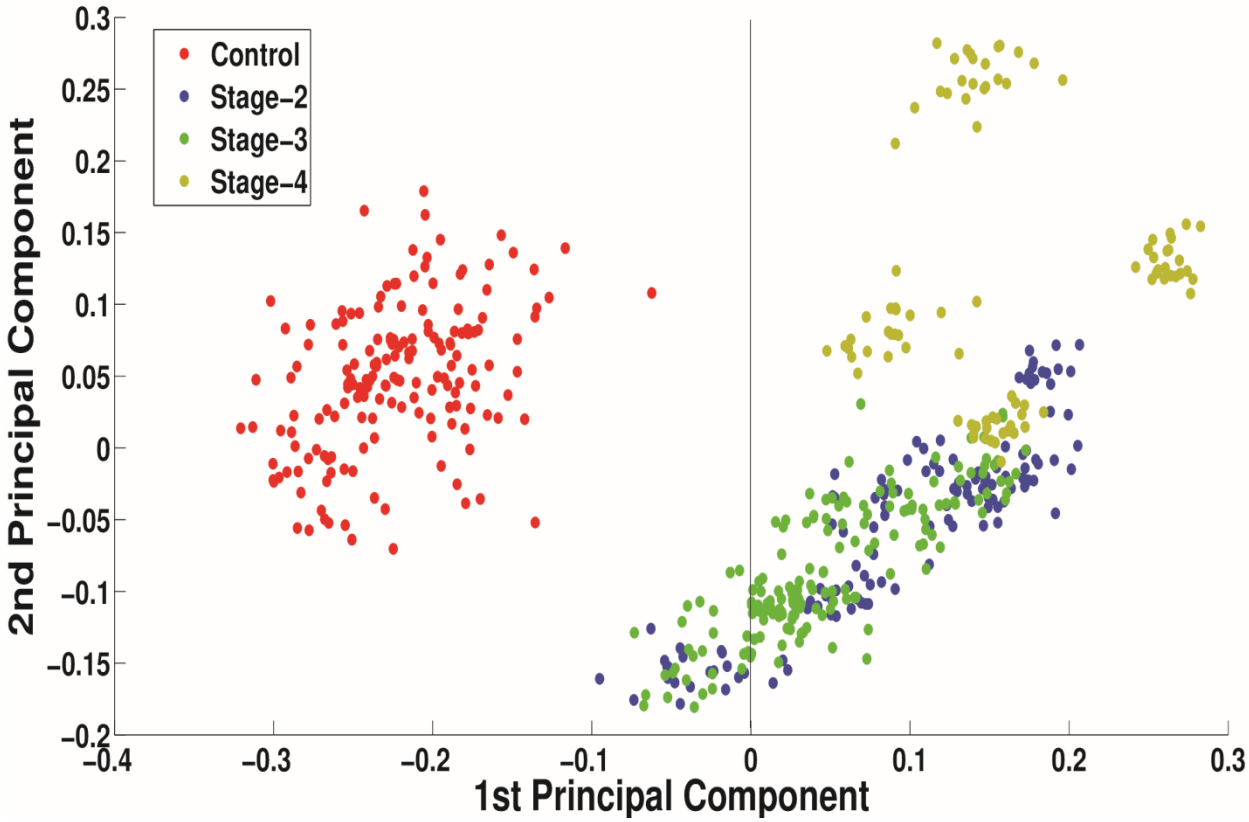

Figure 2: PCA scatter plot of Raman spectral data of healthy versus breast cancer samples.

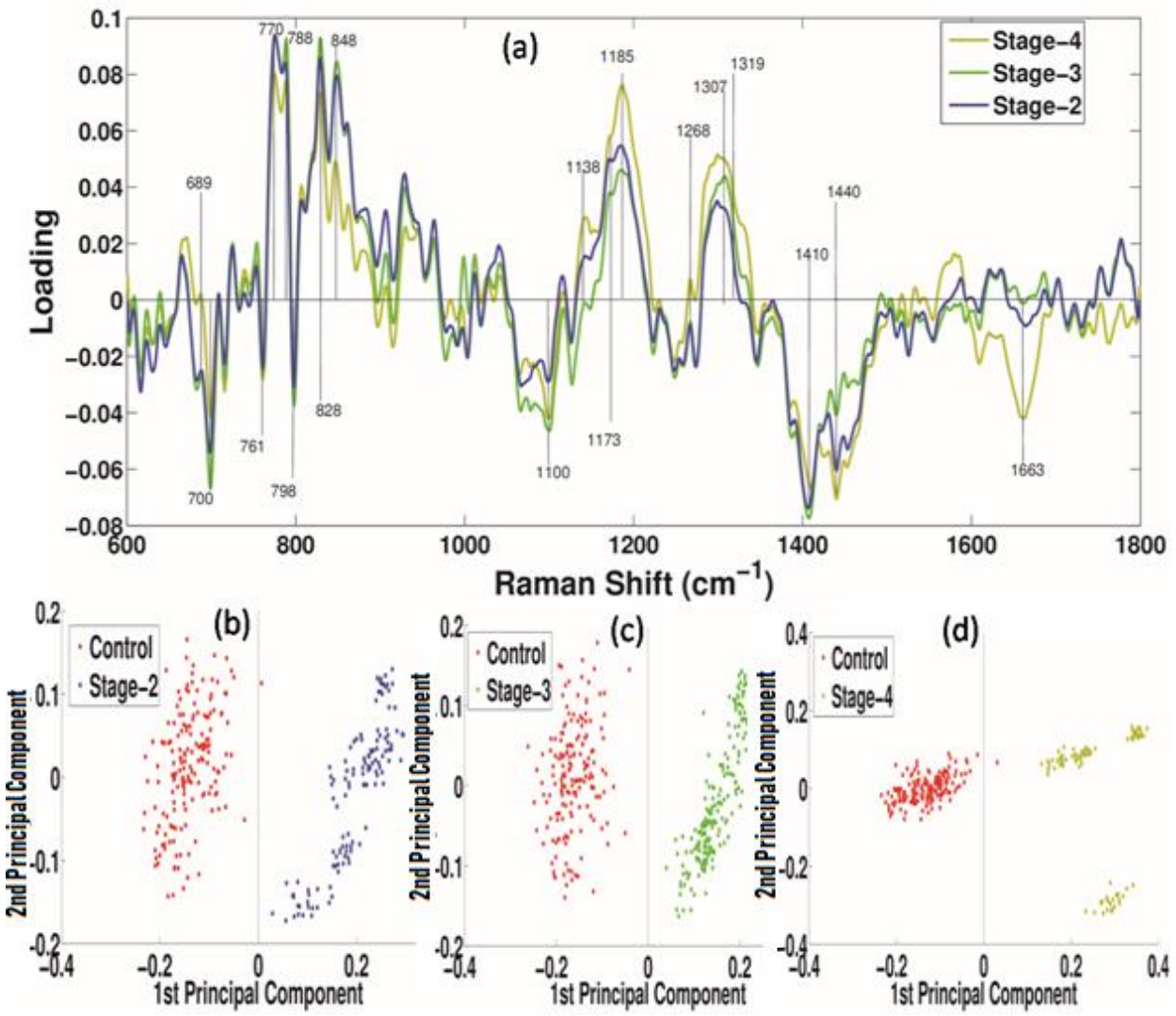

Figure 3: PCA loadings of Raman spectral data of healthy versus different stages of breast cancer samples (a) PCA scatter plots of; control versus stage-2 (b), control versus stage-3 (c) and control versus stage-4 (d).

Figure 3 represents the PCA scatter plots and loadings of the Raman spectral data of control/healthy versus each different stage of the breast cancer blood samples. The changes in biological molecules which are responsible for the clustering of healthy and cancerous samples in the PCA scatter plots are indicated by the loadings of the PC in Figure 3 (a).In all cases, positive loadings show the peaks of cancerous samples, while negative loadings represent the Raman spectral features associated with the healthy samples, as their respective Raman spectra are clustered in the positive and negative axis of the PC-1. The Raman spectral features, observed in

positive loadings of all stages, include those at689 (Nucleotide conformation), 770 (phosphate), 788 (phosphodiester bands in DNA),828 (tyrosine/protein), 848 (single bond stretching vibrations for amino acids and polysaccharides), 1138 (n(C-C)-lipids, fatty acids),1173(cytosine, guanine), 1185 (anti-symmetric phosphate vibrations), 1268 (d (=C-H) (phospholipids)), 1285 (typical phospholipids) 1307 (CH_3/CH_2 twisting, wagging &/or bending modeof collagens & lipids) and 1319 cm^{-1} (guanine (B,Z-marker)).Notably, these Raman featureswere also observed in figure 1 as elevated in the mean Raman spectra of breast cancer blood plasma samples, and therefore these positive loadings can be associated with the presence of disease. Similarly, in the negative loadings, the features at 700 (n (C-S) trans (aminoacid methionine)), 761 (Tryptophan, d (ring)), 798 (CH out of plane deformation) and 1410 cm^{-1} (ns COO_2) have higher intensities in mean Raman spectra of healthy volunteers than patient samples. Other Raman spectral features observed include those at 1100 cm^{-1} (C-C vibration mode of the gauche-bonded chain) as positive loadings and 1440 (CH_2 deformation in normal breast tissue) and 1663 cm^{-1} (DNA) as negative loadings. Moreover, Raman spectral features, including those at 689, 1138, 1173, 1185, 1268, 1319, 1440 and 1663 cm^{-1}, are elevated in breast cancer stage 4, and hence can be considered as a marker associated with the progression of disease.

In Figure 4 (a),the PCA scatter plot shows the comparison of two different breast cancer stages, stage-2 and stage-3. The spectra of the different stages are differentiated according to PC2, indicating that Raman spectroscopy of the blood serum can differentiate these two stages of the disease. The positive loadings of PC2 is associated with the Raman features of stage-3, which have higher intensities than that of stage-2, including those at 625cm^{-1}, 689cm^{-1} (nucleotides), 770 (phosphate), 788 (phosphodiester bands in DNA), 828 (tyrosine/protein), 1285 (phospholipids) 1307 (CH_3/CH_2 twisting, wagging and/or bending modeof collagens & lipids) and 1319 cm^{-1} (Guanine).

Therefore, although stage 2 and stage 3 are not well distinguished in the overall comparison of all patient samples to normal controls (Figure 2), the two stages of disease development are well differentiated in a pairwise comparison (Figure 4 a).

After thediagnosis of breast cancer, it is important to determine the staging of the cancer, because this will help to know that how far the disease has progressed and also to determine the best way to contain and eliminate the breast cancer.Stage 2 breast cancer is still in the earlier stages, but there is evidence that the cancer has begun to grow or spread. It is still contained to the breast area and is generally very effectively treated. On the other hand, stage 3 is known to be an advanced stage of breast cancer, at which the cancer has started to invade the surrounding tissues of the breast. In this regard, Raman spectroscopy can potentially be of benefit in the differentiation of these stages of breast cancer, hence leading to early diagnosis which could be useful for the effective treatment. The technique is shown to do this on the basis of the identification of the underlying biochemistry, the Raman spectral features, as indicated in (Figure 4 b). It is found that positive loadings of PC2 are associated with the Raman features of stage-3 which have higher intensities than that of stage-2, particularly those of nucleic acids. These features associated with DNA are very clear markers of cancer,as circulating DNA is only present in the blood plasma of the cancer patients. Moreover, this also indicates the progression from stage 2 to stage 3, as indicated by the dominant presence of these Raman features in the Raman spectra of the patients of stage-3 patients.

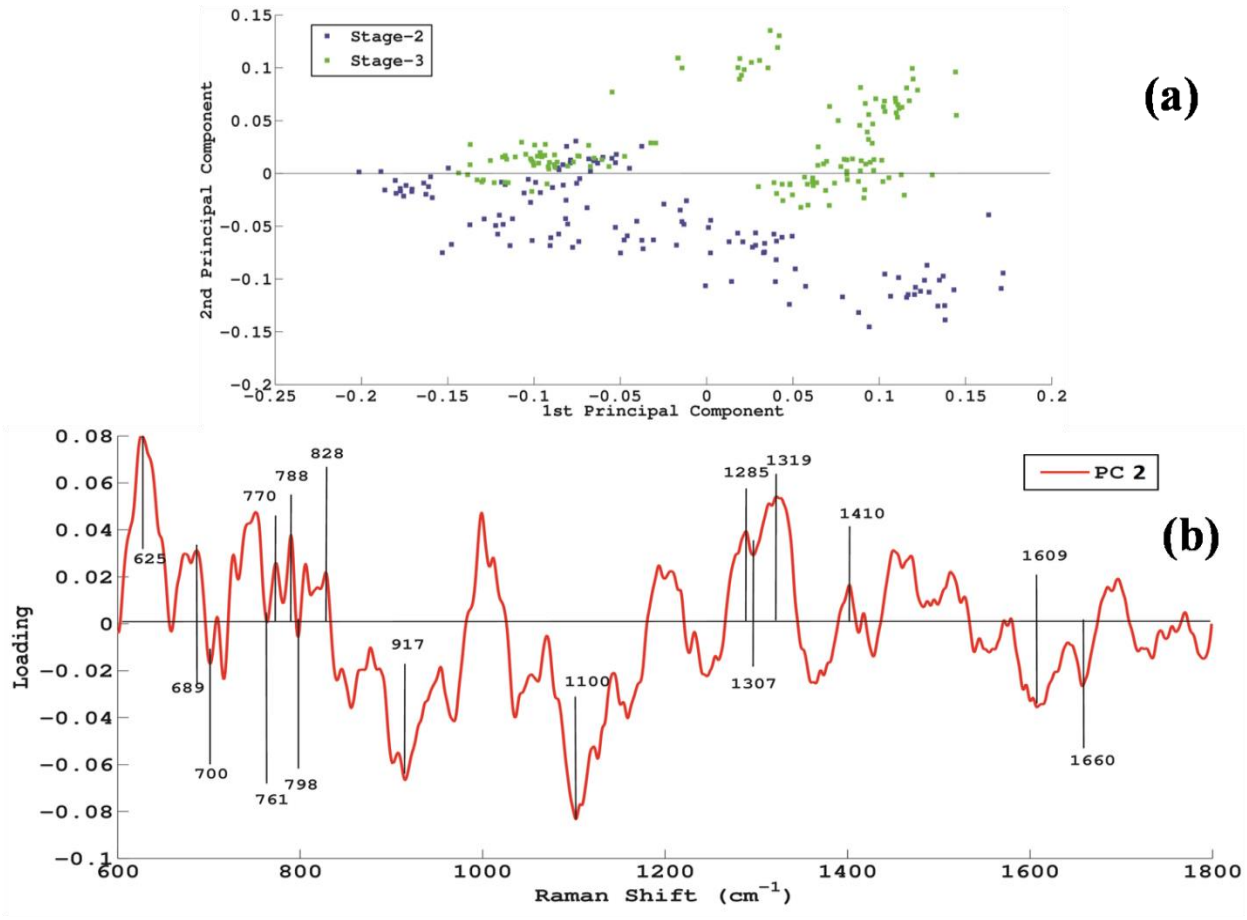

Figure 4: PCA scatter plot (a) and loading (b) of Raman spectral data of breast cancer samples of stage 2 verses stage 3.

Figure 5 (a)showsthe PCA scatter plot of breast cancer stage-2 and stage-4 and it can be seen that Raman spectral data of the stage-2 (blue dots) is differentiated from that of the stage-4 (red dots) by clustering in the negative and positive axis of the PC-1 respectively. InFigure 5 (b),the differences observed as positive loadings have higher intensities in the Raman spectral data of stage-4 than stage-3 patients, including those at 689cm^{-1} (nucleotides), 1185 cm^{-1} (anti-symmetric phosphate vibrations), 1285 cm^{-1} (phospholipids) and 1319 cm^{-1} (guanine (B,Z-marker)).

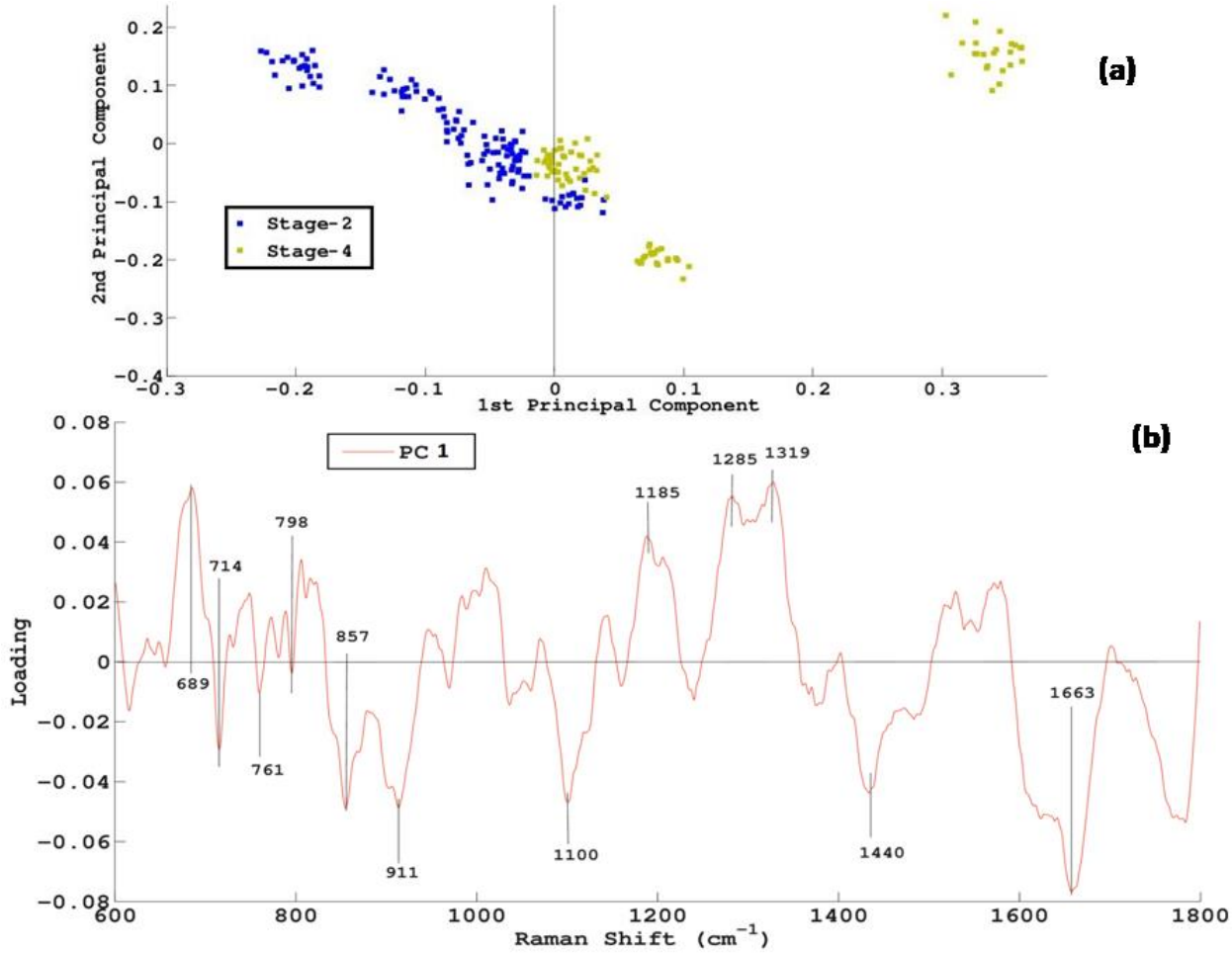

Figure 5: PCA scatter plot (a) and loadings (b) of Raman spectral data of breast cancer samples of stage 2 verses stage 4.

In Figure 6 (a) PCA scatter plot of stage-3 and stage-4 of the breast cancer is shown and separation between these two stages is observed according to PC-1, by clustering in the negative and positive axis respectively. In Figure 6 (b), which shows the loadings of the PC1, the positive loadings represent the Raman spectral features those have higher intensities in stage-4 as compared the stage-3. Notably, these Raman spectral features, including those at 689cm^{-1} (nucleotides), 1185 cm^{-1} (anti-symmetric phosphate vibrations), 1285 cm^{-1} (phospholipids) and 1319 cm^{-1} (guanine), have also been observed (figure 5) as the features associated with stage-4 and in figure 4 these features have higher intensities in stage-3 as compared with stage-2. Hence, these changes observed in the stage wise comparison for the breast cancer indicate the biochemical changes taking place during the development and progression of breast cancer.

In stage 4, thediseasehas spread to the surrounding tissues of the breast and it is considered the most advanced stage of breast cancer. It is very difficult to treat effectively at this stage of the cancer. The differentiation between Stage 2/3 and stage 4 is important because this can lead to early diagnosis which could be useful for the effective treatment.Notably, stage 2 is difficult to diagnose by histopathology unless a number of biopsies are taken. Moreover, this process is time consuming and histology protocols are not necessarily reliable. In this regard, Raman spectroscopic analysis of blood samples, a much more rapid analysis, is shown to be helpful in the detection of the early stage and differentiation of the later stages of breast cancer, predominantly on the basis of increasing DNA contentsduring the progression of the breast cancer from stage 2 to stage 4. The Raman spectral features, including those at 689cm^{-1} (nucleotides), 1185 cm^{-1} (Anti-symmetric phosphate vibrations), 1285 cm^{-1} (phospholipids) and 1319 cm^{-1} (Guanine), are observed (figure 5) as the features associated with stage-4 and in figure 4 these features have higher intensities in stage-4 as compared with stage-2/3. Hence, these changes observed in the stage wise comparison for the breast cancer indicate the biochemical changes taking place during the development and progression of breast cancer. Notably, all the patients of the stage 4 are found very much distinct from each other, indicated by the clustering pattern of the Raman spectral data of this stage in the PCA scatter plots (Figure 5 a and Figure 6a) which may be due to the increase in the degree of the severity of the breast cancer, and the diversity of routes towards metastasis.

This study advances previously published work related to the use of Raman spectroscopy for the characterization/diagnosis of breast cancer by using blood plasma/serum samples from clinically diagnosed patients, as the spectral markers of689 (Nucleotide conformation), 788 (phosphodiester bands in DNA), 828 (tyrosine/protein), 848 (most probably due to single bond stretching vibrations for amino acids and polysaccharides) and 1663 cm^{-1} (DNA) have not previously been identified. In these reports, Bilal et al. applied Partial least squares(PLS) regression analysis to the Raman spectral data acquired from blood serum of breast cancer and healthy individuals and coefficients of regression were determinedwith sensitivity of 90% and specificity 75% which are associated with the changes at the bio-molecular level [23], while in the current study sensitivity and specificity have been found to be as high as 100% in the comparison of control and breast cancer samples. In another study, Pichardo-Molina et al. applied Raman spectroscopy along with principal

components analysis (PCA) and PCA-Linear Discriminate Analysis (LDA) to breast cancer and control blood samples, identifying specific bands which distinguish breast cancer blood serum samples from the healthy ones [24]. In the current study, the technique has been applied to differentiate the different stages of breast cancer by using blood plasma samples and such Raman spectral changes are observed in the stage wise comparison for the breast cancer which indicates the biochemical changes taking place during the development and progression of breast cancer through different stages.

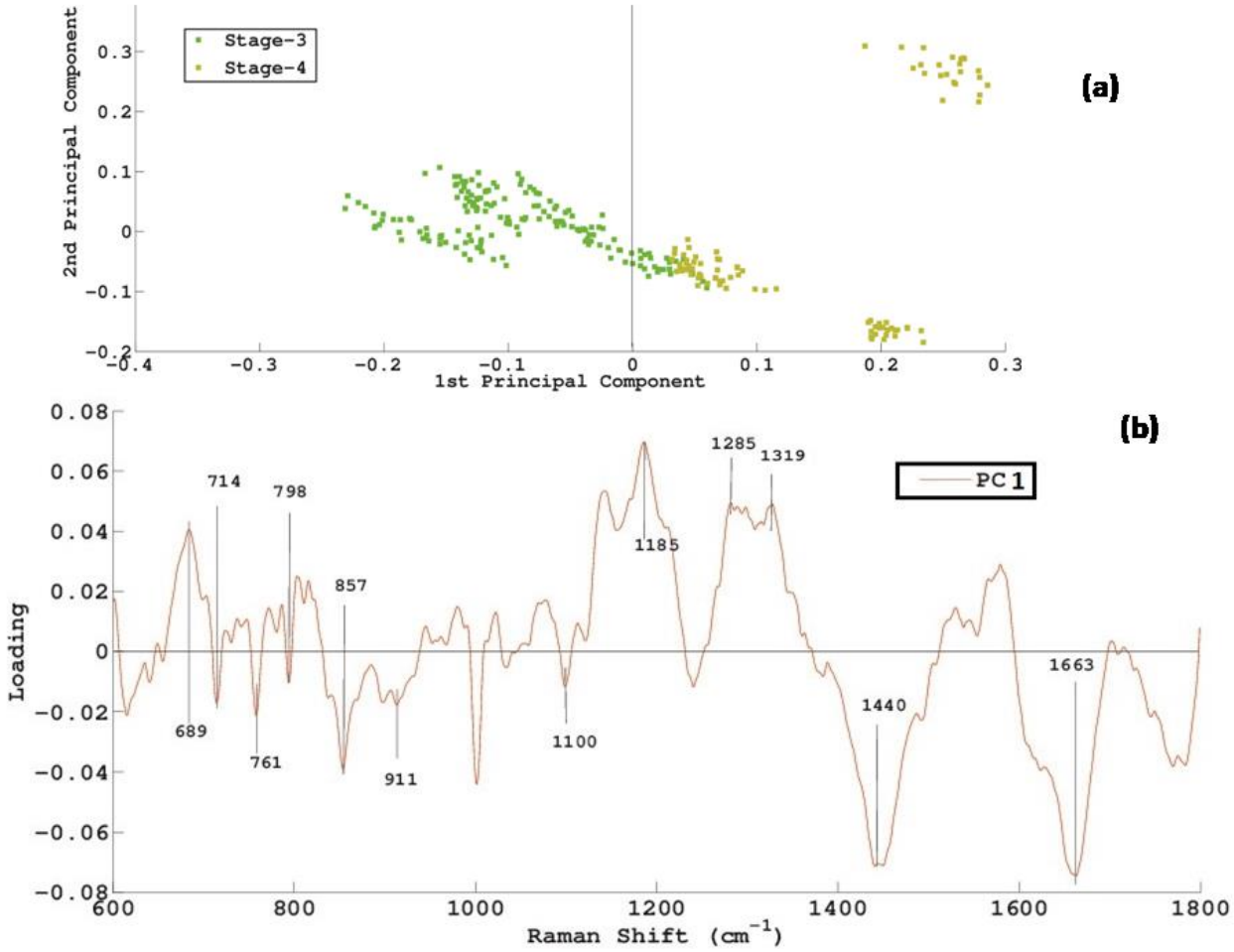

Figure 6: PCA scatter plot (a) and loadings (b) of Raman spectral data of breast cancer samples of stage 3 verses stage 4.

In order to quantify the diagnostic potential of the spectroscopic technique, a further classification of the control and breast cancer using PCA-FDA was carried out. The results are shown in **Table 2**, which indicates that the presence of breast cancer in patients can be detected in plasmawith a

sensitivity of 100% and a specificity of 100%. Furthermore, PCA-FDA shows sensitivity 99.65%and specificity 96.83% for breast cancer stage 2 versus stage 3/4 (**Table 3)**, sensitivity 96.3% and specificity 90.16% for breast cancer stage 3 versus stage 2/4 (**Table 4)** and sensitivity 93.10% and specificity 91.11% for breast cancer stage 4 versus stage 2/3 (**Table 5)**.

	Normal	**Breast cancer**	**Total/group**	**Specificity**
Normal	175	0	175	100%
Breast cancer	0	409	409	**Sensitivity**
		Total spectra	584	100%

Table 2: Results of the PCA-FDA for the Raman spectral data of Normal versus all breast cancer (stage 2, 3&4) samples.

	Breast cancer stage-2	**Breast cancer stage-3/4**	**Total/group**	**Specificity**
Breast cancer Stage-2	122	4	126	99.65%
Breast cancer stage-3/4	1	282	283	**Sensitivity**
		Total spectra	409	96.83%

Table 3: Results of the PCA-FDA for the Raman spectral data of breast cancer stage 2 versus stage 3/4samples.

	Breast cancer stage-3	**Breast cancer stage-2/4**	**Total/group**	**Specificity**
Breast cancer Stage-3	174	19	193	96.3%
Breast cancer stage-2/4	8	208	216	**Sensitivity**
		Total spectra	409	90.16%

Table 4: Results of the PCA-FDA for the Raman spectral data of breast cancer stage 3 versus stage 2/4 samples.

	Breast cancer stage-4	**Breast cancer stage-2/3**	**Total/group**	**Specificity**
Breast cancer Stage-4	82	8	90	93.10%
Breast cancer stage-2/3	22	297	319	**Sensitivity**
		Total spectra	409	91.11%

Table 5: Results of the PCA-FDA for the Raman spectral data of breast cancer stage 4 versus stage 2/3samples.

Conclusions

Raman spectroscopy along PCA has shown its potential to be used for the diagnosis of breast cancer at an early stage, by employing blood plasma samples and healthy controlstaken from clinically diagnosed breast cancer patients. Among several, Raman spectral features associated with DNA and proteins are identified, which are solely observed in the blood plasma samplesof the breast cancer patients. Moreover, there are several other spectral features which are strikingly different in the blood plasma samples of different stages of breast cancer. Raman spectral data of different stages of the disease is compared by employing PCA. PCA scatter plots have shown clear differentiation between healthy and disease as well as different stages of breast cancer.The differentiation between Stage 2/3 and stage 4 is important because this can lead to early diagnosis which could be useful for the effective treatment.

References

1. Backhaus, J., et al., *Diagnosis of breast cancer with infrared spectroscopy from serum samples.* Vibrational Spectroscopy, 2010. **52**(2): p. 173-177.
2. Forouzanfar, M.H., et al., *Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis.* The lancet, 2011. **378**(9801): p. 1461-1484.
3. Siegel, R.L., et al., *Colorectal cancer statistics, 2017.* CA: a cancer journal for clinicians, 2017. **67**(3): p. 177-193.
4. Mariotto, A.B., et al., *Projections of the cost of cancer care in the United States: 2010–2020.* Journal of the National Cancer Institute, 2011. **103**(2): p. 117-128.
5. Martínez, M., *El acceso al continuo de servicios entre niveles asistenciales en dos redes integradas de servicios de salud en Colombia: un estudio de casos múltiples de mujeres con cáncer de mama.* Facultad Nacional de Salud Pública, 2016. **30**(4.1).
6. Campeau, P.M., W.D. Foulkes, and M.D. Tischkowitz, *Hereditary breast cancer: new genetic developments, new therapeutic avenues.* Human genetics, 2008. **124**(1): p. 31-42.
7. Martin, A.-M. and B.L. Weber, *Genetic and hormonal risk factors in breast cancer.* Journal of the National Cancer Institute, 2000. **92**(14): p. 1126-1135.
8. Ekbom, A. and D. Hunter, *Pancreatic cancer.* Textbook of cancer epidemiology, 2008. **33**: p. 233-242.
9. Brandan, M.E. and Y. Villaseñor, *Detección del cáncer de mama: estado de la mamografía en México.* Cancerología, 2006. **1**(3): p. 14-162.
10. Lieber, C.A., et al., *Raman microspectroscopy for skin cancer detection in vitro.* Journal of biomedical optics, 2009. **13**(2): p. 024013.
11. Caspers, P., et al., *In vitro and in vivo Raman spectroscopy of human skin.* BIOSPECTROSCOPY-NEW YORK-, 1998. **4**: p. S31-S40.
12. Krafft, C., et al., *Near infrared Raman spectroscopic mapping of native brain tissue and intracranial tumors.* Analyst, 2005. **130**(7): p. 1070-1077.

13. Kumar, S., et al., *Linking carbon metabolism to carotenoid production in mycobacteria using Raman spectroscopy.* FEMS Microbiol. Lett, 2015. **362**: p. 1-6.
14. Gautam, R., et al., *Raman spectroscopic studies on screening of myopathies.* Analytical chemistry, 2015. **87**(4): p. 2187-2194.
15. Gautam, R., et al., *Raman and mid-infrared spectroscopic imaging: Applications and advancements.* 2015.
16. Caldeira, J.R.F., et al., *CDH1 promoter hypermethylation and E-cadherin protein expression in infiltrating breast cancer.* BMC cancer, 2006. **6**(1): p. 48.
17. Agarwal, G., et al., *Spectrum of breast cancer in Asian women.* World journal of surgery, 2007. **31**(5): p. 1031-1040.
18. Li, Q.-B., et al., *Diagnosis of gastric inflammation and malignancy in endoscopic biopsies based on Fourier transform infrared spectroscopy.* Clinical Chemistry, 2005. **51**(2): p. 346-350.
19. Alfano, R., et al., *Fluorescence spectra from cancerous and normal human breast and lung tissues.* IEEE Journal of Quantum Electronics, 1987. **23**(10): p. 1806-1811.
20. Liu, C., et al. *Human breast tissues studied by IR Fourier-transform Raman spectroscopy*. in *Conference on Lasers and Electro-Optics*. 1991: Optical Society of America.
21. Pu, Y., et al., *Native fluorescence spectra of human cancerous and normal breast tissues analyzed with non-negative constraint methods.* Applied optics, 2013. **52**(6): p. 1293-1301.
22. Teh, S., et al., *Near-infrared Raman spectroscopy for early diagnosis and typing of adenocarcinoma in the stomach.* British Journal of Surgery, 2010. **97**(4): p. 550-557.
23. Bilal, M., et al., *Optical screening of female breast cancer from whole blood using Raman spectroscopy.* Applied Spectroscopy, 2017. **71**(5): p. 1004-1013.
24. Pichardo-Molina, J., et al., *Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients.* Lasers in medical science, 2007. **22**(4): p. 229-236.
25. De Gelder, J., et al., *Reference database of Raman spectra of biological molecules.* Journal of Raman Spectroscopy, 2007. **38**(9): p. 1133-1147.
26. Meade, A.D., et al., *Growth substrate induced functional changes elucidated by FTIR and Raman spectroscopy in in–vitro cultured human keratinocytes.* Analytical and bioanalytical chemistry, 2007. **387**(5): p. 1717-1728.
27. Notingher, I., *Raman spectroscopy cell-based biosensors.* Sensors, 2007. **7**(8): p. 1343-1358.
28. Jess, P., et al., *Dual beam fibre trap for Raman microspectroscopy of single cells.* Optics Express, 2006. **14**(12): p. 5779-5791.
29. Movasaghi, Z., S. Rehman, and I.U. Rehman, *Raman spectroscopy of biological tissues.* Applied Spectroscopy Reviews, 2007. **42**(5): p. 493-541.
30. Nawaz, H., et al., *Evaluation of the potential of Raman microspectroscopy for prediction of chemotherapeutic response to cisplatin in lung adenocarcinoma.* Analyst, 2010. **135**(12): p. 3070-3076.
31. Bertrand, D., et al., *Stepwise canonical discriminant analysis of continuous digitalized signals: Application to chromatograms of wheat proteins.* Journal of Chemometrics, 1990. **4**(6): p. 413-427.
32. Chan, J.W., et al., *Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells.* Biophysical journal, 2006. **90**(2): p. 648-656.
33. Su, Y.-C., et al., *Antifungal activities and chemical compositions of essential oils from leaves of four eucalypts.* Taiwan Journal of Forest Science, 2006. **21**(1): p. 49-61.
34. Shetty, G., et al., *Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus.* British journal of cancer, 2006. **94**(10): p. 1460.
35. Stone, N., et al., *Raman spectroscopy for identification of epithelial cancers.* Faraday discussions, 2004. **126**: p. 141-157.

36. Forrest, G. and R. Lord, *Laser Raman spectroscopy of biomolecules. X-frequency and intensity of the phosphodiester stretching vibrations of cyclic nucleotides.* Journal of Raman Spectroscopy, 1977. **6**(1): p. 32-37.
37. Krafft, C., et al., *Near infrared Raman spectra of human brain lipids.* Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2005. **61**(7): p. 1529-1535.
38. Notingher, I., et al., *Discrimination between ricin and sulphur mustard toxicity in vitro using Raman spectroscopy.* Journal of the Royal Society Interface, 2004. **1**(1): p. 79-90.
39. Lambert, J.B., et al., *Introduction to organic spectroscopy*. 1987: Macmillan Publishing Company.
40. Cheng, W.T., et al., *Micro-Raman spectroscopy used to identify and grade human skin pilomatrixoma.* Microscopy research and technique, 2005. **68**(2): p. 75-79.
41. Gniadecka, M., et al., *Diagnosis of basal cell carcinoma by Raman spectroscopy.* Journal of Raman Spectroscopy, 1997. **28**(23): p. 125-129.
42. Huang, Z., et al., *Raman spectroscopy in combination with background near-infrared autofluorescence enhances the in vivo assessment of malignant tissues.* Photochemistry and photobiology, 2005. **81**(5): p. 1219-1226.
43. Ruiz-Chica, A., et al., *Characterization by Raman spectroscopy of conformational changes on guanine–cytosine and adenine–thymine oligonucleotides induced by aminooxy analogues of spermidine.* Journal of Raman spectroscopy, 2004. **35**(2): p. 93-100.
44. Andrus, P.G. and R.D. Strickland, *Cancer grading by Fourier transform infrared spectroscopy.* Biospectroscopy, 1998. **4**(1): p. 37-46.
45. Lakshmi, R.J., et al., *Tissue Raman spectroscopy for the study of radiation damage: brain irradiation of mice.* Radiation research, 2002. **157**(2): p. 175-182.
46. Malini, R., et al., *Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: a Raman spectroscopy study.* Biopolymers, 2006. **81**(3): p. 179-193.
47. Frank, C.J., R.L. McCreery, and D.C. Redd, *Raman spectroscopy of normal and diseased human breast tissues.* Analytical chemistry, 1995. **67**(5): p. 777-783.
48. Binoy, J., et al., *NIR-FT Raman and FT-IR spectral studies and ab initio calculations of the anti-cancer drug combretastatin-A4.* Journal of Raman Spectroscopy, 2004. **35**(11): p. 939-946.