

# Identification of new spectral signatures associated with dengue virus infected sera

Ayyaz Amin,<sup>a</sup> Nimrah Ghouri,<sup>a</sup> Safdar Ali,<sup>c</sup> M. Ahmed,<sup>b</sup> M. Saleem<sup>b\*</sup> and Javaria Qazi<sup>a\*</sup>

In a quest to use Raman spectroscopy as an optical diagnostic tool, we recorded Raman spectra of 32 dengue virus (DENV)-infected and 28 healthy sera samples in the near-infrared spectral range (540 to 1700  $\text{cm}^{-1}$ ) using laser at 785 nm as the excitation source. We observed clear differences in the Raman spectra of DENV-infected sera as compared with those of healthy individuals. Here, as a result of our study, we report 12 unique Raman bands associated with DENV-infected sera that are not reported earlier. After applying analysis of variance and t-test ( $p < 0.05$ ) on these 12 bands, six Raman bands at 630 (N-acetylglucosamine), 883 (in-plane bending (ring) of deoxyribose), 1218 (amide III- $\beta$  conformation from  $\text{C}_6\text{H}_5\text{-C}$  stretching vibrations of tryptophan and phenylalanine), 1273 (amide-III), 1623 (tryptophan) and 1672  $\text{cm}^{-1}$  (ceramide) were found only in the DENV-infected sera. The remaining six Raman bands at 716 (lipids), 780 (Uracil-based ring breathing mode), 828 (ring breathing tyrosine), 840 ( $\alpha$ -anomers), 1101 ( $\nu(\text{C-N})$  of lipids and DNA) and 1150  $\text{cm}^{-1}$  (glycogen/carotenoids) were only found in healthy sera. Two types of classification models, principal component analysis and linear discriminant analysis, were employed to develop principal component analysis-linear discriminant analysis model that has provided diagnostic accuracy 96.50%, sensitivity 93.44%, and specificity 100%. This indicates that these 12 Raman bands have the potential to be used as biomarkers for optical diagnosis of DENV infection. This study provides a new insight for future research in the field of optical diagnosis using Raman spectroscopy. Copyright © 2017 John Wiley & Sons, Ltd.

**Keywords:** dengue virus infection; Raman spectroscopy; optical diagnosis; partial least square regression (PLSR)

## Introduction

The comprehensive proteomic analyses of different biological fluids, particularly serum and plasma, have attracted substantial attention for the identification of protein biomarkers as early detection tools for humans.<sup>[1]</sup> Raman spectroscopy is an optical technique on the basis of inelastic scattering of light by vibrating molecules. It provides molecular/chemical signatures of cells, tissues and bio-fluids. Because of high chemical specificity, minimal or lack of sample preparation and the ability to use advanced optical technologies in the visible or near-infrared spectral range have recently led to an increase in medical diagnostic applications of Raman spectroscopy.<sup>[2–4]</sup> Majority of the reports published recently have explored the possibility of the use of Raman for cancer diagnostics. In case of infectious diseases, most research reports are on the possibility of diagnosis of malaria.<sup>[5,6]</sup> Only a small number of reports are available on the use of Raman spectroscopy technique for the analysis of chemical signatures arising from viral infections in human blood sera. These viruses include hepatitis B virus, hepatitis C virus, human immunodeficiency virus and dengue virus (DENV).<sup>[7–9]</sup>

Dengue fever is caused by DENV that is an arthropod-borne Flavivirus composing of four distinct serotypes (DENV-1, DENV-2, DENV-3 and DENV-4). Infection with DENV results in a range of clinical manifestations from asymptomatic infection to dengue fever and the severe disease, dengue haemorrhagic fever/dengue shock syndrome. Fortunately, most of DENV infections are asymptomatic or cause mild symptoms, which are characterized by undifferentiated fever with or without rash. According to a re-cent estimate, 390 million DENV infections occur per year

(95% credible interval 284–528 million), of which 96 million (67–136 million) have clinical manifestations.<sup>[10]</sup> Another study about the prevalence of DENV infection estimates that 3900 million people, in 128 countries, are at risk of infection with DENV.<sup>[11]</sup> In recent years in endemic settings (urban and semi-urban), transmission of DENV has increased especially in America, South and Southeast Asia and the Western Pacific. During the last decade, magnitude, distribution and clinical severity of dengue outbreaks have been alarmingly high in the countries of Indian subcontinent including Pakistan.<sup>[4]</sup> Pakistan is endemic to all four serotypes of DENV circulating throughout the year with a peak incidence during the post monsoon period between October and December.<sup>[12]</sup>

The DENV infection can be diagnosed by isolation of the virus, by serological tests or by molecular methods. Although choice of test depends upon multiple factors, however, serological assays are

\* Correspondence to: Javaria Qazi, Department of Biotechnology, Quaid-i-Azam University, Islamabad, Pakistan.  
E-mail: javariaq@yahoo.com

\* Correspondence to: M. Saleem, Agri. & Biophotonics Division, National Institute of Lasers and Optronics (NILOP), Lehtar Road, Islamabad, Pakistan.  
E-mail: salim569mail@gmail.com

a Department of Biotechnology, Quaid-i-Azam University, Islamabad, Pakistan

b Agri. & Biophotonics Division, National Institute of Lasers and Optronics (NILOP), Lehtar Road, Islamabad, Pakistan

c Directorate General National Repository, P.O. Box 1175, Islamabad, Pakistan

most widely employed because of simplicity and low cost of the tests as compared with nucleic acid detection assays. One type of serological assays detects viral glycoprotein (NS1 antigen) that is common to all DENV serotypes and can be used to detect either primary or secondary infections in the earlier stages. Other type of assay detects DENV-specific antibodies, IgG and IgM, confirming primary or secondary diagnosis, respectively. However, these diagnostic procedures for DENV infection have some limitations such as they are liable to blood-borne infection, invasive, delayed diagnosis, tiresome, time-consuming procedures, false positive results and called for well-trained technicians.<sup>[13]</sup> Therefore, for a globally prevalent viral disease like DENV infection, there is always a need for more specific, inexpensive diagnostic tests that can be used for clinical management, surveillance and outbreak investigations. Other than improvement in molecular biology techniques for detection, new techniques such as the use of biomarkers have the potential for improved diagnostics. Bodily fluids contain numerous biomolecules of diagnostic interest that can be measured in solution with Raman spectroscopy. Our previous work on the Raman spectroscopic analysis has revealed interesting information regarding bio-

markers in DENV-infected human blood sera as compared with healthy blood sera.<sup>[5,14]</sup> In this study, 32 DENV-infected samples are compared with 28 healthy sera in four age groups, and both genders were analysed using Raman spectroscopy using laser at 785 nm. Further, for classification of DENV-infected sera and healthy sera, we develop principal component analysis–linear discriminant analysis (PCA–LDA) data reduction and classification-based computational method, which achieved improved diagnosis accuracy.

## Materials and methods

### Sample collection and processing

Blood samples of 32 DENV-infected patients (ELISA confirmed) were collected during 5 October 2015 to 25 November 2015 from the Holy Family hospital and allied hospitals of Rawalpindi Medical College Rawalpindi. Samples were collected from the patients within formed consent. In addition, permission of Medical Review Board of Rawalpindi Medical College, Pakistan Medical Research Council that recently changed its name to Pakistan Health Research Council and Institutional Biosafety Committee of Quaid-i-Azam University was obtained prior to this study. Sampling was carried out at random from the ELISA confirmed cases. ELISA (Viracell Cat No. G1018, M1018) kits were used for confirmation of three types of DENV-specific antibodies, IgG, IgM and NS1 in each of these samples. All the healthy and infected blood samples were handled and transported following the standard biosafety procedures.<sup>[15]</sup> Later on, the samples belonging to both male and female patients were divided into four age groups (Figure S1). In the same age groups, blood samples of both genders of 28 healthy volunteers were obtained. To obtain sera, blood samples were centrifuged at 4000 rpm for 10 min using Eppendorf AG 22331 (Hamburg, Germany) centrifuge. Sera were aliquoted in three different sets for short-term and long-term storage at 4 °C, 20 °C and 80 °C to carry out research.

### Acquiring Raman spectra

Raman spectra were recorded directly from the sample by placing a small quantity (~20 µl) as a drop on the clean aluminum substrate using Raman system (Peak Seeker PRO-785 coupled with micro-scope RSM-785, Agiltron, USA) having a spectral resolution of

$10\text{ cm}^{-1}$ . It employed diode laser emitting at  $785\text{ nm}$  to excite the sample for Raman scattering. From each sample, 20 spectra were recorded from different positions by shining a laser power of  $150\text{ mW}$  for the acquisition time of  $10\text{ s}$ . The spectra were re-corded using  $10\times$  objective lens of microscope for focusing and ac-quiring. Laser was focused just beneath the surface of serum drop to get better Raman scattering. During acquiring Raman spectra, se-rum sample little bit settled down because of liquid nature; there-fore, same focusing level was always confirmed for each measurement with the help of a camera fixed with the microscope. Laser power was optimized carefully to avoid photodegradation.<sup>[16]</sup> This was performed by recording the Raman spectra from the same sample using different laser powers at  $50$ ,  $100$ ,  $150$  and  $200\text{ mW}$  for the same acquisition time of  $10\text{ s}$ . A spectral overlying of the Raman spectra was observed as a function of laser power as depicted in Fig. 1. However, laser power of  $150\text{ mW}$  was selected where Raman signatures are reasonably prominent to use them for analysis and to be on safe side.

[http://learnrnd.com/detail.php?id=Biological\\_microsensor](http://learnrnd.com/detail.php?id=Biological_microsensor)

[http://learnrnd.com/detail.php?id=Biowarfare\\_and\\_Germwarfare](http://learnrnd.com/detail.php?id=Biowarfare_and_Germwarfare)

## Spectral data processing

Raman spectra were obtained from  $300$  to  $1800\text{ cm}^{-1}$  spectral range that were preprocessed in MATLAB and established protocols<sup>[17]</sup> for the identification of molecular bonds and func-tional groups. Spectra from  $540$  to  $1700\text{ nm}$  were selected for spec-troscopic and statistical analysis because this range contained prominent Raman bands in healthy as well as diseased samples. Data preprocessing included smoothing, baseline correction and vector normalization that was performed using protocols.<sup>[17]</sup> All the spectra were smoothed using a Savitzky–Golay smoothing method (order  $5$ ,  $13$ -point window). A rubber band correction method was used for baseline removal in all the spectra that excel-

lently removed the fluorescence background from each spectrum.<sup>[18]</sup> Same protocols<sup>[17]</sup> were used for applying PCA on all the Raman spectra for clustering of individual species based on their Raman spectral fingerprints.

The Raman spectra were tested against one another using the analysis of variance (ANOVA) and t-test to verify the consistency

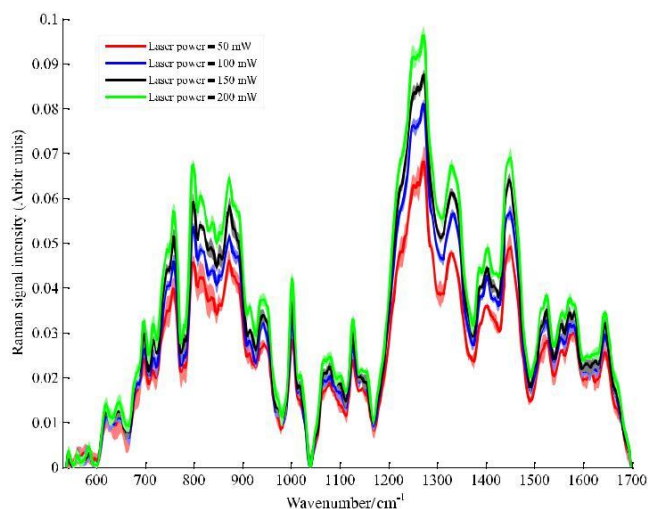


Figure 1. It shows the spectral overlaying of the Raman spectra that were recorded as a function of laser power for  $50$ ,  $100$ ,  $150$  and  $200\text{ mW}$  for the same acquisition time of  $10\text{ s}$ . This was aimed to find the appropriate

of the band, so a set of specific Raman bands is chosen. The ANOVA and t-test were used to decide which specific Raman bands would be statistically significant ( $p < 0.05$ ) and useful for discrimination of the healthy and DENV-infected sera. We reported specific Raman bands, which are statistically significant because the returned p-value is lower than the defined  $\alpha$ -value (0.05). Further, we visually inspected the Raman spectra in order to acquire evident spectral differences between healthy and DENV-infected spectra. The next section describes the development of classification model PCA-LDA.

### PCA-LDA model development

After data preprocessing (see Section on Spectral Data Processing), PCA-LDA models are developed using PCA in conjunction with LDA for classification of DENV-infected and healthy Raman spectra. PCA is a well-known unsupervised, multivariate and dimension-reduction method, which is extensively used to classify biological samples of distinct molecular species. It has previously been applied to study a single cell and the blood.<sup>[17,18]</sup> PCA utilizes variance that is present in the data to generate principal components (PCs), i.e. set of orthogonal features. The deficiency of this correlation means that the PCs represent valuable different 'dimen-sions' of the data. The order of the PCs describes their relative importance in the data set. PC1 describes the highest amount of variation, PC2 the second greatest and so on. The variance is some-times called scores of the PCs. The PCA score plot shows a degree of separation into two classes based on the small molecular variations in two types of spectra.

On the hand, LDA algorithm is a supervised classification method, which is extensively used in machine learning, pattern recognition and statistics for classification of two or more classes. For our two-class problem, classification of disease and healthy sam-ples, we employed linear discriminant function. This linear function generates decision surface (hyperplane) for the separation of two classes. LDA uses an optimized classification criterion, which exploiting within-class and between-class variance for improved classification. Following PCA for dimension reduction, the LDA-based classifier is built by presenting scores of the PCs that have the highest variance between the data. In this way, to obtain im-proved diagnosis accuracy, we develop PCA-LDA-based classifica-tion method, which performed data reduction and classification. During model development, to avoid overfitting, we adopted ten-fold jackknife cross-validation protocol.<sup>[44]</sup> PCA and LDA algorithms are implemented in MATLAB R2014a (Mathworks Inc., USA).

During model development, first, PCA employed to original data set  $D_{rs}$  of Raman spectra (observations), and only first few PC scores are kept for further use. The PC score matrix  $X$  is specified as follows:

$$X = D_{rs} PC_n \quad (1)$$

where  $PC_n$  represents the matrix of first  $n$  PCs. Next, we applied LDA on  $X$  to obtain LD score matrix  $Y$ .

$$Y = XW \quad (2)$$

Our basic goal is to search for the linear combination that optimally separates our multivariate observations into classes. LDA identifies classification criterion using the scatter (or covariance) measure of within-class and between-class variance. LDA algorithm attempts to find the weighting vector  $W$ , which maximizes the rate between the between-class scatter to within-class scatter. Defining the between-class scatter  $CS_b$

$$CS_b = \frac{1}{C} \sum_{i=1}^C \mu_i \mu_i^T - \mu \mu^T \quad (3)$$

and within-class scatter  $CS_w$

$$CS_w = \frac{1}{C} \sum_{i=1}^C \sum_{j=1}^{m_i} x_j x_j^T - \mu_i \mu_i^T \quad (4)$$

where  $C$  represents the number of classes (for our binary problem  $C = 2$ ),  $\mu_i$  denotes mean of class  $i$  and  $\mu$  represents mean of all such means,  $m_i$  denotes the total number of observations of the  $i$ th class,  $x_j$  is one such observation and  $T$  denotes the transport.

The optimizing criterion (i.e. objective function)  $J(W)$  is written using Eqns (3 and 4) as follows:

$$J(W) = \frac{W^T CS_b W}{W^T CS_w W} \quad (5)$$

Find  $W$ , which identifies the axes of the transformed space, such that  $J$  is maximized. The classes are parted by straight line for LDA.

## Results and discussion

### Raman spectral-based analysis

Raman spectroscopy being a rapid and nondestructive procedure is capable of producing information on conformation of biomolecules.<sup>[19–21]</sup> It has already been successfully used for detection of different diseases such as malaria and multiple types of cancers.<sup>[5,22]</sup> This study was aimed at investigating differential spatial signatures in DENV-infected sera as compared with the sera of healthy individuals. Findings are plotted as bicolour bands in Fig. 2, where red colour is assigned to the average spectra of DENV-infected sera and blue colour represents the average spectra of healthy sera. Comparison of Raman shifts among healthy and DENV-infected serum samples is shown in Table S1.

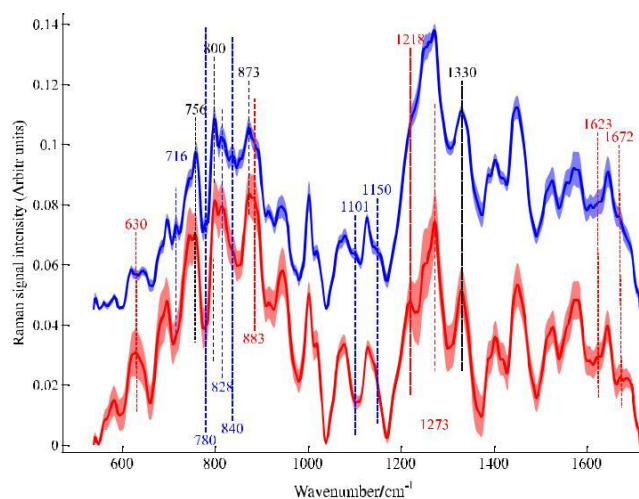


Figure 2. Average spectra of 28 healthy and 32 dengue virus (DENV)-infected sera samples. For each sample, 20 Raman spectra were recorded. It means that this figure contains average of 560 spectra of healthy spectra and average of 640 spectra of DENV-infected sera. Red colour is used to present average spectra of 32 dengue-infected sera samples, and blue colour is used to present average spectra of 28 healthy sera samples. Raman bands with labelled red colour are observed only in DENV-infected sera, labelled with blue are observed only in healthy and disappeared in DENV-infected sera whereas black labelled, the Raman bands that deformed because of DENV infection. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



Most important finding of this study is that we observed 12 distinctive Raman bands uniquely associated with DENV-infected sera that have not been reported earlier. Results of the ANOVA ( $F = 360.63$ ,  $p < 0.0001$ ) highlighted significant variation in the spectra because the returned  $p$ -value is lower than the defined  $\alpha$ -value (0.05) and the  $t$ -test confirms that the means for two groups are significantly different ( $p < 0.05$ ). The ANOVA and  $t$ -test were used to decide which specific Raman bands would be statistically significant ( $p < 0.05$ ) and useful for discrimination of the healthy and DENV-infected sera. We found specific Raman bands, which are statistically significant ( $p < 0.05$ ). These Raman bands were observed at 630, 716, 780, 840, 828, 883, 1101, 1150, 1218, 1273, 1623 and 1672  $\text{cm}^{-1}$  that may be indicating towards an initial set of DENV-specific biomolecular signatures to be tested in future during the prevalence of dengue disease in this region. Specific molecular signatures of these bands and their possible relevance for consistent association with DENV infection are described and discussed here one by one.

The Raman band at 630  $\text{cm}^{-1}$  arose because of N-acetylglucosamine, which is a carbohydrate assignment.<sup>[37]</sup> Crystal structure of protein E of DENV showed that it contains carbohydrate residues such as  $\alpha$ -mannose or N-acetylglucosamine required for virus binding to host cells.<sup>[39]</sup> Navarro-Sanchez et al. reported that E protein of DENV possesses two putative N-linked glycosylation sites and both are utilized for internalization of DENV into host cell. These sites correspond to Asn-67 and Asn-153. The site for Asn-153 is conserved among various flaviviruses, whereas site for Asn-153 is exclusive to DENV. The existence of both N-linked sugar moieties is essential for recognition through dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin<sup>[40]</sup> and suggests that this tetrameric lectin might be needed to interact with two or more sugar at a time for efficient entry of DENV into host cell.<sup>[46]</sup> Raman shift at 883  $\text{cm}^{-1}$  is previously reported for in-plane bending (ring) of deoxyribose (constituent of DNA).<sup>[23]</sup> DENV infection can trigger apoptosis in a variety of cell types, and apoptotic cells are the main source of cell-free DNA in the serum.<sup>[24,25]</sup> Circulating DNA levels of serum were significantly higher in patients with DENV infection than with other febrile illnesses<sup>[26,27]</sup> that give rise to this signature at 883  $\text{cm}^{-1}$  for in-plane bending of DNA. Variations in the amide-III region (1200–1300  $\text{cm}^{-1}$ ) reflect the most persuasive changes in the structure of proteins.<sup>[28]</sup> In this region, clear differences for DENV-infected sera, with two prominent Raman bands at 1218 and 1273  $\text{cm}^{-1}$ , are observed in the current study. These shifts were not observed in average spectra of healthy sera samples. Raman band at 1218  $\text{cm}^{-1}$  has been assigned to amide III- $\beta$  conformation of proteins (with contribution from  $\text{C}_6\text{H}_5\text{-C}$  stretching vibrations of tryptophan and phenylalanine).<sup>[29]</sup> Certain metabolite changes may be more general across different pathogenic infections. For example, Cui et al. reported that the serum level of phenylalanine is significantly raised in DENV infection. Phenylalanine 4-hydroxylase (PHA) having common cofactor Tetrahydrobiopterin with NO synthase (NOS) is required to convert phenylalanine into tyrosine. The increased level of NOS in DENV infection is reported to compete with PHA for cofactor Tetrahydrobiopterin. Thus, preventing the PHA function results in enhanced unconverted phenylalanine level in blood of DENV-infected patient.<sup>[30,31]</sup> Here, we propose that Raman band at 1218  $\text{cm}^{-1}$  is observed only in dengue-infected sera, possibly because of  $\text{C}_6\text{H}_5\text{-C}$  stretching vibrations of phenylalanine. Raman band at 828  $\text{cm}^{-1}$  observed only in average spectrum of healthy

sera is reported for ring breathing tyrosine.<sup>[42]</sup> Consistent to previous discussed results for Raman band at 1218  $\text{cm}^{-1}$ , the increased level of NOS in DENV infection is responsible for inhibition of conversion of phenylalanine into tyrosine suggesting absence of this band from the spectrum of DENV-infected sera. The Raman band at 1273  $\text{cm}^{-1}$  arose because of amide-III, which is a protein assignment.<sup>[32]</sup> Alteration in the concentration of proteins due to DENV infection has been reported in literature.<sup>[33]</sup> This band observed only in DENV-infected sera is likely due to the increased concentration of serum amyloid P. Serum amyloid P prevents infection although binding with residues of glucuronic acid, mannose and terminal N-acetylgalactosamine.<sup>[34]</sup> High levels of various serine protease inhibitors have been reported for dengue and other viral infection. Up regulation of protease inhibitor such as  $\alpha_1$  antitrypsin and  $\alpha_1$  antichymotrypsin may be attributed to host response against DENV infection to prevent vascular damage and inflammation.<sup>[35,36]</sup> The region 1672–1694  $\text{cm}^{-1}$  is clearly different for both healthy and DENV-infected sera where Raman band at 1672  $\text{cm}^{-1}$  observed only in DENV-infected sera is produced because of the presence of ceramide.<sup>[38]</sup> Ceramide with some other factors is involved in the activation of natural killer T-cells. Natural killer T-cells play a major role in the clearance of DENV. Moreover, it has also been reported that mast cell surveillance in DENV infection is significantly influenced by serum ceramide.<sup>[30]</sup> Raman band at 756  $\text{cm}^{-1}$  is reported for symmetric breathing of tryptophan<sup>[45]</sup> that is observed in average spectrum of both healthy and DENV-infected sera samples. However, this band showed higher intensity in average spectrum of DENV-infected sera samples. The higher intensity of this band validates the assignment of Raman band at 1218  $\text{cm}^{-1}$  for amide III- $\beta$  conformation of proteins (with contribution from  $\text{C}_6\text{H}_5\text{-C}$  stretching vibrations of phenylalanine and tryptophan). Raman shift due to ceramide and five previously discussed bands is distinctive for DENV-infected sera and can provide a useful insight for the use of this data towards Raman-based diagnostic test for dengue infection.

After discussing, Raman bands specific to DENV-infected sera, we will move towards the bands observed only in healthy sera such as Raman band at 716  $\text{cm}^{-1}$  depicting the presence of lipids.<sup>[9]</sup> Significant decrease of serum concentration of high-density lipoprotein, cholesterol and low-density lipoprotein in DENV-infected patients has been reported. Reduced plasma cholesterol level can be attributed to liver damage resulted because of dengue infection. Biosynthesis of cholesterol in human takes place in liver depending upon the cellular level of cholesterol, while high-density lipoprotein and low-density lipoprotein are involved in the transport of cholesterol.<sup>[41]</sup> The Raman band at 780  $\text{cm}^{-1}$  found only in spectra of control (healthy) samples has been assigned to Uracil-based ring breathing mode.<sup>[42]</sup> Viruses like DENV depend on host cellular factors to complete their life cycle because of small genome size. Therefore, viruses interact with host cellular proteins to get activities that are not encoded by their own genome.<sup>[49]</sup> Absence of Raman band at 780  $\text{cm}^{-1}$  in DENV sera samples might be due to an overall synthesis of new genetic code of DENVs. The Raman band at 840  $\text{cm}^{-1}$  found only in spectra of control samples has been assigned to  $\alpha$ -anomers.<sup>[50]</sup> However, its possible relevance to DENV infection is not reported in literature.

Raman band at 1150  $\text{cm}^{-1}$  observed only in average spectrum of healthy sera is reported for glycogen and carotenoids.<sup>[42]</sup> The peak intensity of this band for carotenoids is always higher in healthy sera samples when excited at laser of 532 nm. This is due to the

resonance effect for carotenoids absorption band (blue-green) for the excitation laser at 532 nm.<sup>[43]</sup> Using excitation laser at 785 nm, the peak intensity of this band is very low. The reason is that carotenoids are colour pigments and during centrifugation, these are separated from serum and only their residues left behind. However, because of their resonance absorption at 532 nm, a strong Raman signal is generated. Therefore, here, we proposed that Raman band at  $1150\text{ cm}^{-1}$  appeared because of glycogen. In a study conducted by Agarwal et al., minimal glycogen level was observed in skeletal muscle of dengue-infected mice up to day 5 gradually from the beginning of infection, which might be due to metabolic stress caused by virus replication.<sup>[47]</sup> Moreover, the reduction of glycogen reserves in hepatocytes<sup>[48]</sup> of dengue-infected mice was also reported by Milligan et al. These observations suggest that the decrease in serum concentration of DENV-infected patients might be due to depletion of glycogen reserves of the liver and skeletal muscles.

Our most interesting finding in this study is the 12 Raman bands that are uniquely associated to DENV-infected sera, and in the future, more studies utilizing more sample data can provide a useful insight for the use of this data towards Raman-based diagnostic test for dengue infection.

### Statistic-based analysis

Statistic-based analysis is carried out using PCA with LDA to classify the disease and healthy Raman spectra. From the analysis of PCA results, it is observed that the total variation of 94% in the spectrum is explained by PC1 (88%) and PC2 (6%). For further analysis and development of LDA classification model, only the first two PCs (PC1 and PC2) were considered.

Figure 3 depicts the classification of DENV-infected sera and healthy sera using proposed PCA-LDA method. The PCA-LDA plot is produced from 32 DENV-infected and 28 healthy samples. Because 20 Raman spectra were recorded from each sample, therefore, the PCA-LDA plot contains 640 DENV-infected spectra and 560 healthy spectra. In Fig. 3, the red colour circles

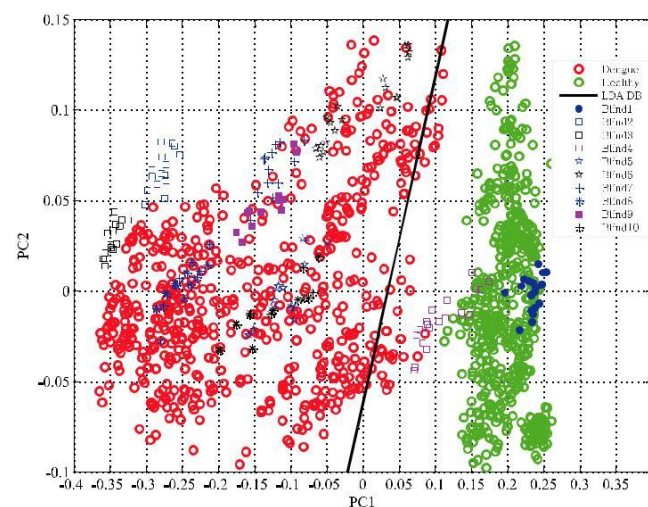


Figure 3. Classification of DENV-infected sera and healthy sera using PCA-LDA method. It shows the scatter plot of PC1 (88%) with PC2 (6%). Solid line corresponds to the DB, which is developed by the LDA classifier. It also contained ten double blindly tested samples. These were found 100% accurately classified. DB, decision boundary; LDA, linear discriminant analysis; PC, principal component; PCA, principal component analysis. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

represent the DENV-infected, and the green circles represent the healthy samples. Solid line (diagnostic line) represents the classification boundary generated by the LDA classifier. This solid line is obtained from Fisher's discriminant function of LDA  $39.62 * PC1 + 21.95 * PC2 = 1.38$ . The PCA-LDA model is evaluated using tenfold cross-validation technique. The performance measures of accuracy, sensitivity and specificity are used to assess the classification performance of the model. Sensitivity represents correctly recognized DENV-infected samples, whereas specificity correctly sorts out healthy samples.<sup>[44]</sup> The proposed PCA-LDA model provided diagnostic accuracy of 96.50%, sensitivity of 93.44% and specificity of 100%. It is deduced that when PCA combined with LDA, the diagnostics accuracy of the classification improved. Overall, the results achieved are quite interesting and promising.

After the successful implementation of PCA-LDA on the spec-tra, ten samples were blindly tested. Their clinical results kept hidden during the testing of PCA-LDA model. In Fig. 3, the pre-diction of the blind samples on the PCA-LDA plot interestingly depicted 100% accurate result when these were compared with the clinical results. There were two samples named blind-1 and blind-4; those were clustered in the healthy region of PCA-LDA, and the rest were clustered in the DENV-infected region. The description of the blind samples is pictorially shown on the PCA-LDA plot in Fig. 3 with different colours and different shapes in the legend.

### Conclusion

Dengue fever, caused by the infection of mosquito-born DENV, is a severe public health concern specially in tropical and sub-tropical regions around the globe. Lack of specific treatment and inadequacy of conventional diagnostic tools is the major problems in the management of this lethal type of infection. Present study aimed to reveal the capacity of Raman to be a possible tool for DENV infection diagnosis. The average spectra of healthy and DENV-infected sera samples are discriminated in these studies at 630, 716, 780, 840, 828, 883, 1101, 1150, 1218, 1273, 1623 and  $1672\text{ cm}^{-1}$  using ANOVA and t-test ( $p < 0.05$ ). These Raman bands corresponding to unique molecular compositions associated with DENV infection and can therefore be claimed potential biomarkers for the diagnosis of DENV infection. Raman bands, statistically significant ( $p < 0.05$ ), at 630, 883, 1218, 1273, 1623 and  $1672\text{ cm}^{-1}$  were consistently absent in healthy sera samples, indicating the potential opportunity for the fast, accurate and early diagnosis of DENV infection. In this study, we have used identified Raman bands to build classification models (PCA-LDA) for discrimination of DENV-infected sera and healthy sera. These models were evaluated using tenfold cross-validation technique. A diagnostic accuracy 96.50%, sensitivity 93.44% and specificity 100% were yielded by the developed PCA-LDA model. The acquired results are very interesting and promising. However, further experiments involving large number of samples with consistent data are required to accept this set of signatures as Raman-based diagnostic tool for clinical diagnosis.

### Acknowledgements

This work is supported by a research grant provided by Pakistan Medical & Research Council (4-22-25/13/RDC/QAU). In addition, we are thankful to the staff of pathology lab at Holy Family hospital

Rawalpindi for proving the samples and Mrs Fatima/Mr Irfan (Lab technicians at NILOP) for acquiring Raman spectra used in the pres-ent studies.

## References

- [1] S. Ray, P. J. Reddy, R. Jain, K. Gollapalli, A. Moiyadi, S. Srivastava, *Proteomics* 2011, 11, 2139.
- [2] [http://learnrmd.com/detail.php?id=Age\\_Effects\\_on\\_Brain\\_with\\_Symptoms\\_and\\_their\\_Solutions](http://learnrmd.com/detail.php?id=Age_Effects_on_Brain_with_Symptoms_and_their_Solutions)
- [3] K. Kong, C. Kendall, N. Stone, I. Notinghe, *Adv. Drug Deliv. Rev.* 2015, 89, 121.
- [4] S. Li, L. Li, Q. Zeng, Y. Zhang, Z. Guo, Z. Liu, M. Jin, C. Su, L. Lin, J. Xu, S. Liu, *Sci. Rep.* 2015, 5, 9582.
- [5] A. Sharma, S. Sharma, A. Zarrow, R. A. Schwartz, W. C. Lambert, *Indian J. Dermatol. Venereol.* 2016, 61, 1.
- [6] M. Bilal, M. Saleem, S. T. Amanat, H. A. Shakoor, R. Rashid, A. Mahmood, M. Ahmed, *J. Biomed. Opt.* 2015, 20, 017002.
- [6] Y. Clement, Q. Liu, *J. Biomed. Opt.* 2012, 17, 0170051.
- [7] S. Firdous, M. Ahmed, A. Rehman, M. Nawaz, S. Anwar, S. Murtaza, *Laser Phys. Lett.* 2012, 9, 317.
- [8] A. Shahzad, S. Firdous, *Laser Phys. Lett.* 2015, 12, 076001.
- [9] L. Sitole, F. Steffens, D. Meyer, *Curr. Metabolomics*. 2015, 3, 65.
- [10] S. Bhatt, P. W. Gething, O. J. Brady, J. P. Messina, A. W. Farlow, C. L. Moyes, J. M. Drake, J. M. Brownstein, J. S. Hoen, A. G. Sankoh, M. F. Myers, *Nature* 2013, 496, 504.
- [11] S. Q. Zaman, E. Roshan, A. Mahmood, S. Mahmood, *Pak Armed Forces. Med J.* 2016, 66, 16
- [12] [http://learnrmd.com/detail.php?id=PAK-CHINA\\_ECONOMIC\\_CORRIDOR](http://learnrmd.com/detail.php?id=PAK-CHINA_ECONOMIC_CORRIDOR)
- [13] A. R. M. Radzol, Y. L. Khuan, and W. Mansor, *IEEE/EMBS Spec. Top. Conf. Microtechnol. Med. Biol.*, 3rd. 2013, pp. 1438-1441.
- [14] M. Saleem, M. Bilal, S. Anwar, A. Rehman, M. Ahmed, *Laser Phys. Lett.* 2013, 10, 035602.
- [15] [http://learnrmd.com/detail.php?id=Age\\_Effects\\_on\\_Brain\\_with\\_Symptoms\\_and\\_their\\_Solutions](http://learnrmd.com/detail.php?id=Age_Effects_on_Brain_with_Symptoms_and_their_Solutions)
- [16] J. L. Pichardo-Molina, C. Frausto-Reyes, O. Barbosa-García, R. Huerta-Franco, J. L. González-Trujillo, C. A. Ramírez-Alvarado, G. Gutiérrez-Juárez, C. Medina-Gutiérrez, *Lasers Med. Sci.* 2007, 22, 229.
- [17] Knief P (2010) PhD Thesis Dublin Institute of Technology.
- [18] [http://learnrmd.com/detail.php?id=Types\\_of\\_Obesity\\_Its\\_Symptoms\\_and\\_Prevention](http://learnrmd.com/detail.php?id=Types_of_Obesity_Its_Symptoms_and_Prevention)
- [19] H. J. Byrne, P. Knief, M. E. Keating, F. Bonnier, *Chem. Soc. Rev.* 2016.
- [20] S. R. Panikkanvalappil, M. James, S. M. Hira, J. Mobley, T. Jilling, N. Ambalavanan, M. A. El-Sayed, *J. Am. Chem. Soc.* 2016.
- [21] S. Khan, R. Ullah, M. Saleem, M. Bilal, R. Rashid, I. Khan, A. Mahmood, M. Nawaz, *Optik* 2016, 127, 2086.
- [22] W. Wang, J. Zhao, M. Short, H. Zeng, *J. Biophotonics* 2015, 8, 527.
- [23] S. O. L. Janus, K. S. Małek, M. G. Głogowska, T. Walski, M. Komorowska, W. Witkiewicz, C. Pezowicz, M. Kobielarz, S. Szotek, *Acta Bioeng. Biomech.* 2012, 14, 101.
- [24] A. V. Lichtenstein, S. M. Hovsep, L. D. Tomei, S. R. Umansky, *Ann. N. Y. Acad. Sci.* 2001, 945, 249.
- [25] K. S. Myint, T. P. Endy, D. Mongkolsirichaikul, C. Manomuth, S. Kalayanaroj, D. W. Vaughn, A. Nisalak, S. Green, A. L. Rothman, F. A. Ennis, D. H. Libraty, *J. Infect. Dis.* 2006, 194, 607.
- [26] D. V. John, Y. S. Lin, G. C. Perng, *J. Biomed. Sci.* 2015, 22, 1.
- [27] T. T. N. Ha, N. T. Huy, L. A. Murao, N. T. P. Lan, T. T. Thuy, H. M. Tuan, C. T. P. Nga, V. V. Tuong, T. Van Dat, M. Kikuchi, M. Yasunami, *PLoS One* 2011, 6, e25969
- [28] I. Choi, Y. S. Huh, D. Erickson, *Microfluid. Nanofluid.* 2012, 12, 669.
- [29] S. R. Panikkanvalappil, S. M. Hira, M. A. El-Sayed, *Chem. Sci.* 2016, 7, 1133.
- [30] L. Cui, Y. H. Lee, Y. Kumar, F. Xu, K. Lu, E. E. Ooi, S. R. Tannenbaum, C. N. Ong, *PLoS Negl. Trop. Dis.* 2013, 7, e2373
- [31] R. Takhampunya, R. Padmanabhan, S. Ubol, *J. Gen. Virol.* 2006, 87, 3003.
- [32] S. G. Demos, A. J. Vogel, A. H. Gandjbakhche, *J. Mammary Gland, Biol. Neoplasia.* 2006, 11, 165.
- [33] H. C. Chiu, H. Hannemann, K. J. Heesom, D. A. Matthews, A. D. Davidson, *PLoS One* 2014, 9, e93305
- [34] A. P. Kaplan, K. Joseph, M. Silverberg, *J. Allergy, Clin. Immunol.* 2002, 109, 195.
- [35] S. Ray, R. Srivastava, K. Tripathi, V. Vaibhav, S. Patankar, S. Srivastava, *OMICS* 2012, 16, 527.
- [36] X. F. Tan, S. Wu, S. Li, Z. Chen, F. Chen, *Virol. J.* 2011, 8, 1.
- [37] G. Li, H. Yang, Y. Xu, Z. Zhang, *Sci. China Ser. C Life Sci.* 2002, 45, 405.
- [38] L. M. Almond, J. Hutchings, G. Lloyd, H. Barr, N. Shepherd, J. Day, O. Stevens, S. Sanders, M. Wadley, N. Stone, C. Kendall, *Gastrointest. Endosc.* 2014, 79, 37.
- [39] J. de Jesús Martínez-Barraganand, R. M. Del Angel, *J. Virol.* 2001, 75, 7827.
- [40] E. Navarro-Sanchez, R. Altmeyer, A. Mara, O. Schwartz, F. Fieschi, J. L. Virelizier, F. Arenzana-Seisdedos, and P. Despre's, *EMBO Rep.* 2003.
- [41] H. H. Biswas, A. Gordon, A. Nuñez, M. A. Perez, A. Balmaseda, E. Harris, *PLoS Negl. Trop. Dis.* 2015, 9, e0003904
- [42] Z. Movasaghi, S. Rehman, I. U. Rehman, *Appl. Spectrosc. Rev.* 2007, 42, 493.
- [43] M. E. Darvin, I. Gersonde, H. Albrecht, M. Meinke, W. Sterry, J. Lademann, *Laser Phys. Lett.* 2006, 3, 463.
- [44] S. Ali, A. Majid and A. Khan, *Amino Acids* 2014, 46, 977.
- [45] L. Su, Y. F. Sun, Y. Chen, P. Chen, A. G. Shen, X. H. Wang, J. Jia, Y. F. Zhao, X. D. Zhou, J. M. Hu, *Laser Phys.* 2012, 22, 311.
- [46] F. A. Rey, *Proc. Natl. Acad. Sci. U. S. A.* 2003, 100, 6901.
- [47] D. K. Agrawal, P. Tandon, U. C. Chaturvedi, A. Kumar, *J. Gen. Virol.* 1978, 40, 408.
- [48] G. N. Milligan, V. V. Sarathy, E. Infante, L. Li, G. A. Campbell, P. R. Beatty, E. Harris, A. D. Barrett, N. Bourne, *PLoS One* 2015, 10, e0125476
- [49] S. Khadka, A. D. Vangeloff, C. Zhang, P. Siddavatam, N. S. Heaton, L. Wang, R. Sengupta, S. Sahasrabudh, G. Randall, M. Gribskov, R. J. Kuhn, R. Perera, D. J. LaCount, *Mol. Cell. Proteomics* 2011, 10, 012187
- [http://learnrmd.com/detail.php?id=Chapter\\_wise\\_Corpus\\_Based\\_Study\\_of\\_the\\_Style\\_in\\_Shakespearean\\_Play\\_Hamlet](http://learnrmd.com/detail.php?id=Chapter_wise_Corpus_Based_Study_of_the_Style_in_Shakespearean_Play_Hamlet)