

2017

## Development of a High Throughput (HT) Raman Spectroscopy Method for Rapid Screening of Liquid Blood Plasma From Prostate Cancer Patients

Dinesh Medipally

*Technological University Dublin*

Adrian Maguire

*Technological University Dublin, [adrian.maguire@mydit.ie](mailto:adrian.maguire@mydit.ie)*

Jane Bryant

*Technological University Dublin, [jane.bryant@tudublin.ie](mailto:jane.bryant@tudublin.ie)*

*See next page for additional authors*

Follow this and additional works at: <https://arrow.tudublin.ie/radart>

 Part of the [Medicine and Health Sciences Commons](#)

---

### Recommended Citation

Medipally, D. et al (2017). Development of a high throughput (HT) Raman spectroscopy method for rapid screening of liquid blood plasma from prostate cancer patients. *Analyst*, 142, pp.1216-1226 doi:10.1039/C6AN02100J

This Article is brought to you for free and open access by the Radiation and Environmental Science Centre at ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact [arrow.admin@tudublin.ie](mailto:arrow.admin@tudublin.ie), [aisling.coyne@tudublin.ie](mailto:aisling.coyne@tudublin.ie), [vera.kilshaw@tudublin.ie](mailto:vera.kilshaw@tudublin.ie).

---

**Authors**

Dinesh Medipally, Adrian Maguire, Jane Bryant, John Armstrong, Mary Dunne, Marie Finn, Fiona Lyng, and Aidan Meade

## **Development of a high throughput (HT) Raman spectroscopy method for rapid screening of liquid blood plasma from prostate cancer patients.**

Dinesh K. Medipally<sup>1, 2</sup>, Adrian Maguire<sup>1, 2</sup>, Jane Bryant<sup>2</sup>, John Armstrong<sup>3</sup>, Mary Dunne<sup>3</sup>, Marie Finn<sup>3</sup>, Fiona M. Lyng<sup>\*1, 2</sup> and Aidan D. Meade<sup>\*1, 2</sup>.

*1. School of Physics, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland.*

*2. DIT Centre for Radiation and Environmental Science, Focas Research Institute, Dublin Institute of Technology, Camden Row, Dublin 8, Ireland.*

*3. Department of Radiation Oncology, Saint Luke's Radiation Oncology Network, St Luke's Hospital, Dublin, Ireland.*

*\* These authors contributed equally*

### **Abstract**

Extensive research has been undertaken on the examination of tissue biopsies using vibrational spectroscopic techniques. However, fewer studies have focused on less invasive and commonly acquired blood samples. Recent studies have shown the ability of Raman and Fourier transform infrared spectroscopy (FTIR) spectroscopy to discriminate between non-cancer controls and cancer cases using blood serum or plasma. Even though many studies have proposed Raman spectroscopy as a potential diagnostic tool in various cancers, the Raman spectroscopic technique has not been introduced as a routine clinical technology. This is due to multiple drawbacks with the application of the technique, including sample preparation, the requirement for expensive substrates and long acquisition times. The current study aims to overcome these limitations and focuses on the translation of Raman spectroscopy into a high throughput clinical diagnostic tool for prostate cancer.

In this study, the effect of various instrumental and sample preparation parameters were investigated, with the aim of identifying a combination that would reduce the overall acquisition time for spectra from peripheral blood plasma, reduce the complexity of sample preparation and retain the classification accuracy from Raman spectroscopic diagnostics. A high throughput (HT) system was developed and Raman spectroscopic measurements were performed on plasma samples from 10 prostate cancer patients and 10 healthy volunteers. The spectra were pre-processed and classified by principal component analysis - linear discriminant analysis (PCA-LDA) in the R environment. Statistically significant differences were observed

between Raman spectra of prostate cancer patients and non-cancer controls. The (HT) classification resulted in a sensitivity and specificity of 96.5% and 95% respectively. Overall, this study has overcome some of the limitations associated with clinical translation of Raman spectroscopy. The HT-Raman spectroscopy method developed in this study can be used for rapid and accurate diagnosis of prostate cancer using liquid plasma samples.

**Keywords:** Raman spectroscopy, High throughput, Prostate cancer, Biofluids, Blood plasma

## **Introduction**

In western countries, prostate cancer is the second most frequently diagnosed cancer and the third most common cause of death from cancer in men <sup>1</sup>. Clinically there are several means of detecting prostate cancer including B-mode ultrasound, biopsy, and tumor marker screening, although each of these techniques possess some disadvantages. B-mode ultrasound images the formed solid tumor in the prostate gland and thus can only detect patients that have progressed through the early stages of cancer <sup>2</sup>. Biopsy is the gold standard for detection of cancer, but it is invasive and impractical for a high-risk patient with multiple suspicious lesions, and particularly so for deep-seated cancers such as that of the prostate. Tumor marker screening, in particular the measurement of prostate specific antigen (PSA), has significantly improved early diagnosis. The PSA test has a good specificity (91%) but is limited in sensitivity (20.5%). PSA levels can be high in some subjects without prostate cancer leading to overdiagnosis; also PSA has poorer discriminating ability in men with symptomatic benign prostatic hyperplasia <sup>3</sup>.

The potential of Raman and FTIR spectroscopy for diagnostic applications has been well investigated and demonstrated<sup>4,5,6,7</sup>. The detailed information obtained from Raman spectroscopy provides information about the molecular structure and the composition of cells, tissues and biofluids, and ultimately promises an analysis of disease origin and progression. The analysis of biofluids can enable powerful minimally invasive diagnostics for many diseases <sup>8</sup>.

Infrared (IR) spectroscopy is based on the absorption of infrared radiation by the sample under study and the fact that molecules absorb specific frequencies of the incident light which are characteristic of their structure. FTIR spectroscopy is a non-destructive, label-free method for

studying molecular composition and the structure of macromolecules, either in their isolated form<sup>9</sup> or within complex biological systems such as cells and tissues<sup>10</sup>. It can quantitatively or qualitatively analyse spectral variations within the sample that may be linked to various molecular constituents, such as nucleic acids, carbohydrates proteins or lipids<sup>11</sup>. The modality has been recently applied to biofluids for screening diseases, such as Alzheimer's disease, galactosemia, hepatic fibrosis and hepatocellular carcinoma with high specificity and sensitivity (> 85%)<sup>7,11,12,13</sup>. However there are some disadvantages associated with FTIR spectroscopic analysis of biofluids. Biofluids must be measured with FTIR spectroscopy in their dried state due to absorbance of water in the mid IR region. As glass has a spectral signature in the mid-IR region, glass slides cannot be used as substrates. In addition, because the FTIR method imposes constraints on sample thickness, uniformity and dilution to avoid saturation, some sample preparation is needed<sup>14</sup>.

Raman spectroscopy is based on inelastic scattering of monochromatic light, usually from a laser source. Inelastic scattering means that the frequency of photons in monochromatic light changes upon interaction with a sample<sup>15</sup>. Like FTIR spectroscopy, Raman spectroscopy is a label free, non-destructive technique for studying molecular composition. Key advantages include that it requires little to no sample preparation and has no interference from water. This is due to polar molecules, such as water, producing relatively weak Raman signals<sup>16</sup>. However, unlike FTIR spectroscopy, long spectral acquisition times are needed to achieve adequate SNR<sup>17</sup>. Raman spectroscopy has recently been applied to biofluids (serum or plasma) for discriminating between non cancer controls and head and neck cancer patients<sup>18</sup>, breast cancer patients<sup>20</sup> and cervical cancer patients<sup>21</sup> with sensitivity and specificity above 75%.

Despite its proven ability, Raman spectroscopy lags behind FTIR spectroscopy for the purposes of clinical diagnosis with biofluids due to certain limitations. Raman spectroscopic studies of biofluids to date have mainly been performed on dried samples deposited on spectroscopically neutral substrates such as CaF<sub>2</sub>. However, sample homogeneity and measurement reproducibility can be a problem<sup>22,23</sup>. The CaF<sub>2</sub> substrates are also very expensive and cannot be reused in clinical diagnosis. Due to the relatively low interaction cross-section for the Raman interaction, this also necessitates relatively long acquisition times in comparison with applications involving FTIR spectroscopy.

Therefore, the present study aims to overcome these limitations and contribute to the translation of Raman spectroscopy of biofluids to the clinic. In the study, a HT-Raman spectroscopy

method was developed for rapid screening of prostate cancer patients. Different instrumental and sample preparation parameters were tested for the identification of a suitable parameter set for the HT-Raman spectroscopy method. This HT method was developed using low cost substrates and liquid plasma samples drawn from prostate cancer patients and healthy volunteers. The classification performance of this method provides an overall higher sensitivity (96.5%) and specificity (95%) when compared to the conventional PSA test. In addition this method can analyse plasma samples in a 96-well microplate in less than one hour. Overall, this HT-Raman spectroscopy approach has the potential to translate into clinics for HT screening of prostate cancer patients using samples of liquid plasma samples. It also has potential to be adaptable to any other human body fluid.

## **Materials and Methods**

### **Ethical approval**

Blood samples were obtained from healthy volunteers after informed consent. Ethical approval has been obtained from the DIT Research Ethics Committee for this arm of the work. Blood samples from prostate cancer patients were obtained through an ongoing collaboration with Prof John Armstrong, Consultant Radiation Oncologist, St Luke's Hospital (Dublin, Ireland). Ethical approval has been obtained from St Luke's Research Ethics Committee. The study is also sanctioned as part of an All-Ireland Cooperative Oncology Research Group (ICORG) translational study, ICORG 08-17.

### **Patient and control demographics**

The healthy controls were a cohort of both males and females, smokers and non-smokers, all between the ages of 21 and 64, while the patients were male with an age range from 58 to 85.

### **Plasma separation**

Whole blood was drawn into Lithium-heparin tubes from 10 healthy volunteers and 10 prostate cancer patients. Plasma was isolated from these blood samples by centrifugation at 3500 x g for 5 minutes at 18<sup>0</sup>C. The samples were subsequently stored at -80<sup>0</sup>C prior to Raman acquisition.

## **Raman spectroscopy**

### **Optimisation study**

Extensive optimisation work has been carried out for the development of the HT Raman spectroscopy method. Plasma samples from healthy volunteers were used for the optimisation study. A Horiba Jobin Yvon LabRAM HR 800 was used to acquire Raman spectra from healthy plasma samples. The instrument was calibrated with the spectra using silicon ( $520.7\text{cm}^{-1}$ ). A range of parameters (Table 1) were tested during the optimisation of HT-Raman spectroscopy.

### **HT-Raman spectroscopic analysis of blood plasma samples**

Frozen plasma samples, after passive thawing, were subjected to Raman spectroscopy by placing 20  $\mu\text{l}$  plasma on a cover glass bottomed 96 well plate and the spectra were recorded automatically from each well where the spectrometer was programmed using an in house developed high throughput macro template.

Spectra of 1,4-Bis (2-methylstyryl) benzene were acquired prior to each daily measurement as a further wavenumber calibration, and sample spectra from each day were wavenumber corrected to the standard spectrum of 1,4-Bis (2-methylstyryl) benzene. The plasma samples were excited with the 785nm, 660nm, 532 nm and 473nm laser focused through a 10x objective (N.A. 0.25). Spectra were recorded using a diffraction grating ruled with a grating of 300 lines/mm (for 785nm and 660nm) and 600 lines/mm (for 532nm and 473nm) giving a spectral resolution of  $\sim 2.1\text{cm}^{-1}$ . Each spectrum was acquired in the region of 400–1800  $\text{cm}^{-1}$ . Multiple spectra were recorded from each individual sample using 785nm and 532nm laser lines at various acquisition times as shown in Table 2. Ten spectra per sample were recorded except for the acquisition time of 20 seconds and 2 accumulations using the 785 nm laser line whereby only 6 spectra per sample were recorded to avoid inconsistent spectra due to the drying of liquid plasma.

Approximately 1-2mg of lyophilized  $\beta$ -carotene was deposited on calcium fluoride slide and Raman spectra were recorded with both 532 and 785nm laser lines. The laser was focused through a 50x objective (N.A. 0.55) using a diffraction grating ruled with a grating of 600 lines/mm for 532nm and 300 lines/mm for 785nm. Each spectrum was acquired in the region of 600–1800  $\text{cm}^{-1}$ . Five spectra per sample were recorded with an acquisition time of 5 seconds and 3 accumulations.

## Data analysis

The raw Raman spectra were baseline-corrected (Rubber band), smoothed (Savitzky–Golay, polynomial order 5, window width 15) and vector-normalized. These processed spectra were used for spectral comparisons and for computing the difference spectra. The mean and difference spectra were also computed to illustrate the intra-group heterogeneity. The spectra were also subjected to principal component analysis-linear discriminant analysis (PCA-LDA).

In brief, PCA is a routinely used method for multivariate data compression and visualization. It describes data variance by identifying a new set of orthogonal features, called principal components (PCs) or factors<sup>24</sup>. In LDA, the discriminating plane is identified by maximising the ratio of between class variance and within class variance. LDA can be used in conjunction with PCA (PCA-LDA). PCA-LDA is a supervised method of using unsupervised PCA scores to define a linearly discriminating plane for classifying data sets using the feature space of the principal components. The advantage of doing this with Raman spectra is ultimately that the approach removes noise and covariance from the data and the generation of independent features that may be important for classification. PCA-LDA models were optimised using a leave-one-out-cross-validation (LOOCV), where the choice of the number of latent variables (principal components) was chosen from where the accuracy of classification reached the first maximum and where additional principal components did not contribute to the overall accuracy of the PCA-LDA performance.

LOOCV is a type of rotation estimation, a technique used for assessing the performance of a predictive model with a hypothetical validation set when an explicit validation set is not available. Leave-one-out involves using a single observation from the original sample as the validation data, and the remaining observations as training data. This is repeated such that each observation in the sample is used once as the validation data and all observations are used to form a confusion matrix. Though LOOCV may yield over optimistic results, in most cases it gives a reasonable first estimate of the model performance and is thus normally employed for small datasets<sup>25</sup>. Since multiple spectra were acquired from each sample, data analysis is carried out using a spectrum-wise approach, where each spectrum is treated as an independent sample and all spectra are individually analysed by PC-LDA. Signal to noise ratio (SNR) and classification accuracies of the HT-Raman spectroscopy method were also calculated.



Algorithms for all of these analyses were generated in-house implemented the R-based statistical software (Version 0.99.896)<sup>26, 27, 28</sup>.

## **Results and discussion**

### **Optimisation results**

#### **Laser lines**

Four different laser lines (785nm, 660nm, 532nm and 473nm) were tested for the development of HT-Raman spectroscopy method, and sample spectra acquired with each of the laser lines are shown in Figure 1. Only the 785nm laser line gave a good biochemical signature of liquid plasma samples. Fluorescence was observed when liquid plasma samples were excited with the 660nm laser line and a resonance Raman effect was observed when samples were excited with the 532 and 473nm laser lines. This was also observed in previous studies<sup>19</sup>. Along with the 785nm laser line, the 532nm line was also selected for developing the HT-Raman technology due to the highly intense Raman bands that can be observed in the regions 1155 and 1523cm<sup>-1</sup>. These bands are tentatively assigned to  $\beta$ -carotenoids and are only observed when blood plasma is excited at lower wavelengths<sup>19</sup>. These specific  $\beta$ -carotenoids were previously found to be more highly expressed in healthy controls compared to cancer patients<sup>19</sup>.

#### **Substrates**

The 785nm laser line was used for the optimisation study of various parameters (Figure 2). Various types of glass (glass slide, no. 0 cover glass (0.13 mm thick) and no. 1.5 cover glass (0.19 mm thick)), CaF<sub>2</sub> (1 mm & 0.15 mm thick), aluminium (2mm thick), and polypropylene (1.14mm thick) substrates were tested. Liquid plasma samples on glass slides did not produce any Raman signal when excited with the 785nm laser line. Liquid plasma samples on the cover glass substrates produced similar quality spectra to those observed on the CaF<sub>2</sub> substrates. No significant contribution of glass was observed in the Raman spectra of liquid plasma samples acquired using cover glass substrates. Both no. 0 and 1.5 cover glass substrates produced similar quality spectra despite the difference in thickness. The cover glass bottomed 96 well plate produced better quality spectra compared to aluminium and polypropylene well plates. The SNR of both cover glass (43.6 dB) and CaF<sub>2</sub> (43.4 dB) substrates were relatively similar. These cover glass substrates are inexpensive compared to the standard CaF<sub>2</sub> substrates. To minimise costs, no. 1.5 coverglass bottomed 96 well plate (MatTek, USA) was selected as a substrate for the further development of the HT-Raman spectroscopy method.

### **Microscope geometry and choice of objective lenses**

Liquid plasma spectra were recorded in both upright and inverted geometry. Spectra recorded using an upright geometry had a higher SNR (47.3dB) compared to those measured using an inverted geometry (43.6dB). However, for small volumes of plasma (20  $\mu$ l), difficulties were encountered focussing into the liquid plasma using the upright geometry set up. Thus, the inverted geometry set up was selected for further optimisation.

Various dry and immersion objectives were tested for the HT-Raman method (Figure 3). Immersion objectives gave slightly better quality spectra with higher SNR (43.6dB) compared to dry objectives (43.3dB) due to the increased numerical aperture. However, it was not possible to keep the drop of water or oil between the objective lens and each well consistently positioned across the 96 well plate due to smearing of the drop during movement of the stage from one well to the next. A low magnification 10x (N.A. 0.25) dry objective was selected for this study because it could more easily focus through the glass coverslip into the liquid plasma than a higher magnification objective lens. Use of an inverted geometry together with a low magnification objective lens allowed focussing of the laser into the liquid plasma and collection of the Raman scattered light without any interference from the glass coverslip.

### **Filtration and sample volume**

As this study focused on liquid plasma samples, these samples were filtered using centrifugal filter devices (3K and 10K Amicon Ultra 0.5ml, Merck, Germany) in an attempt to concentrate them<sup>29</sup>. Raman spectra from unfiltered and filtered plasma samples did not show any significant difference in spectral signature nor in spectral quality (Figure 4A) and had similar SNR (43.6dB).

Various plasma volumes of 20 to 300 $\mu$ l (for increments refer to Table 1) were also tested. No significant difference in the spectral information or quality was observed in plasma samples excited with the 785nm laser line in the inverted set up (Figure 4B). This was also confirmed by SNR, all the sample volumes had the same SNR of 43.7dB. As a result, 20  $\mu$ l of liquid plasma was selected for the development of the HT-Raman method to minimise the quantity of plasma required for the spectral acquisition.

## **HT-Raman spectroscopy method**

The parameters that were selected from the optimisation study are listed in Table 3. A motorised XYZ sample stage was used for the development of the HT-Raman spectroscopy method to move the multiwell plate automatically in a horizontal and vertical direction. The motorised stage allows acquisition of spectra from the center of each well throughout the 96 well plate. Multiple spectra were recorded from each sample using the Labspec software. The laser was focused on the bottom of the 96 well plate and using the motorised stage the focus was adjusted into the sample (-550  $\mu\text{m}$  from the bottom of the plate into the sample), allowing acquisition of 10 or more spectra at various depths of plasma volume ( $\pm 10\mu\text{m}$  from the centre of the well). This was used as a method to take multiple spectra from each sample, not necessarily to record spectra from different volumes of the same sample. Focusing the laser into the sample and away from the substrate also avoids the contribution of the glass substrate in the spectra. Overall, this HT-Raman spectroscopy method is capable of automatically recording multiple spectra from each well throughout the 96 well plate. A schematic representation of the HT- Raman spectroscopy method is presented in Figure 5. This in house developed method was employed for rapid screening of liquid plasma samples from healthy volunteers and prostate cancer patients.

## **HT-Raman spectroscopy of liquid plasma samples from healthy volunteers and prostate cancer**

### **Spectral features**

Vector normalised mean and difference spectra of liquid plasma from healthy volunteers and prostate cancer patients are presented in Figure 6. Differences in the form of intensity related variations were observed across these mean spectra. Significant differences were observed in the region around 1155 and 1523  $\text{cm}^{-1}$  between cancer and control spectra. These bands were observed with high intensity in normal spectra compared to mean cancer spectra. To explain the differences between healthy volunteers and prostate cancer patients, difference spectra were computed by subtracting the mean of the spectra from the healthy volunteers from the mean of the spectra from the prostate cancer patients.

Figure 6A shows the mean and difference spectra of liquid plasma samples from healthy volunteers and prostate cancer patients recorded with the 785nm laser line using an integration time of 10 seconds averaged over 2 accumulations. The difference spectrum clearly shows highly intense bands at 1155 and 1523 $\text{cm}^{-1}$ . Figure 6B also shows the mean and difference

spectra of liquid plasma samples from healthy volunteers and prostate cancer patients recorded with the 532nm laser line with an integration time of 1 second averaged over 3 accumulations. Similar discriminating features were found with both the 785nm and 532nm laser lines and the discriminating features were near resonance at 532nm. These discriminating spectral features can be tentatively assigned to  $\beta$  carotenoids<sup>19</sup>. As a means of validating this assignment, spectra from  $\beta$ -carotene were recorded with both 532 and 785nm laser line. The  $\beta$ -carotene spectra resembles closely the difference spectra with high intense bands in the regions 1155 and 1523 $\text{cm}^{-1}$  (figure 7). The difference spectra in this study can therefore be tentatively assigned to  $\beta$  carotenoids or other related molecules.

PCA was also performed on the vector normalised spectra of cancer and control groups. Figure 8 shows the scatter plot of the PCA classification of cancer and control group spectra acquired with the 532nm laser line. The first two principal components (PC) were used to visualise the classification between the groups. Two minimally overlapping clusters for cancer and normal groups were observed. The PC's showed a high correlation to that of the difference spectra shown in Figure 6. PCA showed that the two  $\beta$  carotene related peaks had differences between the control and cancer group. As mentioned earlier, these carotenoid levels are often found to be higher in the control group compared to the group of patients with cancer

### **Classification efficiency of HT-Raman using 532 and 785nm laser lines**

Blood plasma samples are colloidal dispersions, and can demonstrate heterogeneous behaviour in a micro-probing area, possibly because of the non-uniform distribution of the plasma constituents in the small volume under consideration. Thus, 10 spectra were recorded from each sample and the data were analysed using spectrum-wise approaches to investigate the uniformity of the distribution of the plasma analytes. The spectrum-wise approach, where each spectrum is treated as an independent entity, was adopted to explore if sample heterogeneity can influence the classification between control and cancer samples.

The preprocessed spectra were subjected to supervised PCA-LDA, using the spectrum wise approach, and validated using leave-one-spectrum-out analysis (as described in materials & methods). The results of PCA-LDA are depicted in the form of a scatter plot and a confusion matrix. The confusion matrix is generated to understand the discrimination between the groups obtained by accounting for the contribution of all factors selected for analysis.

The vector-normalized spectra from the control and cancer groups were imported into R software for processing using in-house developed algorithms and subsequent classification using PCA-LDA. Figure 9A shows scatter plot of control (n=10) and cancer (n=10) plasma spectra recorded with 785nm laser line. An LDA classifier was built using the first five principal components and the LDA scores from the classification are shown in the scatter plot. Two well differentiated clusters for control and cancer groups were observed. As shown in the confusion matrix in Table 4A, 58/60 cancer spectra and 57/60 control spectra were correctly classified with a sensitivity and specificity of 96.5% and 95% respectively.

Figure 9B shows the scatter plot of control (n=10) and cancer (n=10) plasma spectra recorded with 532nm laser line. An LDA classifier was built using the first two principal components and the LDA scores from the classification are shown in the scatter plot. Two well differentiated clusters for control and cancer groups were observed. As shown in the confusion matrix in Table 4B, 80/100 cancer spectra and 100/100 control spectra were correctly classified with a sensitivity and specificity of 80% and 100% respectively.

As multiple spectra were recorded with two different laser lines at various acquisition times, the classification accuracy of the two laser lines was compared in figure 10A. Classification accuracies were determined by the sum of the diagonal elements of the confusion matrix divided by total sum of all elements of the confusion matrix. The accuracy of the classification increased with increasing acquisition time for the 785nm laser line but the classification accuracy remained the same after 1 second for the 532nm laser line. The classification accuracies of 785 and 532nm laser lines at different acquisition times are shown in Table 5. Figure 10B shows the comparison of SNR values for the 785 and 532nm laser lines. SNR increased with increasing acquisition time.

## **Conclusion**

In the study, a high throughput methodology has been developed that can screen prostate cancer patients with a higher sensitivity and specificity than the conventional PSA test. Spectral differences are mainly observed in the region around 1155 and 1523  $\text{cm}^{-1}$  across the analysed groups. These bands were highly expressed in healthy volunteers compared to prostate cancer

patients. Control versus cancer cases could be classified with similar discriminating features using both 785 and 532nm laser lines and the discriminating features were resonant at 532nm. The discriminating feature in this study is tentatively assigned to  $\beta$  carotenoids or other related molecules. Carotenoids have been shown to inhibit cancerous changes in several organs including mammary gland, lung, liver and colon <sup>30</sup>. It has been previously reported that the amount of  $\beta$  carotene present in blood decreases during cancer conditions. Reports suggest that  $\beta$  carotene levels are significantly reduced in blood plasma of women with histopathologically diagnosed cervical dysplasia and an inverse association between  $\beta$  carotene plasma levels and increasingly severe graded cervical histopathology <sup>31</sup>. Epidemiological studies have suggested that high endogenous levels of pro-oxidants and deficiencies in levels of antioxidants are likely to be an important risk in the progression of pre-cancer to cancer <sup>32</sup>. An increased breast cancer risk was also observed in subjects with lower levels of  $\beta$  carotene <sup>20</sup>. Recent studies also reported lower levels of  $\beta$  carotene in the serum of oral cancers, buccal mucosa cancer, tongue cancer and cervical cancer as compared to healthy controls <sup>21,33</sup>. The higher levels of  $\beta$  carotene in the healthy controls compared to cancer patients observed in this study are in concordance with the existing reports. However, further studies are required to confirm the precise molecule(s) that differentiate control and cancer cohorts. PCA-LDA findings indicate the possibility of classifying the prostate cancer patients from healthy volunteers with a sensitivity of 96.5% (at 20 seconds and 2 accumulations) and 80% (at 0.5 seconds and 2 accumulations) using the 785 and 532nm laser lines respectively.

Classification accuracy and SNR increased with increasing acquisition times for the 785nm laser line but classification accuracy remained the same after 1 second for the 532nm laser line. A classification accuracy of 95.8% could be achieved with an acquisition time of 40 seconds (20s with 2 integrations) using the 785nm laser line but a classification accuracy of 90% could be achieved with an acquisition time of 1 second using the 532nm laser line.

It is acknowledged that a limitation of the present study is that age matched controls were not used. However, the main objective of this study was to develop a HT Raman spectroscopy method for rapid screening of liquid plasma samples and the control and cancer samples were used to illustrate this method. Future work will further investigate the discrimination of non cancer controls and cancer cases using age matched samples. As the data set is small, LOOCV was used but for future validation a larger sample set would be required.

The cost of substrates is one of the factors limiting the clinical translation of Raman spectroscopy. Raman spectroscopy uses expensive substrates (such as CaF<sub>2</sub>) for analysis of biological samples making Raman based diagnostics more expensive than existing methods. Moreover these substrates cannot be reused in a clinical setting. Identifying suitable inexpensive substrates will help to overcome this limitation. This study has revealed that inexpensive cover glass substrates can be used rather than more expensive alternatives.

HT-Raman spectroscopy of blood plasma may be a more practical and patient friendly approach for screening of prostate cancer patients. This technology could produce a result within hours of patient admission as in total, 96 blood plasma samples could be analysed in less than one hour using this method. Future work will be based on validating the results of this HT-Raman study on a larger sample size with appropriate age matched controls.

### **Acknowledgement:**

This research was funded by the Dublin Institute of Technology Fiosraigh Fee and Materials Support Award and Science Foundation Ireland (11/RFP.1/BMT/3317).

### **References**

1. Ferlay, J., I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman & F. Bray, Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*(2015), 136, E359-86.
2. Sarkar, S. & S. Das, A Review of Imaging Methods for Prostate Cancer Detection. *Biomed Eng Comput Biol* (2016), 7, 1-15.
3. Adhyam, M. & A. K. Gupta, A Review on the Clinical Utility of PSA in Cancer Prostate. *Indian J Surg Oncol* (2012), 3, 120-9.
4. Lyng, F. M., D. Traynor, I. R. Ramos, F. Bonnier & H. J. Byrne, Raman spectroscopy for screening and diagnosis of cervical cancer(2015). *Anal Bioanal Chem*, 407, 8279-89.
5. Jermyn, M., K. Mok, J. Mercier, J. Desroches, J. Pichette, K. Saint-Arnaud, L. Bernstein, M. C. Guiot, K. Petrecca & F. Leblond, Intraoperative brain cancer detection with Raman spectroscopy in humans. *Sci Transl Med* (2015), 7, 274ra19.
6. Theophilou, G., K. M. Lima, M. Briggs, P. L. Martin-Hirsch, H. F. Stringfellow & F. L. Martin, A biospectroscopic analysis of human prostate tissue obtained from different time periods points to a trans-generational alteration in spectral phenotype. *Sci Rep* (2015), 5, 13465.
7. Lacombe, C., V. Untereiner, C. Gobinet, M. Zater, G. D. Sockalingum & R. Garnotel Rapid screening of classic galactosemia patients: a proof-of-concept study using high-throughput FTIR analysis of plasma. *Analyst* (2015), 140, 2280-6.

8. Good, D. M., V. Thongboonkerd, J. Novak, J. L. Bascands, J. P. Schanstra, J. J. Coon, A. Dominiczak & H. Mischak, Body fluid proteomics for biomarker discovery: lessons from the past hold the key to success in the future. *J Proteome Res* (2007), 6, 4549-55.
9. Mainreck, N., S. Brézillon, G. D. Sockalingum, F. X. Maquart, M. Manfait & Y. Wegrowski, Rapid characterization of glycosaminoglycans using a combined approach by infrared and Raman microspectroscopies. *J Pharm Sci* (2011), 100, 441-50.
10. Fernandez, D. C., R. Bhargava, S. M. Hewitt & I. W. Levin, Infrared spectroscopic imaging for histopathologic recognition. *Nat Biotechnol* (2005), 23, 469-74.
11. Zhang, X., G. Thiéfin, C. Gobinet, V. Untereiner, I. Taleb, B. Bernard-Chabert, A. Heurgué, C. Truntzer, P. Ducoroy, P. Hillon & G. D. Sockalingum, Profiling serologic biomarkers in cirrhotic patients via high-throughput Fourier transform infrared spectroscopy: toward a new diagnostic tool of hepatocellular carcinoma. *Transl Res* (2013), 162, 279-86.
12. Carmona, P., M. Molina, M. Calero, F. Bermejo-Pareja, P. Martínez-Martín & A. Toledano, Discrimination analysis of blood plasma associated with Alzheimer's disease using vibrational spectroscopy. *J Alzheimers Dis* (2013), 34, 911-20.
13. Scaglia, E., G. D. Sockalingum, J. Schmitt, C. Gobinet, N. Schneider, M. Manfait & G. Thiéfin, Noninvasive assessment of hepatic fibrosis in patients with chronic hepatitis C using serum Fourier transform infrared spectroscopy. *Anal Bioanal Chem* (2011), 401, 2919-25.
14. Baker, M. J., J. Trevisan, P. Bassan, R. Bhargava, H. J. Butler, K. M. Dorling, P. R. Fielden, S. W. Fogarty, N. J. Fullwood, K. A. Heys, C. Hughes, P. Lasch, P. L. Martin-Hirsch, B. Obinaju, G. D. Sockalingum, J. Sulé-Suso, R. J. Strong, M. J. Walsh, B. R. Wood, P. Gardner & F. L. Martin, Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc* (2014), 9, 1771-91.
15. Keller, M., Kanter E. & Mahadevan A, Raman spectroscopy for cancer diagnosis. *Spectroscopy* (2006), 21, 33-41.
16. Clemens, G., J. R. Hands, K. M. Dorling & M. J. Baker, Vibrational spectroscopic methods for cytology and cellular research. *Analyst* (2014), 139, 4411-44.
17. Huang, W., Li, M. & Jarvis, R, Shining light on the microbial world the application of Raman microspectroscopy. *Advances in applied microbiology* (2010), 70, 153-186.
18. Harris, A., Lungari, A. & Needham, C, Potensial for Raman spectroscopy to provide cancer screening using peripheral blood sample. *Head and neck oncology* (2019), 1,34.
19. Abramczyk, H., Placek, I., Brozek-Pluska, B., Kurczewski, K., Morawiec, Z., & Tazbir, M. Human breast tissue cancer diagnosis by Raman spectroscopy. *Spectroscopy* (2008), 22, 113-121. <https://doi.org/10.3233/SPE-2008-0337>.
20. Pichardo, JL., Frausto, C, Barbosa, O, Huerta, R, Gonzalez, JL, Ramirez, CA, Guitierrez, G & Medina, C, Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients. *Laser in medical science* (2007), 22, 229-236.
21. Gonzalez, A., Martinez, L, Torres, L, Aguilar, Jave-Suarez, L, A & Palomares, P, Cervical cancer detection based on serum sample Raman spectroscopy. *Lasers in medical science* (2014), 29, 979-985.
22. Poon, K. W., F. M. Lyng, P. Knief, O. Howe, A. D. Meade, J. F. Curtin, H. J. Byrne & J. Vaughan, Quantitative reagent-free detection of fibrinogen levels in human blood plasma using Raman spectroscopy. *Analyst* (2012), 137, 1807-14.
23. Filik, J. & N. Stone, Analysis of human tear fluid by Raman spectroscopy. *Anal Chim Acta* (2008), 616, 177-84.
24. Krafft, C., G. Steiner, C. Beleites & R. Salzer (2009) Disease recognition by infrared and Raman spectroscopy. *J Biophotonics*, 2, 13-28.



25. Varmuza, K., Filzmoser, P, Introductuion ti multivariate statistical analysis in chemometrics(2009) CRC press, Newyork.
26. Ripley,W. Modern applied statistics with S (2002).
27. Wickham, H. Elegant graphics for data analysis (<http://ggplot2.org>), 2009.
28. Borchers, H. Practical numerical math functions (<http://cran.r-project.org/package=pracma>), 2015.
29. Bonnier, F., F. Petitjean, M. J. Baker & H. J. Byrne (2014) Improved protocols for vibrational spectroscopic analysis of body fluids. *J Biophotonics*, 7, 167-79.
30. Hata, T. R., Scholz, T. A., Ermakov, I. V., McClane, R. W., Khachik, F., Gellermann, W., & Pershing, L. K. Non-invasive Raman spectroscopic detection of carotenoids in human skin. *Journal of Investigative Dermatology* (2000), 115(3), 441–448. <https://doi.org/10.1046/j.1523-1747.2000.00060.x>
31. Palan, P. R., Mikhail, M. S., Basu, J., & Romney, S. L. Plasma levels of antioxidant beta-carotene and alpha-tocopherol in uterine cervix dysplasias and cancer (1991). *Nutrition and Cancer*, 15(1), 13–20. <https://doi.org/10.1080/01635589109514106>
32. L Aravindh , P Jagathesh , S. Shanmugam , Sonali Sarkar , P. Mahesh Kumar, S. R. Estimation of plasma antioxidants beta carotene, vitamin C and vitamin E levels in patients with OSMF and Oral Cancer - Indian population. *International Journal of Biological & Medical Research* (2012), 3(2): 1655.
33. Sahu, A., S. Sawant, H. Mamgain & C. M. Krishna, Raman spectroscopy of serum: an exploratory study for detection of oral cancers. *Analyst* (2013), 138, 4161-7.

## Tables

*Table 1. Parameters tested for the development of HT-Raman spectroscopy method.*

| <b>Parameters</b>            |  |
|------------------------------|--|
| <b>Laser lines</b>           | 785nm, 660nm, 532nm and 473nm  |
| <b>Substrates</b>            | Glass, Calcium fluoride (CaF <sub>2</sub> ), Aluminium and Polypropylene |
| <b>Geometry</b>              | Upright and inverted   |
| <b>Microscope objectives</b> | Dry (10X), water (60X) and oil (60X) immersion objectives.               |
| <b>Filtration</b>            | Unfiltered and filtered plasma (3K and 10K filtration)                   |
| <b>Sample volume</b>         | 20,40,60,80,100,200 and 300µl  |
| <b>Gratings</b>              | 300 and 600 lines/mm   |

|   |  |
|---|--|
| <b>Confocal hole size</b>                   | 100 to 400 $\mu$ m   |
| <b>Acquisition times</b>                    | 0.5, 3,6,10,20 40 and 60 seconds   |
| <b>Autoscan setup (for HT-Raman method)</b> | Motorized XYZ sample stage was tested for the development of HT template |

*Table 2. Different acquisition times tested for the HT-Raman spectroscopy method.*

| <b>Laser lines</b> | <b>Acquisition times</b> | <b>Spectra recorded/sample</b> |
|--------------------|--------------------------|--------------------------------|
| <b>785nm</b>       | <b>3s x 2</b>            | <b>10</b>                      |
|                    | <b>5s x 2</b>            | <b>10</b>                      |
|                    | <b>10s x 2</b>           | <b>10</b>                      |
|                    | <b>20s x 2</b>           | <b>06</b>                      |
| <b>532nm</b>       | <b>0.25s x 2</b>         | <b>10</b>                      |
|                    | <b>0.5s x 2</b>          | <b>10</b>                      |
|                    | <b>1s x 3</b>            | <b>10</b>                      |
|                    | <b>3s x 3</b>            | <b>10</b>                      |

*Table 3. Parameters selected for HT-Raman spectroscopy method.*

| <b>Selected parameters for HT-Raman spectroscopy method</b> |                                      |
|---|--------------------------------------|
| Laser line  | 785 and 532nm                        |
| Substrate   | 96 well plate (cover glass bottomed) |
| Objectives  | X10 Dry                              |
| Geometry  | Inverted                             |
| Sample state  | Unfiltered liquid plasma             |
| Sample volume   | 20 $\mu$ l                           |

*Table 4. Confusion matrix for PCA-LDA after LOOCV of prostate cancer and normal groups (diagonal elements are true positive predictions and ex-diagonal elements are false positives predictions). (A) 785nm laser line. (B) 532nm laser line.*

**A.**

|         | Cancer | Control | Total | Efficiency          |
|---------|--------|---------|-------|---------------------|
| Cancer  | 58     | 2       | 60    | 96.5% (Sensitivity) |
| Control | 3      | 57      | 60    | 95% (Specificity)   |

**B.**

|         | Cancer | Control | Total | Efficiency        |
|---------|--------|---------|-------|-------------------|
| Cancer  | 80     | 20      | 100   | 80% (Sensitivity) |
| Control | 0      | 100     | 100   | 100%(Specificity) |

*Table 5. Classification accuracies of 785 and 532nm laser lines at different acquisition times.*

| <b>Laser lines</b> | <b>Acquisition times</b> | <b>Classification accuracy</b> |
|--------------------|--------------------------|--------------------------------|
| <b>785nm</b>       | <b>3s x 2</b>            | <b>78.5%</b>                   |
|                    | <b>5s x 2</b>            | <b>79.5%</b>                   |
|                    | <b>10s x 2</b>           | <b>85.5%</b>                   |
|                    | <b>20s x 2</b>           | <b>95.8%</b>                   |
| <b>532nm</b>       | <b>0.25s x 2</b>         | <b>89%</b>                     |
|                    | <b>0.5s x 2</b>          | <b>90%</b>                     |
|                    | <b>1s x 3</b>            | <b>90%</b>                     |
|                    | <b>3s x 3</b>            | <b>90%</b>                     |

## Figures

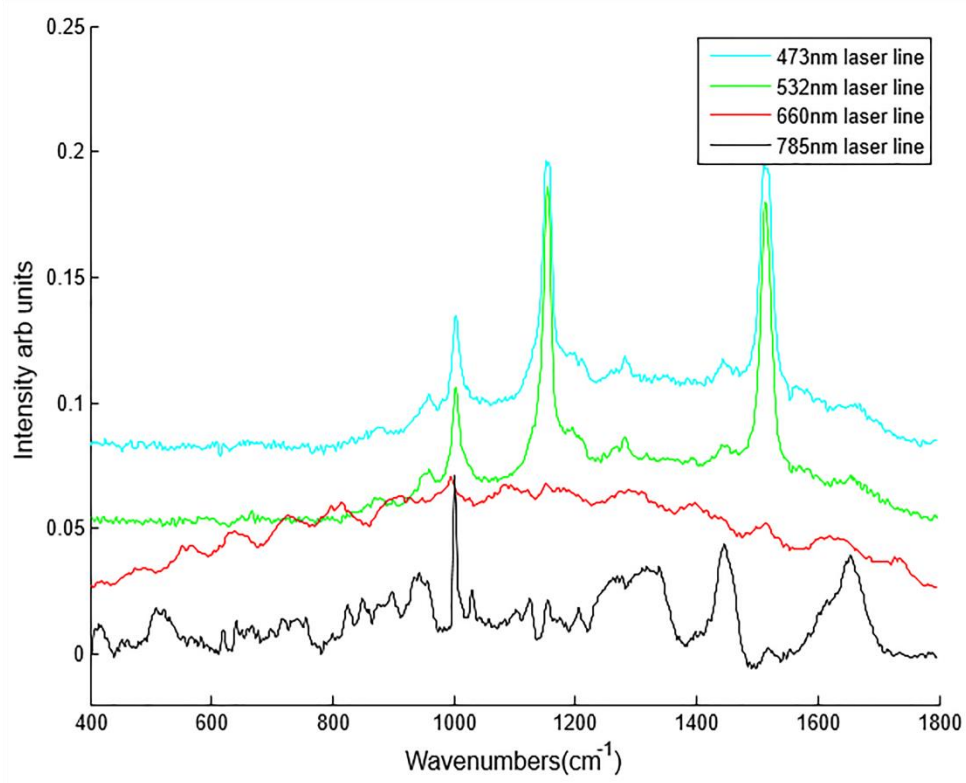
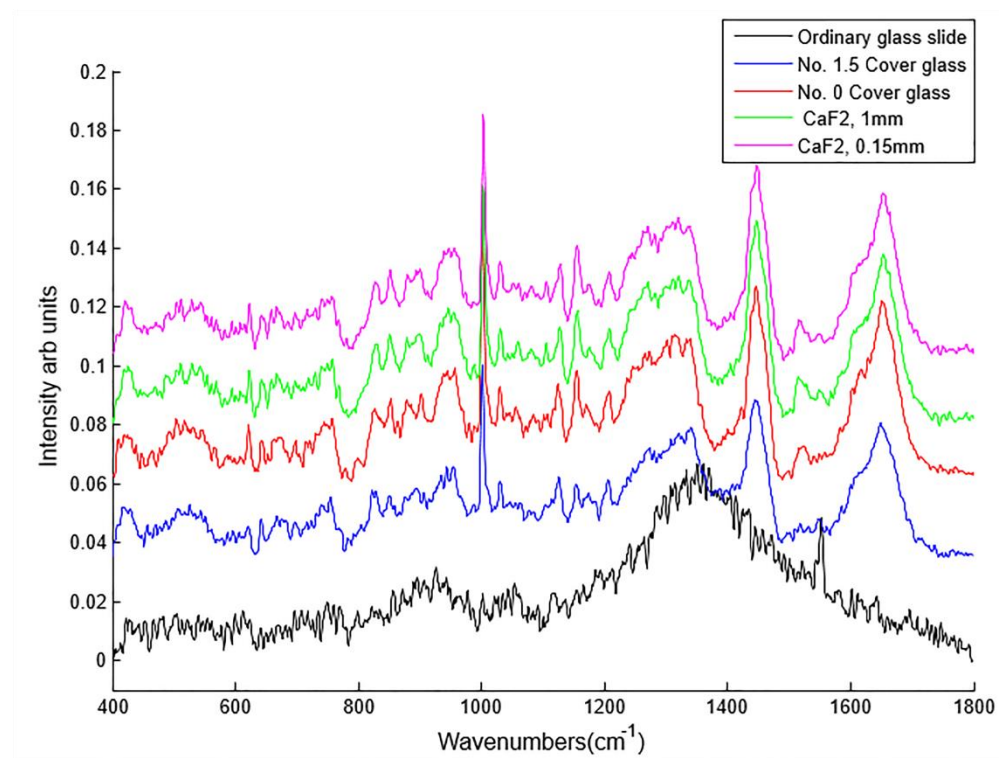


Figure 1. Optimisation of laser lines. Excitation of liquid plasma samples with 785nm laser line produces good spectral information compared to other laser lines.

A.



**B.**

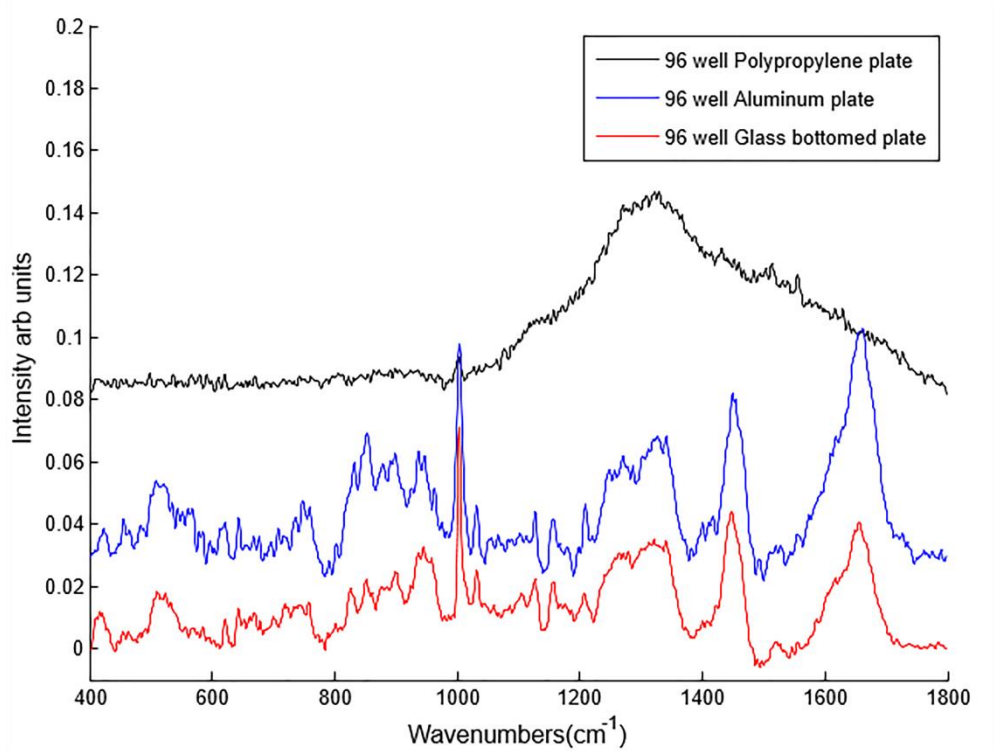
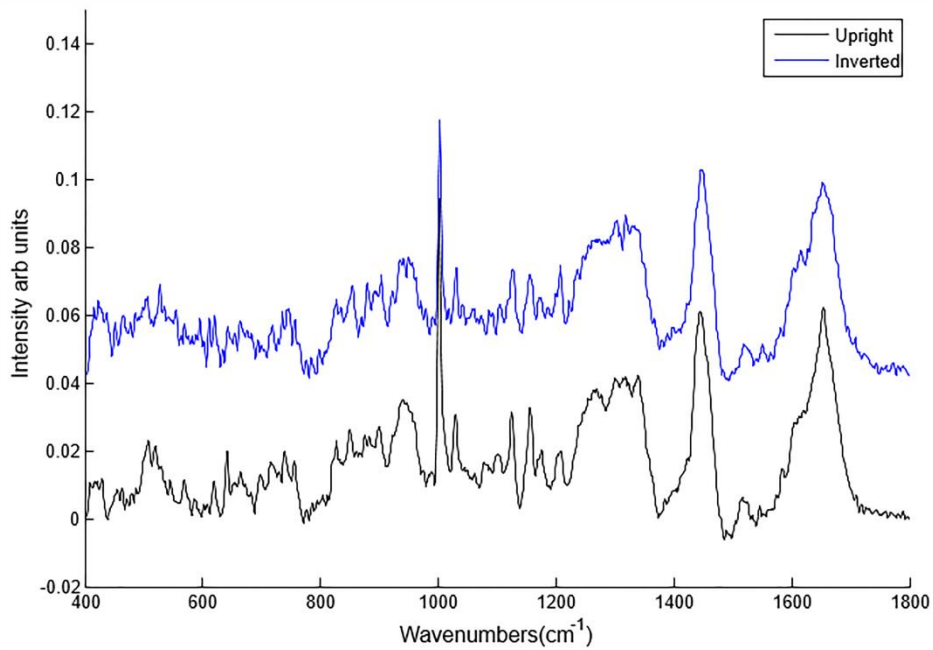


Figure 2. Optimisation of substrates. (A) Optimisation of different grades of glass and CaF<sub>2</sub> substrates. (B) Optimisation of multiwell plates.

**A.**



**B.**

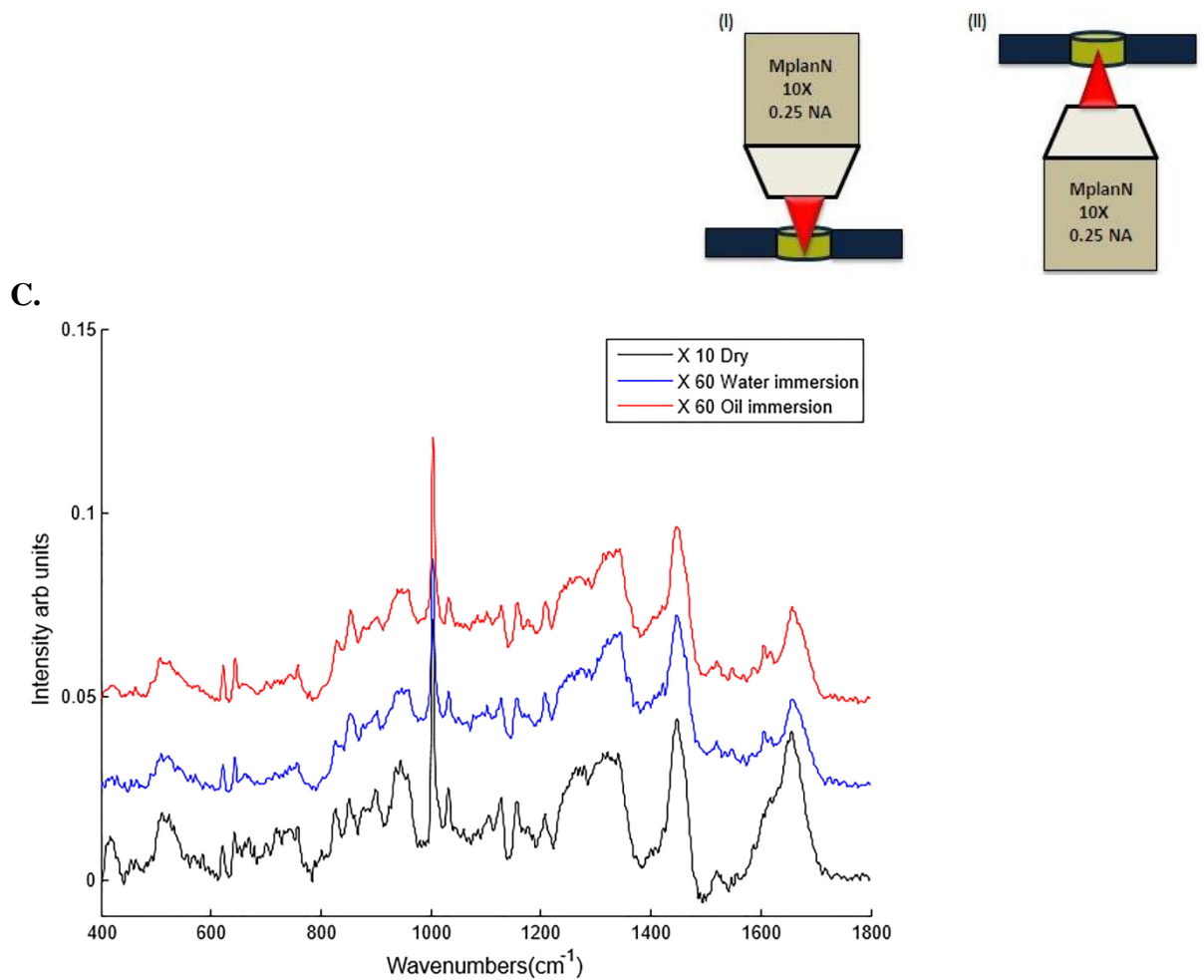


Figure 3. Optimisation study of (A) Geometry (B) Schematic representation of (I) Upright geometry and (II) Inverted geometry (C) Optimisation study of Microscope objectives.

**A.**

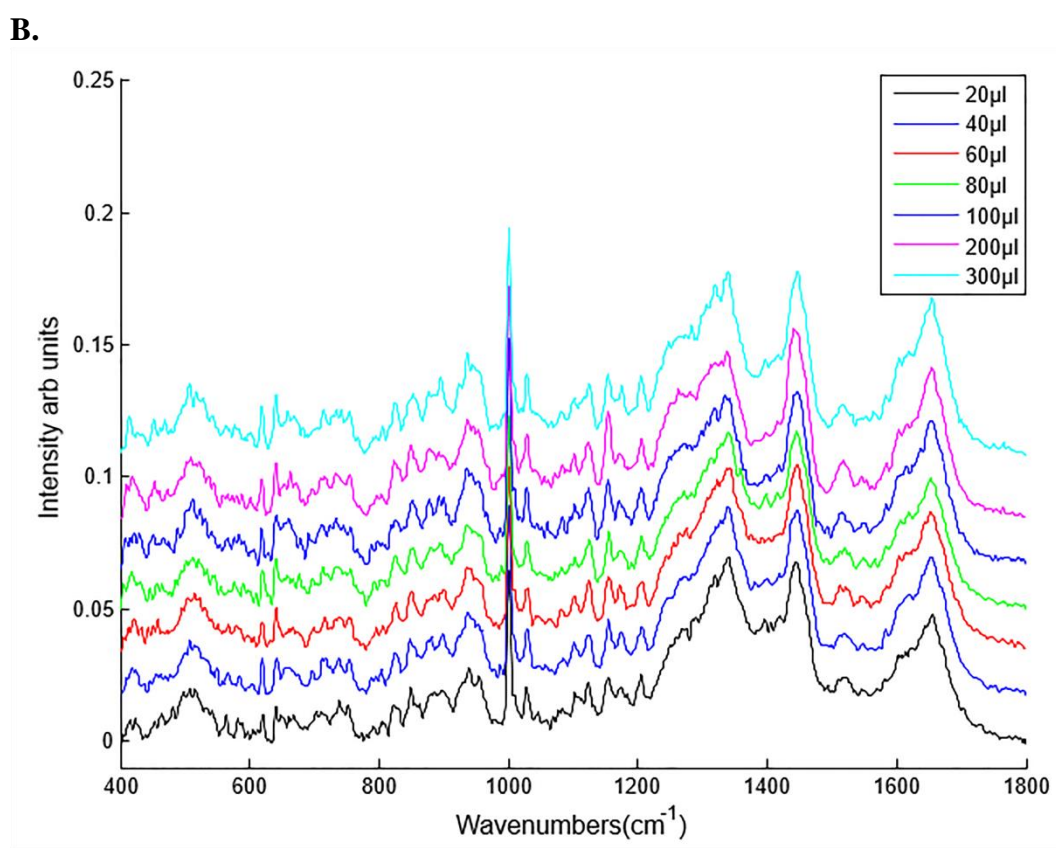
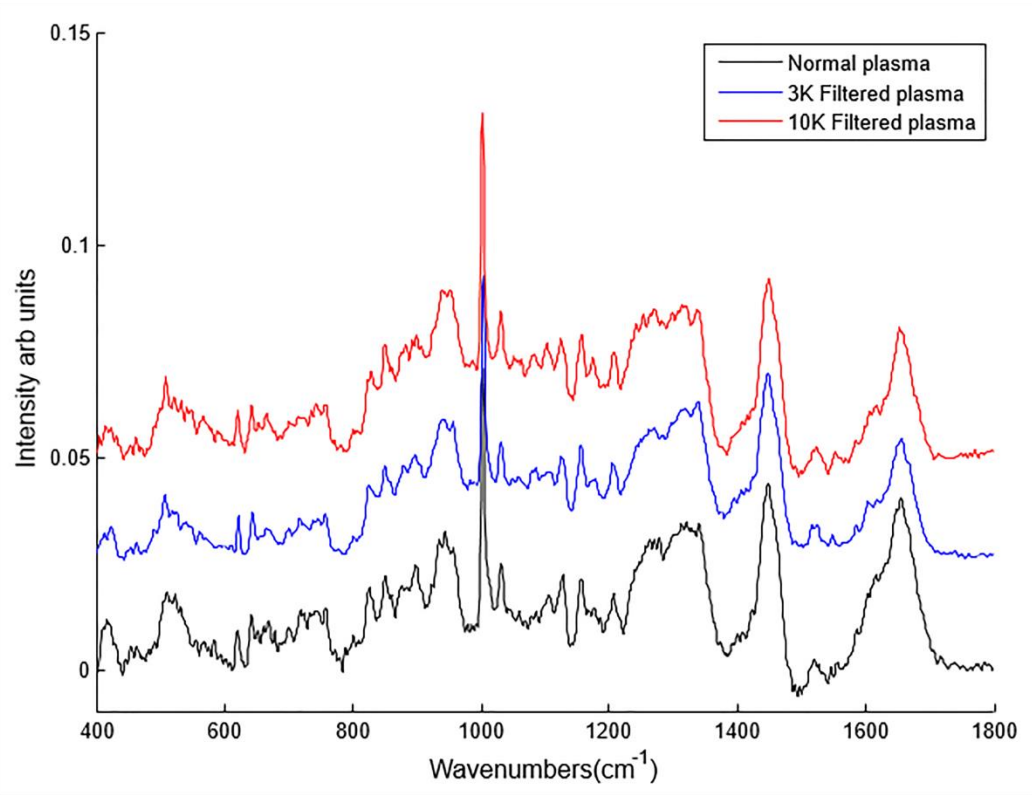
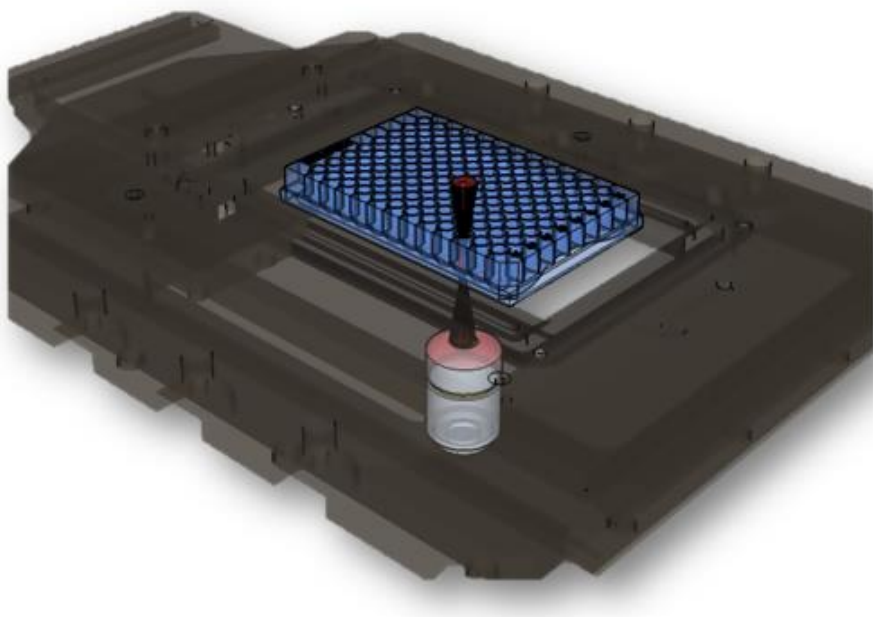


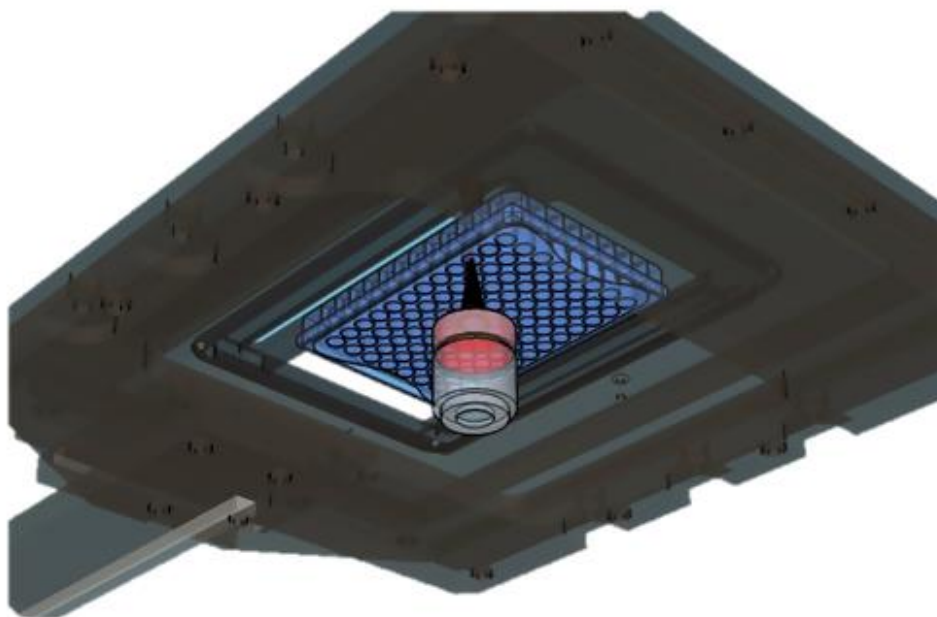
Figure 4. Optimisation study of (A) Filtration of plasma samples and (B) sample volumes.



A.

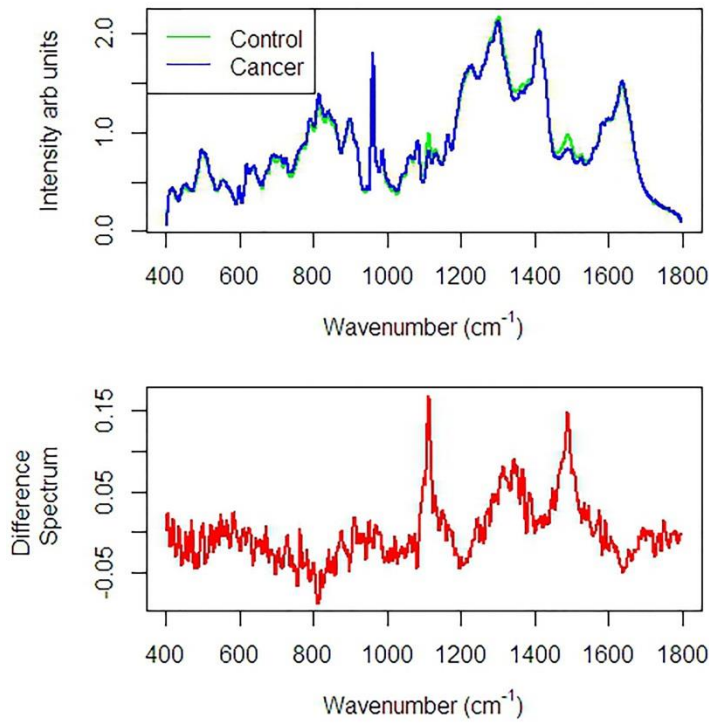


B.



*Figure 5. Schematic representation of HT-Raman spectroscopy method. (A) Top view (B) Bottom view (These schematics are developed using Google Sketch up software).*

A.



**B.**

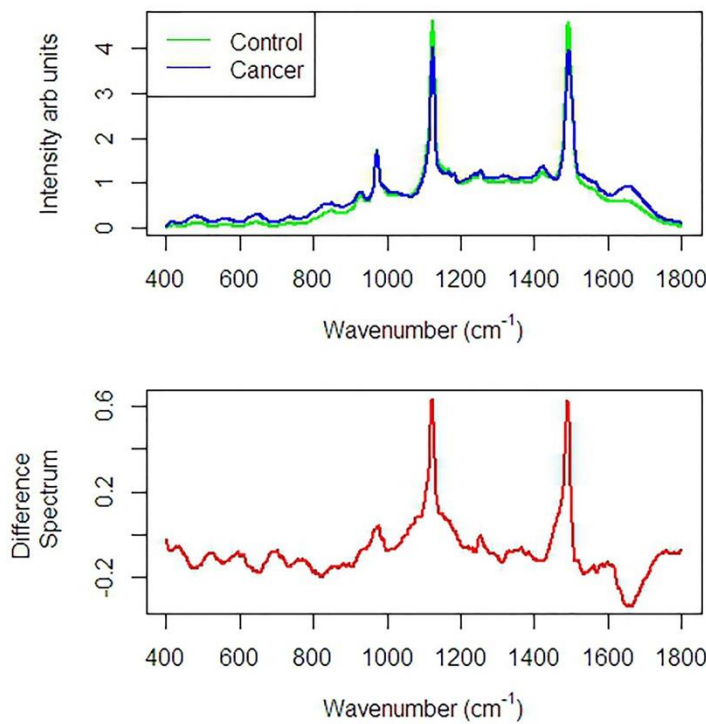


Figure 6. HT-Raman results of healthy control and prostate cancer patient samples. (A) Mean and difference spectra of control and cancer plasma samples acquired with 785nm laser line. (B) Mean and difference spectra of control and cancer plasma samples acquired with 532nm laser line.

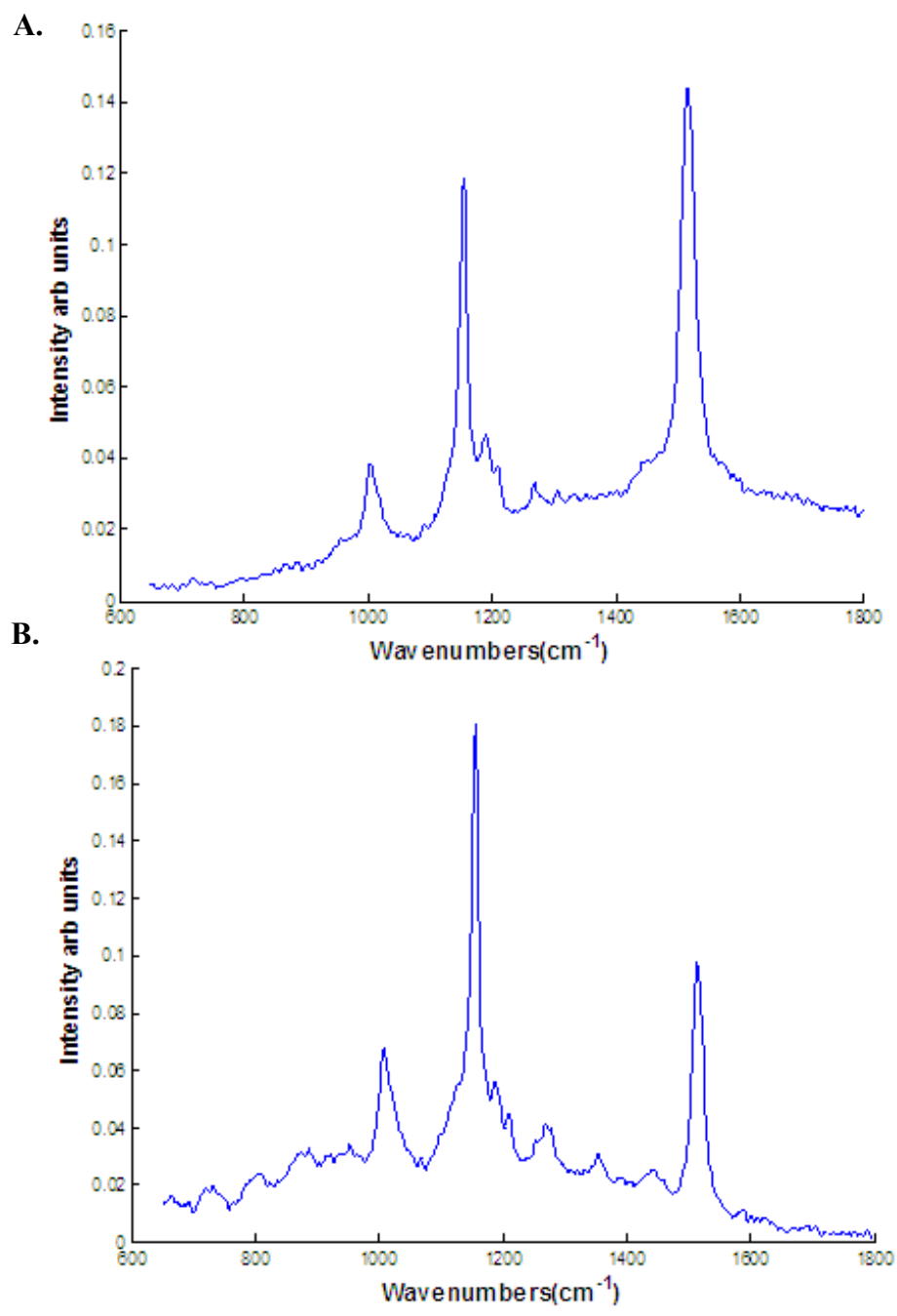


Figure 7: Beta carotene spectra recorded using (A) 532nm laser line (B) 785nm laser line.

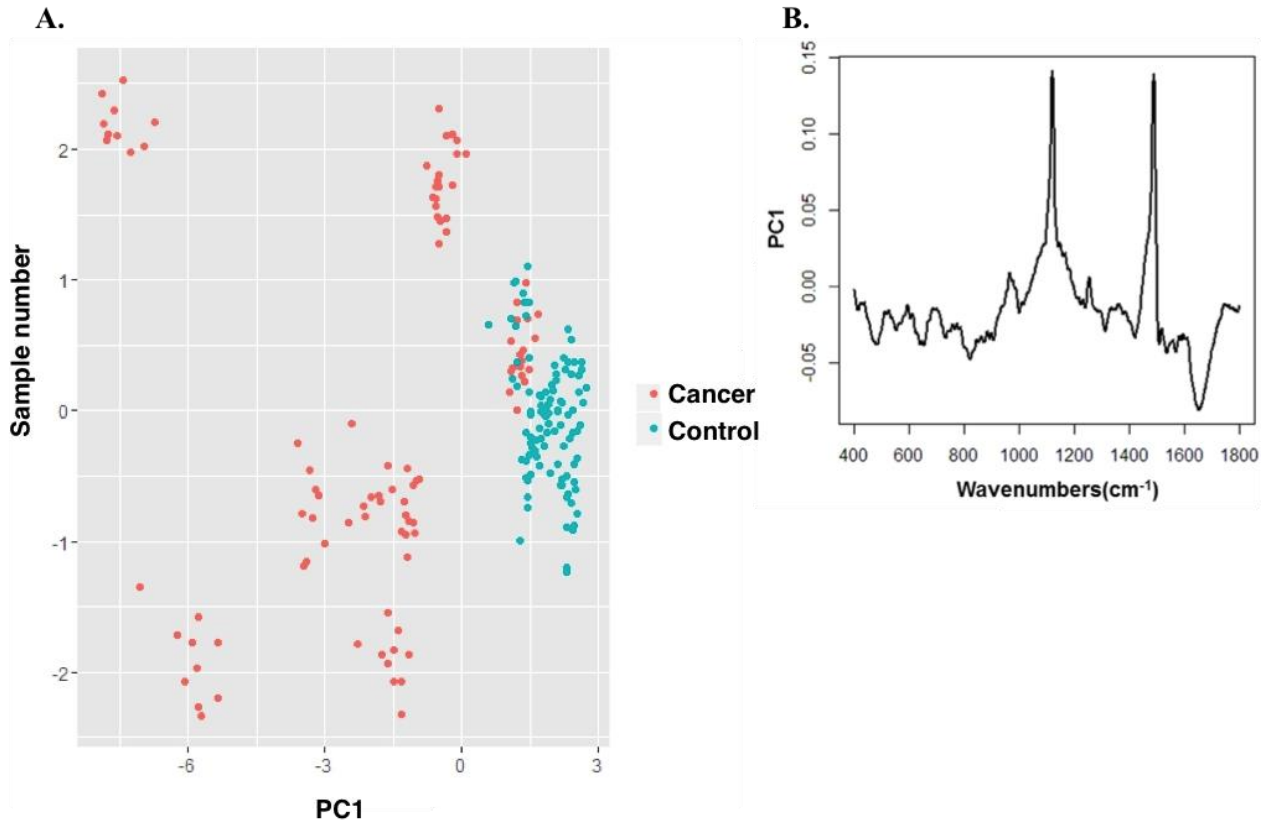
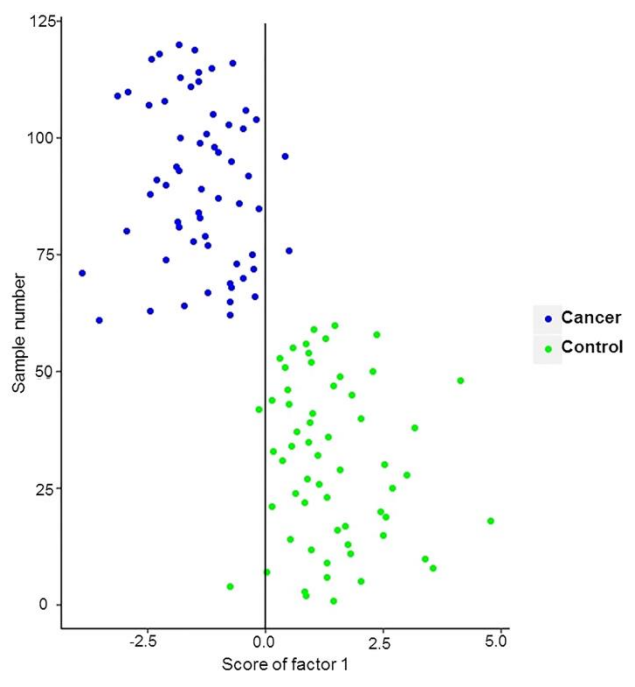


Figure 8: PCA for control and cancer plasma. (A) Scatter plot (B) Principal component 1.

**A.**



**B.**

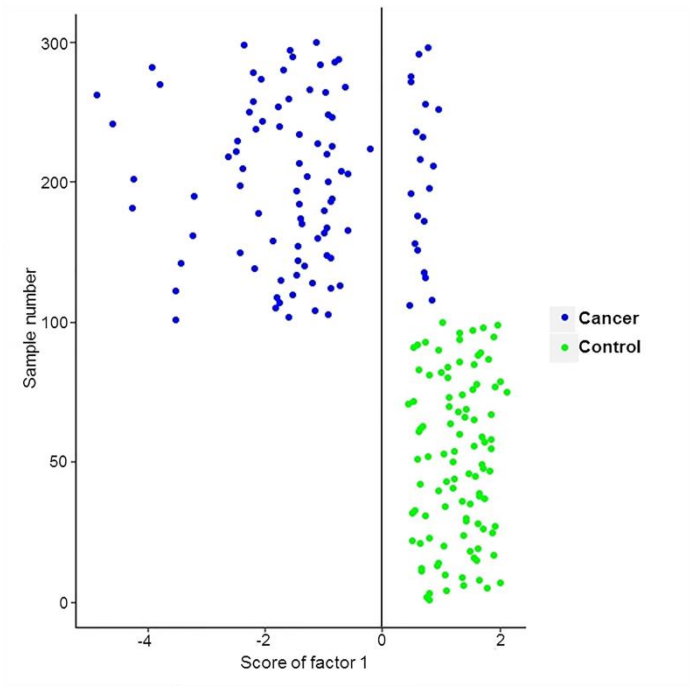
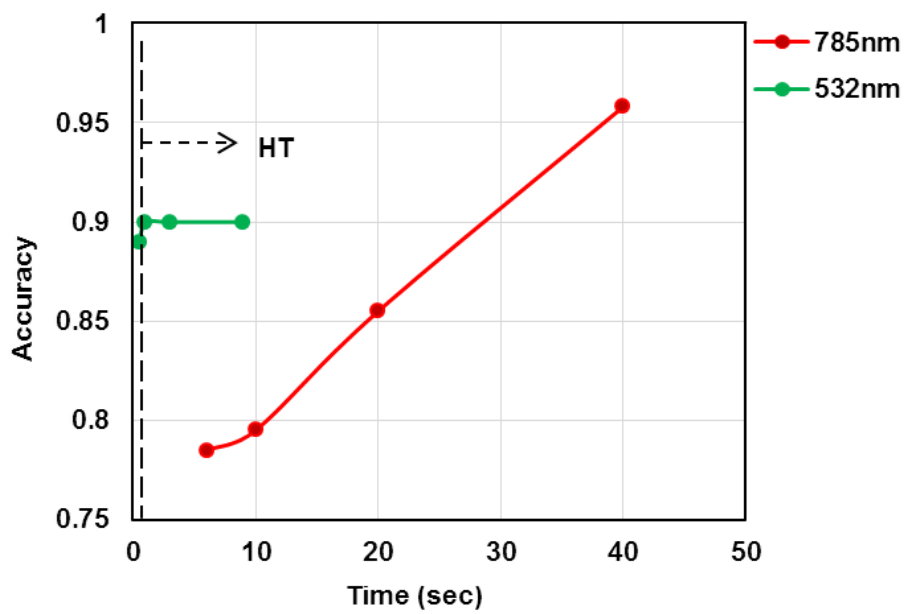


Figure 9. PCA-LDA for healthy control and prostate cancer plasma spectra recorded with (A) 785nm laser line (B) 532nm laser line.

A.



B.

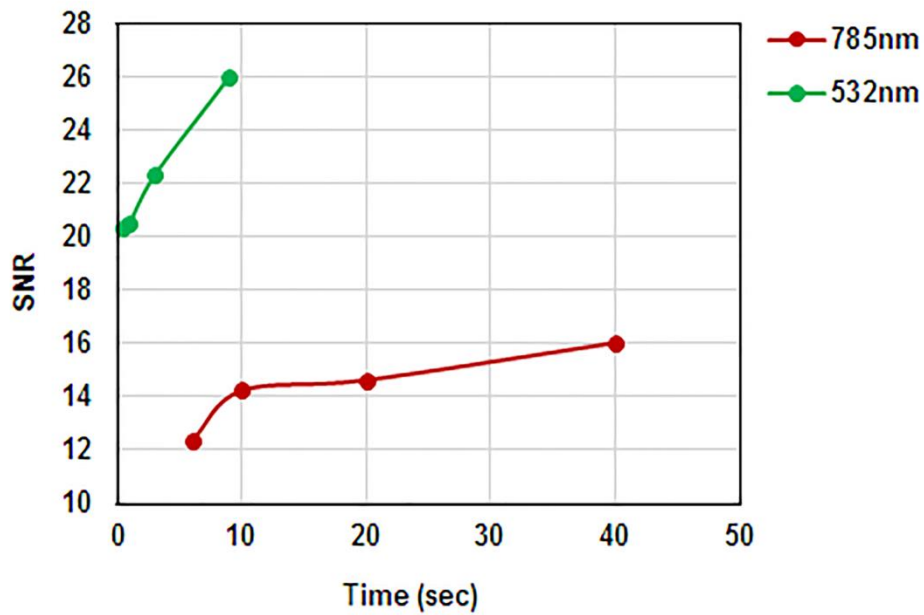


Figure 10. HT-Raman results of healthy control and prostate cancer patient patients. (A) Classification accuracy of 784 and 532nm laser line with increase in acquisition time. (B) SNR of 785 and 532nm laser line with increase in acquisition time.