Development Of Novel Assays For The Optical Detection Of Radiation Exposure

THESIS

Submitted in partial fulfilment of the requirements of

BITS F 421T, Thesis

By

Achint Kumar

ID No.: 2012B5A3325G

Under the supervision of

Dr. Sangeeta Murugkar

Assistant Professor, Carleton University, Canada

BITS PILANI-KK BIRLA GOA CAMPUS

Date: 8/10/2016

CERTIFICATE

This is to certify that the Thesis entitled, Development Of Novel Assays For The Optical Detection Of Radiation Exposure is submitted by Achint Kumar, ID No. 2012B5A3325G in partial fulfilment of the requirements of BITS F421T Thesis embodies the work done by him under my supervision.

Signature of the supervisor

Name: Sangeeta Murgukar

Designation: Assistant Professor,

Carleton University,

Canada

Date: 8/10/2016

Acknowledgement

This thesis is an outcome of help and guidance from various people without whom this work would never have reached its present form. First and foremost, I would like to thank my guide, Prof. Sangeeta Murugkar for introducing me to the very interesting field of Medical Physics. My home institution professor, Dr. P. Nandakumar introduced me to Prof. Murugkar and her work. I thank him deeply for educating and encouraging me throughout my undergraduate years. It was his directions that led me to pursue this project in the first place.

Being inexperienced in the field, I was provided with very helpful advice and kind guidance from my colleagues, Hamid Moradi, Harry Allen and Christopher Dedek. I was also helped by Dr. Nyiri Balazs in data analysis and Dr. Vinita Chouhan with the biology parts of the project.

Finally, I thank my parents and brother for supporting my visit to Canada to pursue the project and for being a constant source of encouragement and upliftment for me. To all people above I express my deep love and gratitude.

Contents

Chapter1

Introduction

Medical procedures, airplane travel, human space explorations all brings humans in contact with higher than ordinary dosages of radiation. The particularly concerning type of radiation experienced by people in these environments is the ionizing radiation which includes the X-rays and the gamma rays. Exposure to ionizing radiation can lead to permanent damage of DNA and thereby leading to cancer or heritable genetic defects. This project employs the use of Raman micro-spectroscopy and multivariate statistical analysis to detect subtle changes in human lens epithelial cells exposed to different dosages of radiation.

The human lens epithelial cells exposed to varying levels of radiation were provided to the laboratory by Health Canada. The author studied the subtle changes brought about by radiation exposure by analysing the spectra of cells using multivariate statistical tools like Principal Component Analysis combined with Linear Discriminant Analysis.

The Chapter 2 includes a discussion of the principles and statistical techniques used in the experiment. In Chapter 3, a discussion of the experimental setup is discussed. In Chapter 4, the actual human lens epithelial cells experiment is discussed. In Chapter 5, a side experiment discussing the contamination in the quartz samples is discussed, Finally, in Chapter 6, future work and extensions of the study is discussed.

It is hoped that this study would help in developing the safety standards while performing medical procedures, adequate shielding standards in airplane and rockets carrying humans to outer space.

Chapter 2

Physical Principles and Data Analysis

The project involved the applications of various physical and statistical analysis components. The following is a brief description of the same as applied to the project:

Raman Spectroscopy: It is based on the principle of Raman scattering which results from the inelastic scattering of light interacting with matter. When laser light falls on a sample most of the light bounces off unchanged in frequency which is a phenomenon known as Rayleigh scattering. However, a small fraction (one in ten million) experiences a change in frequency due to the energy transfer to molecular vibration leading to Raman scattering.

In the project, the cells samples were exposed to laser light and the Raman spectrum of the scattered light was collected to obtain the molecular finger print of the cell. The spectra form the samples exposed to the different dosage levels were then compared to see biochemical difference arising from them.

Principal Component Analysis: The data set collected from the experiment is a collection of intensity spectra obtained from the nucleoli at about 160 different positions in various cells (80 being 0Gy and other 80 being 5Gy). Each nucleoli spectrum is sampled at 1024 different wavenumbers.

These 1024 wavenumbers correspond to the “variables” in our data set and the 160 intensities at each wavenumber correspond to the “observations.” So, this data set can be represented as a 1024 X 160 (variables X observations) matrix. This matrix has two main problems:

1) The data matrix has high dimensionality. So, analysing it is computationally expensive.
2) It might be that many of the variables are collinear (highly correlated), so they don’t contribute any new information.

The Principal Component Analysis(PCA) is a technique to reduce the dimensionality of the data set in such a way that least amount of information is lost. Information in the statistical sense a measure of the variability in the data. So, in PCA we transform our variables in the directions of maximum variability. These directions with maximum variability correspond to the Principal components(PC). We keep only the PC which correspond to high variability and ignore those with low variability. The PC are obtained from the original variables in such a way that the PC are orthogonal to each other so their correlations are zero. The detailed mathematical treatment of PCA can be found in Jollife [1]. The steps in the algorithms applied in the experiment are outlined below:

1) The mean of each variable is subtracted from the observations, so that the new variables are centred, i.e. their mean is zero.
2) Using the new centred variables, the covariance matrix is calculated.
3) The covariance matrix is diagonalized and the eigenvectors corresponding to the eigenvalues are calculated.

The eigenvectors corresponding to the highest eigenvalue is the first principal component(PC1), the second highest eigenvalue gives second principal component(PC2) and so on. This procedure creates a new set of variables, PC1, PC2… which are orthogonal and have maximum variability. Now, we drop the principal components which have low eigenvalues compared to the rest. This procedure has thus led to a lower dimensional data set which is further used in classification.

Linear Discriminant Analysis: The reduced data set obtained is now labelled to observe the features that are different. This procedure helps find the directions in spectral space that maximize the ratio of between-class variance to within-class variance. It is a supervised classification algorithm for discrimination of sample groups. The process involves two steps:

1) Training Phase: In this phase, labelled training data set is used to find pattern in the data
2) Predication Phase: In this phase, the model developed from the previous phase is used to identify an unseen data.

Chapter 3

Experimental Instruments

The experiment was performed on a custom-built Raman confocal Raman micro-spectrometer as shown in Figure 3.1

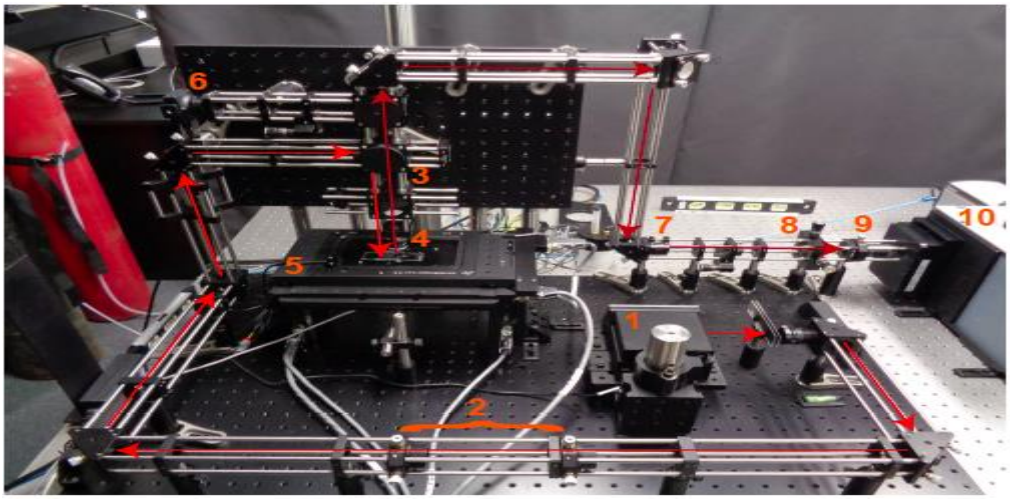


Figure 3.1: Custom confocal microscope coupled to a Raman spectrometer.1) 785nm laser 2)Optics for beam collimation and expansion 3) Dichroic mirror 4) Microscope objective 5) Automated x-y-z stage 6) CCD camera for bright-field imaging 7) Laser Rejection Filter 8) Pinhole 9) Focussing Optics 10) Spectrometer and CCD detector

The different parts of the setup are explained as following:

1) **785nm laser:** The intensity of Raman signal is inversely proportional to the fourth power of the excitation wavelength so shorter wavelengths lead to stronger Raman signal and better signal to noise ratio. However, shorter wavelengths also cause fluorescence in biological samples which have complex organic molecules. This leads to further swamping of the already faint Raman signal. The use of 785nm laser is optimal for Raman spectroscopy as it largely eliminates fluorescence and produces a reasonable SNR for the Raman signal. It also minimizes photo damage to the cells due to its low power.

2) **Beam collimation and expansion:** It is achieved through two lens (a plano-concave lens and another bi-convex lens) in the laser path to ensure that the pupil of the objective is filled by the laser and therefore focused to a tighter spot size.

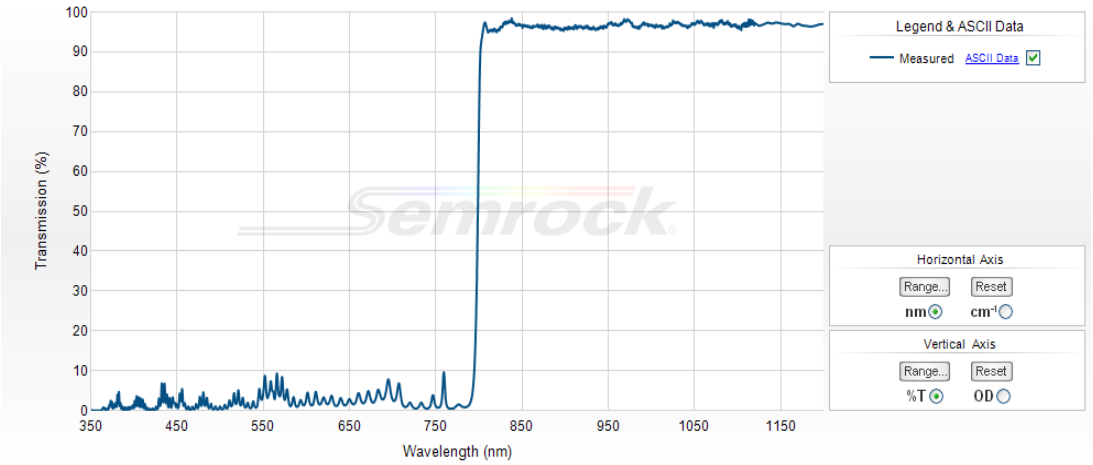
3) **Dichroic mirror:** The laser beam is reflected downward by the dichroic mirror and guided into the objective. The dichroic acts like a long-pass optical filters which in this experiment reflects the 785 nm laser light, but pass light of longer wavelengths with a high transmission efficiency (>90%).

4) **Microscope objective:** A 60X water immersion objective is used in the experiment. The water immersion objective is selected over the oil immersion since it does not contribute to strong Raman peaks in the concerned bandwidth of the experiment. Water immersion also increases the resolution of the microscope since immersing both the lens and the specimen in water (which has higher refractive index than air) helps increase the numerical aperture of the objective lens.

5) **Automated x-y-z stage:** This stage helps to manoeuvre the specimen to bring it to focus and provide movement over different spots in the sample.

6) **CCD camera for bright field imaging:** There is a LED source below the sample which helps identify the spot to be focussed in the sample by producing an image in the CCD camera.

7) **Laser Rejection filter:** This filter which is long pass filter eliminates the laser and Rayleigh scattered light. Only the Raman scattered light passes through.
The transmission efficiency cure of the Dichroic and the laser rejection filter is shown in Figure 3.2:



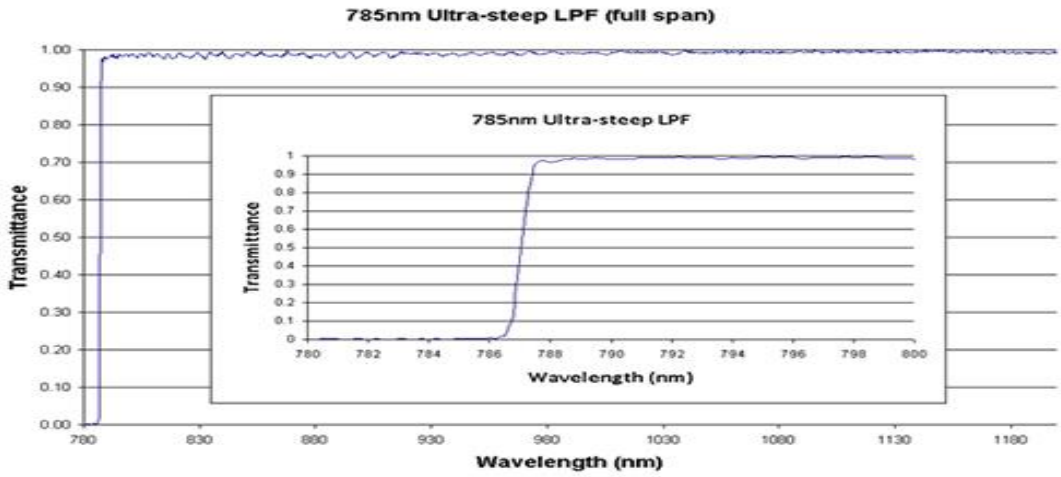


Figure 3.2) Transmission efficiency for a) Dichroic mirror b) Laser rejection filter

The scattered Raman signal follows the path back into the objective and gets transmitted by the dichroic (since its wavelength is larger than 785nm). Following that the light encounters the laser rejection filter which eliminates the laser and Rayleigh scattered light. The last three parts of the setup deal with the collection of the Raman signal.

8) **Pinhole:** This helps bring about the confocal nature of the setup. It blocks the out-of-focus light so that the light originating only from a small region of the sample is observed. Therefore, a confocal microscope can achieve better axial resolution than a conventional microscope and can image a thin section of the sample, referred to as optical sectioning. The smaller the pinhole, the better the discrimination and the thinner the optical section. However, a smaller pinhole yields weaker intensities of light. A compromise must be made between the spatial resolution and the signal intensity. In the experiment pinhole of size 100 µm was selected so as to optimize the signal intensity and resolution.

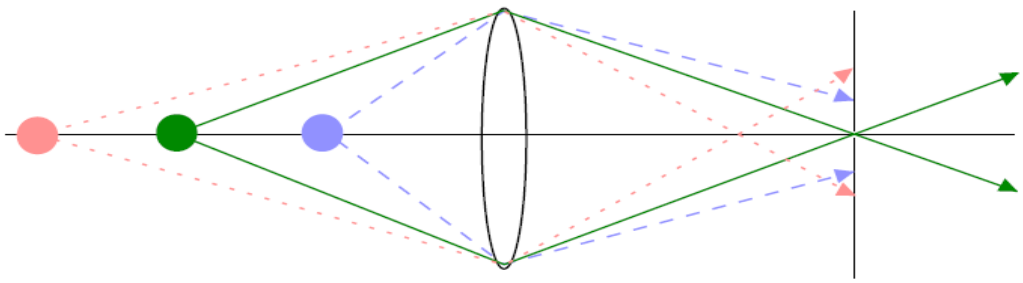


Figure 3.3) Pinhole as optical sectioning device. Green dot is in the focal plane while red and blue dots are not and hence are blocked by the pinhole

9) **Focussing optics:** The coupling of the microscope with the spectrometer involves intermediary optical elements that must match the numerical aperture of the spectrometer

10) **Spectrometer and CCD detector:** The spectrometer used in the experiment was the Shamrock SR-303I with Czerny Turner configuration (See Figure 4.3). It is a fully automated high throughput flat-filed imaging triple grating turret spectrograph. The spectrometer has a f/4 lens placed in front of the spectrometer to ensure the best focussing on the slit. The longer the focal length (e.g., the distance between the dispersing grating and detector) of the spectrometer, the higher the spectral resolution. The focal length of the spectrometer is about 300mm. The spectral resolution, spectral coverage and blaze efficiency dictate the choice of the grating.

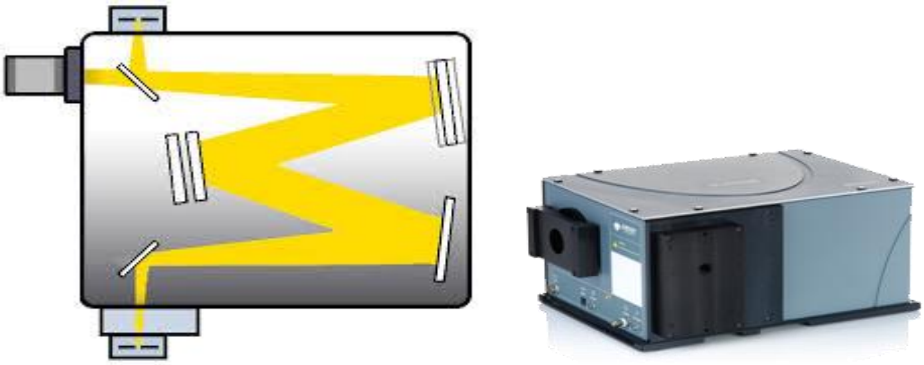


Figure 3.3: Spectrometer

The higher the groove density of the grating (typically measured in grooves per millimetre), the higher the spectral resolution. In the experiment, a groove density of 1000lines/mm with blaze wavelength of 900 nm was used, which is suitable for measurements of Raman spectral lines in the region 750-800 nm.

Chapter 4

Human Lens Epithelial Cells Experiment

Background: Raman spectroscopy has been previously shown to be able to differentiate between cancer and normal ovarian cells. In this experiment, we try to see if it's possible to differentiate human lens epithelial (HLE) cells exposed to 0 Gy and 5Gy radiation. The research opens possibility to early prognosis of eye diseases like cataract by identifying subtle changes in the lens' Raman spectra.

Objective: The experiment envisages to explore the possibility of characterization of 0 and 5Gy HLE cells using the principle of Raman micro-spectroscopy combined with the machine learning tools like PCA-LDA.

Instruments: A custom made confocal Raman micro-spectrometer built at Carleton University was used for collecting and analysing spectra of the HLE cells. The cells were exposed to 0 and 5Gy radiation in a radiator at Health Canada and brought for examination to Carleton University.

Protocol: While collecting the spectra of the cells the following protocol was followed:

1) Before beginning the experiment, measure the spectra of silicon disk to confirm if the alignment is correct.
2) Place the dish in the Raman stage and raise the stage until the objective lens just immerses in the solution. After that raise the stage using the fine-adjustment and continue raising it until the laser is focussed on the cells
3) Once the cells are brought to focus, position the laser focus on any of the nucleolus. Save the picture of the field of view- with the laser spot as well as without the laser spot.
4) Remove the mirror, turn-off the back-illumination LED, remove the filter, turn-off the table light and pull the curtains to completely darken the optic table room.
5) Open the Andor software and start the 10 samples/30 second acquisition time per sample program.
6) Once the spectra has been collected, view the averaged spectra and check for any abnormalities. If some abnormality is observed it can in most cases be removed by confirming if Step-4 has been performed correctly. If the abnormality appears in the form of fluorescence it can be removed by pushing down the stage and bringing it back up again.
7) If the spectra look normal, use the snipping tool to save the spectra as well as the configuration of the setup.
8) Save the file and begin again to take the next set of measurements.

9) Once the experiment is over, measure the spectra of silicon disk again to confirm if the alignment is correct.

This protocol was followed to collect 336 samples. There were 188, 0 Gy samples and 148, 5 Gy samples in total. Taking 75% as training set and 25% as test set, a PCA-LDA analysis was carried out.

Data Analysis:

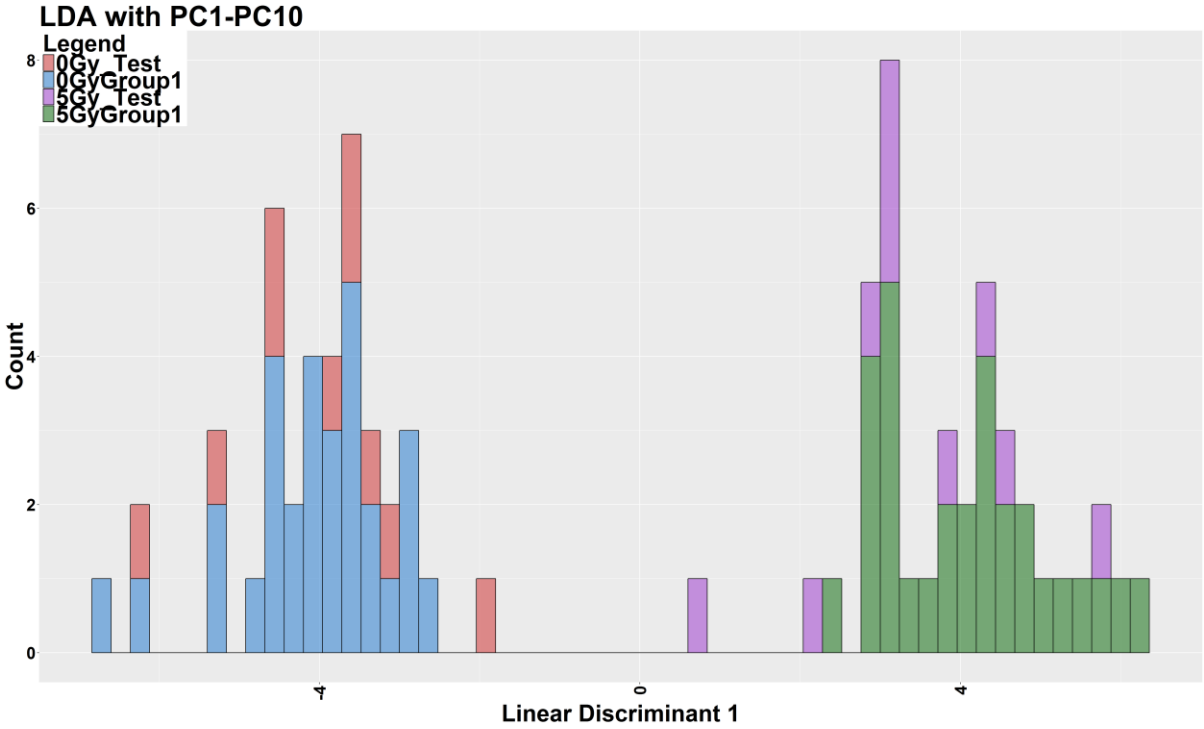
1) **Same batch:** Performing classification on 80 samples collected on 3rd and 4th November showed 100% accuracy. Ten principal components were used and 30 samples of each class was taken as training examples and 10 samples of each class were used as test sets. Following is the result:

LOOCV Results [Training Data]

	0GyGroup1	5GyGroup1
0GyGroup1	30	0
5GyGroup1	0	30

Test Data Classification Results

	0GyGroup1	5GyGroup1
0Gy_Test	10	0
5Gy_Test	0	10



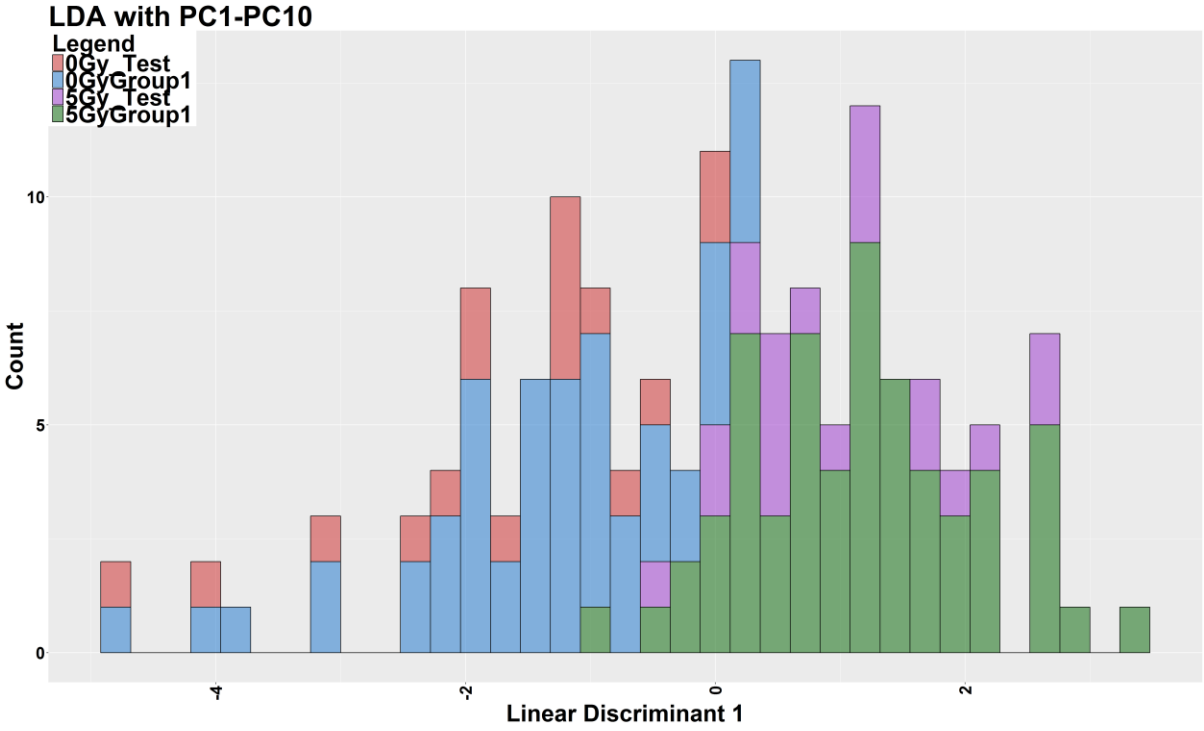
2) **Different batches:** 80 samples from 3-4 November and 70 samples from 10th November was taken into consideration. Ten principal components were used for classification. 52 0Gy cells and 61 5 Gy cells were used as training set while 17 0Gy cells and 20 5 Gy cells were used as test set. An accuracy is found to be 89%

LOOCV Results [Training Data]

	0GyGroup1	5GyGroup1
0GyGroup1	41	11
5GyGroup1	3	58

Test Data Classification Results

	0GyGroup1	5GyGroup1
0Gy_Test	15	2
5Gy_Test	1	19



3) Blind Test: A blind sample containing 40 samples was analysed using 10 PC and 150 training samples (81 of 0 Gy and 69 of 5 Gy) . The accuracy obtained was found to be 80%. Next, the number of principal components were increased to 25 and then the accuracy was found to be 83%.

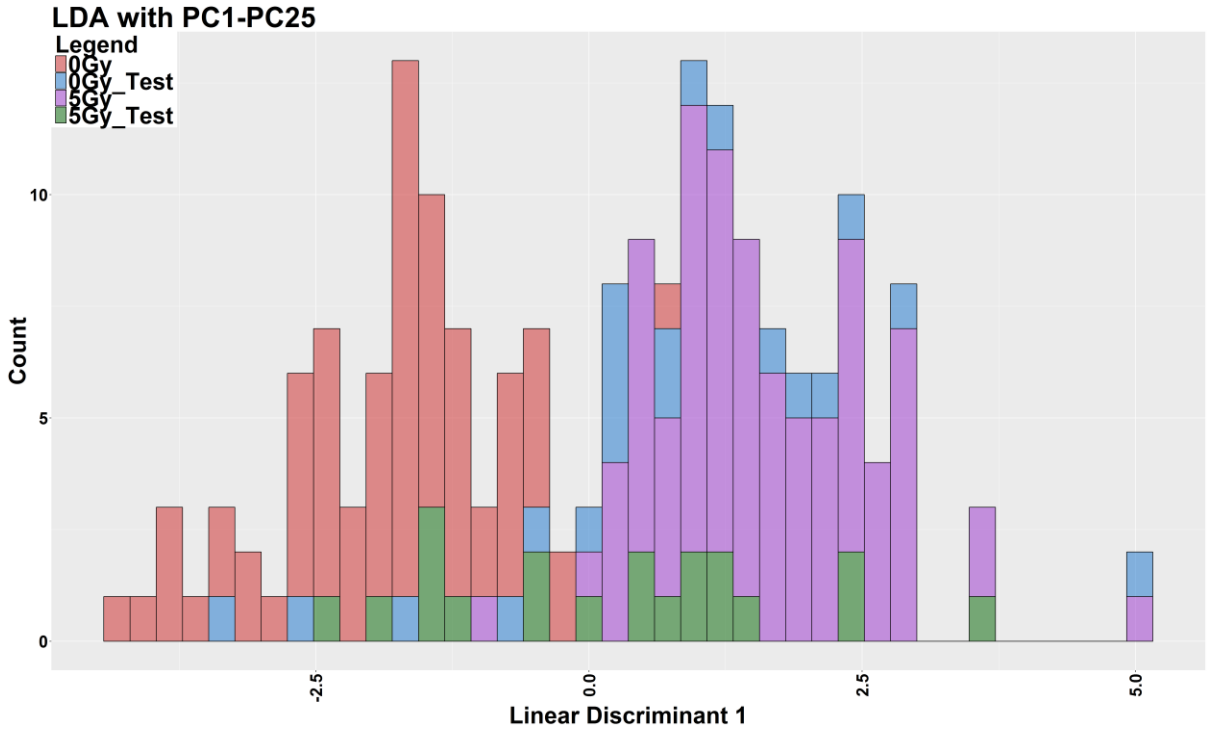
Following is the result with 25 PC and 83% accuracy:

LOOCV Results [Training Data]

	0Gy	5Gy
0Gy	63	6
5Gy	6	75

Test Data Classification Results

	0Gy	5Gy
0Gy_Test	5	15
5Gy_Test	8	12



Conclusion:

The technique of confocal Raman micro-spectroscopy combined with multivariate techniques has been shown to be successful in differentiating cells exposed to different dosages of radiation with an accuracy of 82%.

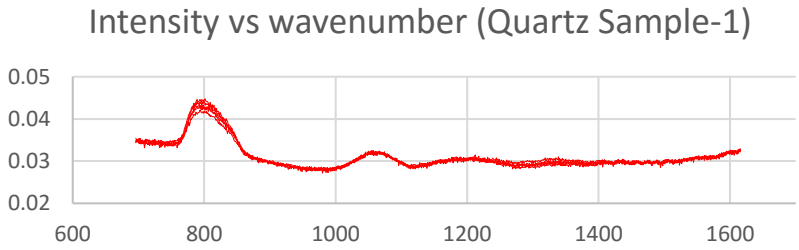
Chapter 5

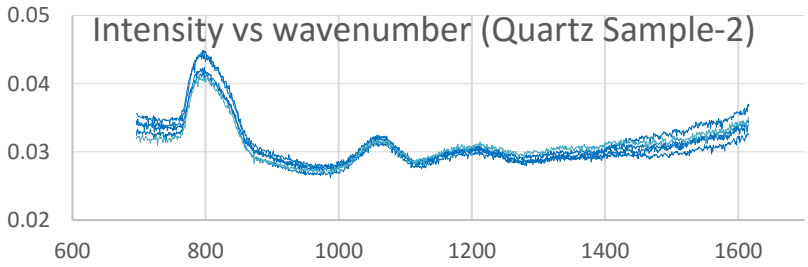
Study of Contamination of Quartz substrates

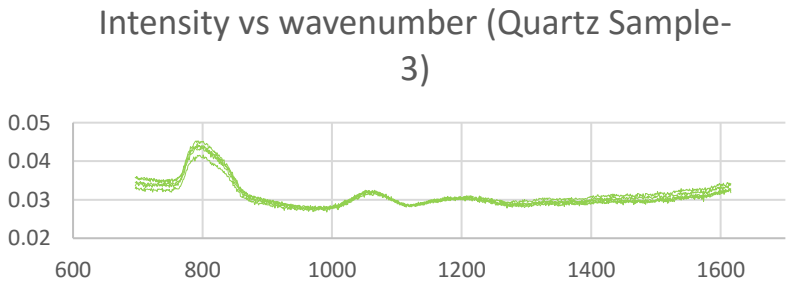
Objective: To ascertain whether the substrate to which the cells would be added at later stage of the experiment are free from contaminants.

Procedure: Three quartz samples were placed under the objective of a Raman spectroscope one after the other and their spectra was collected. For each sample, five spots were chosen and the spectra was obtained ten times for each spot. The spectra were then normalized, plotted and the correlation between the spectra was calculated. Since the samples were of the same type, it was expected that their correlation coefficients would be close to one.

Results: The normalized spectra of the three substrates are plotted below:







The similar peaks in each point of every sample in all the spectra show that the quartz samples are indeed free of contaminants. There are also no point contaminants, which would have led to a sharp peak in the spectra. The correlation between the spectra was calculated to be above 90% in each case thereby confirming our assessment that the substrate is free from contaminants.

Chapter 6

Challenges and Opportunities

1) After characterization, the next step would be to attempt to figure out the exact proteins/amino acids that are undergoing biochemical changes to produce difference in spectra. This can be done comparing the spectra of proteins/amino acids with the spectra of the cells. If the contribution from a particular biochemical is present in cell spectra of one class but not in the other then we can deduce the presence of that protein.
2) The current method of placing the cells in magnetic clamp system poses the problem of replenishing the cells with nutrients periodically at intervals of about half an hour. To resolve this, a route to add the nutrients automatically needs to be found. A possible solution is to use glass bottom dishes which can store comparatively higher amounts of nutrients.
3) The acquisition time is very long. It takes about 4 minutes to obtain spectra at each point of the sample. This time can be reduced by lowering the number of spectra obtained at each point from 10 to 5.
4) The magnetic clamp keeps cracking the delicate and brittle quartz disks. A possible resolution could be to use a different substrate which is won't crack- like fused quartz disks. Another way could be to use dual layer of quartz for added reinforcement.
5) Whenever a new sample is put under the objective, a mirror needs to inserted into the setup to help focus the laser again. Experiment time would significantly reduce if a way to automate this process is found.

References:

[1] I.T. Jolliffe. Principal Component Analysis. Springer-Verlag New York, Inc., New York, N.Y., 2002.

[2] Hecht E," Optics", Addison –Wesley, 2002

[3] Butler, H. J., Ashton, L., Bird, B., Cinque, G., Curtis, K., Dorney, J. Martin, F. L. (2016). Using Raman spectroscopy to characterize biological materials. Nature Protocols, 11(4), 664–87.

[4] Dieing, T., Hollricher, O., Toporski, J., Eds. Confocal Raman microscopy; Springer: Berlin, 2011.