Volume 1 | Number 1 | Jan 2013 | Pages 1–100

Analytical Methods

www.rsc.org/methods

ROYAL SOCIETY OF CHEMISTRY

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

Investigating the use of Raman and Immersion Raman Spectroscopy for Spectral Histopathology of Metastatic Brain Cancer and Primary Sites of Origin

Leanne M. Fullwood[1], Graeme Clemens[1], Dave Griffiths[2], Katherine Ashton[3], Timothy P. Dawson[3], Robert W. Lea[2], Charles Davis[3], Franck Bonnier[4], Hugh J. Byrne[4], and Matthew J. Baker[1*]

[1] Centre for Materials Science, Division of Chemistry, University of Central Lancashire, Preston, PR1 2HE, UK
[2] School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston, PR1 2HE, UK
[3] Lancashire Teaching Hospital NHS Trust Royal Preston Hospital, Sharoe Green Lane, Fulwood, Preston PR2 9HT, UK
[4] FOCAS Research Institute, Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland
*Corresponding Author: MJBaker@UCLan.ac.uk

Abstract

It is estimated that approximately 13,000 people in the UK are diagnosed with brain cancer every year; of which 60% are metastatic. Current methods of diagnosis can be subjective, invasive and have long diagnostic windows. Raman spectroscopy provides a non-destructive, non-invasive, rapid and economical method for diagnosing diseases. The aim of this study was to investigate the use of Raman and immersion Raman spectroscopy for diagnosing metastatic brain cancer and identifying primary sites of origin using brain tissue. Through spectral examination, the Raman peaks at 721 cm^{-1} and 782 cm^{-1} were identified as being the most distinct for discriminating between the glioblastoma multiforme (GBM), metastatic and normal brain tissue spectra. A ratio score plot of these peaks calculated the classification sensitivities and specificities as 100% and 94.44% for GBM, 96.55% and 100% for metastatic brain, and 85.71% and 100% for normal brain tissue. Principal Component-Linear Discriminant Analysis (PC-LDA) also showed discrimination between normal, GBM and metastatic brain tissue spectra. We also present, for the first time, the use of Raman spectroscopy to investigate primary site of origin for metastatic brain cancer and any biochemical differences between different primary and metastatic cancer using linked samples. This study revealed interesting spectral differences in the amide regions showing changes in the biochemistry of the metastatic brain cancer from the primary cancer

Keywords: Immersion Raman, Vibrational Spectroscopy, Diagnosis, Brain Cancer, Metastatic

Introduction

It is estimated that a third of people in the UK will develop cancer at some point during their lifetimes; and approximately 50% of these will die as a result of the disease within five years after diagnosis [1]. Brain metastases are the most common form of intracranial neoplasms in adults and are predicted to develop in 20-40% of cancer patients [2]. It is reported that around 13,000 people in the UK are diagnosed with brain cancer every year, of which, 60% are metastatic tumours that have originated from primary cancers located outside the central nervous system [3]. All cancers have the potential to metastasise and their probability of doing so is influenced by their location; lung cancer is the most common cancer to metastasise and is responsible for 50% of all brain metastases, breast cancer is the second most common and accounts for 15-25%, melanoma accounts for 5-20%, and the remaining 5-30% are as a result of other cancers [4]. Identifying the site of origin of the primary tumour increases the efficiency of treatment and thus patient survival; however, in 15% of metastatic cases, the location of the primary cancer is unknown [5].

The current method of diagnosis for brain metastases and their primary sites is histopathological analysis. The cancer type and the origin of the primary cancer can be identified based on the tissue architecture observed [6]. This method requires a trained neuropathologist, has a long diagnostic window, and can be subjective [7]. As well as these issues, disagreement amongst pathologists is reported to occur in up to 43% of specimens [8]. The accurate and rapid diagnosis of disease allows early intervention of appropriate treatment, thus increasing life expectancy and reducing healthcare costs [9]. Therefore, there

is a requirement for non-subjective techniques that can rapidly and accurately identify disease.

Raman spectroscopy is an analytical technique that utilises the phenomenon of inelastic light scattering to produce characteristic spectra unique to specific samples. When a tissue becomes diseased, molecular changes occur, which should be reflected in its spectrum. The identification of spectral differences between healthy and diseased tissue should enable objective diagnosis of specific diseases. Raman spectroscopy is not significantly affected by water and thus can be used *in vivo* to reduce and aid biopsy [10]. Raman spectroscopy is also a non-destructive, rapid and economical technique, making it an attractive method for diagnosis [11].

The potential of Raman spectroscopy as a diagnostic technique for diseases has been demonstrated by a variety of studies. Many papers have shown the ability of Raman spectroscopy to diagnose cancers, such as: brain cancers using tissue on low-E microscope slides and CaF_2 slides respectively [12, 13], lung tumours through the analysis of bronchial tissue sections [14], gastric adenocarcinomas using tissue samples [15, 16], non-melanoma skin cancers from the analysis of tissue *in vivo* [17], laryngeal cancers from tissue at endoscopy [18], breast cancers using fresh and frozen tissue specimens respectively [19, 20], cervical cancers from tissue *in vivo* and *ex vivo* [21, 22], bladder and prostate cancers from tissue sections [23, 24] and oesophagus and colon cancers through the analysis of snap frozen tissue on CaF_2 [25].

Gajjar *et al.* reported the ability of Raman spectroscopy to differentiate between brain tumour and healthy brain tissue [12]. They observed successful discrimination between different brain tumour types in their study. They reported that metastatic brain tissue could be recognised from healthy brain tissue on low-E substrates based on spectral peaks at: 997 cm^{-1}

(phospholipids and glucose-I-phosphate), 1077 cm^{-1} and 1446 cm^{-1} (lipids and proteins), 1241 cm^{-1} (Amide III), $\approx$ 1460 cm^{-1} (cytosine) and 1654 cm^{-1} (Amide I) [12].

Tissue samples are usually fixed using formalin and impregnated with paraffin wax to preserve the tissue for future analysis. The addition of paraffin also provides support for the microtomy process, enabling the sectioning of the tissue block for microscopic examination. However, prior to histological and Raman analysis, tissue sections must undergo a dewaxing process to remove the added paraffin. This then allows histological staining of the tissue for pathological examination, and reduces the paraffin contributions in Raman spectra. Ó Faoláin *et al*. investigated the efficacy of dewaxing procedures on formalin fixed, paraffin preserved (FFPP) cervical tissue on glass slides [26]. They reported that dewaxing procedures employing the commonly used, xylene and histoclear solvents, do not completely remove all of the paraffin wax; they found hexane to be a much more efficient dewaxing agent [26] However it requires 18 hours of tissue submersion and is not clinically used for dewaxing procedures, conventional clinical dewaxing procedures were employed here [27].

Immersion Raman spectroscopy utilises an immersion lens which is in direct contact with an appropriate liquid, such as deionised water covering a sample under investigation. Bonnier *et al.* describe and successfully demonstrate the use of immersion Raman spectroscopy for both live cell and the first example of *in vitro* tissue specimens on CaF$_2$, and observed an improvement in spectral quality, sample stability and the reduction of spectral background [28, 29]. Bonnier *et al.* [28, 29] demonstrated that fluorescence, contributing to spectral background, is only significant from sources producing short wavelengths of light, as proteins typically only fluoresce when sources producing wavelengths below 500nm are used and therefore, the background contribution seen from spectra acquired when using 785 nm lasers should not be attributed to fluorescence, but to morphology dependant scattering of the incident light, and Raman lines that cause non-collimated entry into the spectrometer as stray

light. The exchange of a tissue/air interface with a tissue/liquid interface results in more uniform refractive indices, thus reducing this level of stray light, and in turn, the intensity of the spectral background. Submerging the sample in liquid also protects the tissue specimen from photo-damage; enabling more powerful lasers of shorter wavelengths to be used, longer acquisition times and higher numbers of accumulations to be employed hence, improving spectral quality.

We present a comparison of Raman and immersion Raman recorded spectra, with a study highlighting the use of immersion Raman spectroscopy for the rapid diagnosis of site of origin of metastatic brain tumours. Vibrational spectroscopy has the ability to revolutionise the clinical environment allowing for increased efficiency within the diagnostic regime with corresponding decreases in mortality, morbidity and economic impact upon the health services [30].

Experimental

Study participants

Tissue sections were cut using a microtome. Parallel tissue sections of 4 μm in thickness were cut for glass microscope slides and 10 μm for CaF_2 (Crystran, UK) substrates. Tissue was obtained from formalin fixed paraffin preserved (FFPP) tissue blocks from the Brain Tumour North West (BTNW) bio-bank under ethical approval (BTNW/WRTB 13_01). Patient data consisted of histological information, patient gender, and date of birth and origin of metastasis/histological subtype. A total of 48 tissue specimens were obtained from 41 different patients. Tissue consisted of normal brain samples (n=7), glioblastoma multiforme (GBM) brain samples WHO (World Health Organisation) grade IV (n=5), metastatic brain samples (n=29) and primary cancer tissue samples (n=7). Table 1 displays further information about the tissue specimens.

Dewaxing and Haematoxylin and Eosin staining of tissue

Tissue sections on the microscope slides were stained with Haematoxylin and Eosin (H&E) for histological examination So as to remove the paraffin, tissue sections were de-waxed through 2 x 5 minute baths of Histoclear followed by 2 x 5 minute baths of ethanol. Prior to H&E staining, tissue sections were washed in distilled water for 5 minutes after de-waxing. Sections were then bathed in haematoxylin for 5 minutes and washed in warm tap water to allow the nuclei to turn blue. The sections were then covered in eosin for 4 minutes and washed off with distilled water. Finally, the tissue sections were dehydrated in 2 x 5 minute baths of ethanol and cleared in 2 x 5 minute baths of Histoclear tissue sections were protected and preserved through the application of Histomount and a coverslip. The sections were then microscopically examined in order to identify the metastatic sites present in the tissue.

Figure 1 shows the microscopic images of the H & E stained tissue samples at x 100 magnification. It can be observed that normal, metastatic and GBM brain tissue differ from one another architecturally, with differences also being observed between the differing metastatic types.

Tissue section preparation for Raman spectroscopic analysis

The tissue sections on the CaF_2 substrates also needed to be de-waxed prior to Raman analysis, in order to reduce the paraffin peak contributions in the spectra. The de-waxing procedure consisted of 3 x 5 minute baths of Histoclear followed by 3 x 5 minute baths of ethanol. The sections were left to air dry for 30 minutes, placed in a Petri dish and stored in a desiccator until spectroscopic analysis.

Raman spectroscopy instrumentation and analysis

Spectroscopic measurements were carried out on a Horiba Jobin-Yvon LabRAM HR800 spectrometer. An air cooled CLDS point mode diode 785 nm laser with a single edge filter (cut off to 100 cm^{-1}) and an output power of 300mW was used to acquire spectra; which was used with a grating of 300 gr/mm and blazed at 1000 nm. Non-immersion point spectra were acquired using a 0.9 numerical aperture (x 100) (MPlanN) objective and immersion point spectra were acquired using a 0.75 numerical aperture (x 60) objective (LUMPlanFLN, Olympus). The confocal hole was set at 100 μm for 785 nm spectral collections. The detector used was an Andor electromagnetic (EM) charged coupled device (CCD). A video camera within the Raman system was used to take images of the specimens.

The instrumentation was calibrated before operation to silicon at the spectral line of 520.8 cm^{-1}. Spectra were acquired using the 785 nm laser at 100% exposure for 30 s and accumulated twice, with both the immersion and non-immersion objectives. A total of 1574 spectra were generated through the production of 20 spectra from three to five areas on each tissue sample, ensuring a representative area was used to gather data. Immersion Raman spectroscopy was carried out by submerging the tissue sample in deionised water for spectral collection.

Data processing and multivariate analysis

Pre-processing and multivariate analysis was carried out on the raw data using LabSpec 6 spectroscopy software suite (HORIBA Scientific) and MATLAB version 7.11.0 (R2010b) (The MathWorks, Inc., USA) using in-house written software. Pre-processing methods were kept to a minimum to enable better reproducibility; data were background subtracted through the application of a fifth order polynomial and smoothed using 7 point smoothing (Labspec 6) and vector normalised. Paraffin peaks were removed from the spectra

for multivariate analysis at the following wavenumbers: 882 cm^{-1} - 912 cm^{-1}, 1051 cm^{-1} - 1071 cm^{-1}, 1115 cm^{-1} - 1143 cm^{-1}, 1163 cm^{-1} – 1187 cm^{-1}, 1284 cm^{-1} – 1305 cm^{-1} and 1407 cm^{-1} – 1501 cm^{-1}. Multivariate analysis was then carried out on the data set using PC-DFA. Martin *et al.* demonstrate the successful use of PCA-LDA for showing clustering and identifying the major contributory variables [31].

Principal component analysis (PCA) is an unsupervised multivariate algorithm used to find spectral differences between recorded spectra. Being unsupervised means that the algorithm has no priori information on the grouping of the recorded spectra, PCA can only find orthogonal dimensions of spectral variance. One advantage to PCA is that can be used to reduce the dimensionality of the dataset prior to discriminate classification algorithms, such as Linear Discriminant Analysis (LDA). Without dimension reduction, LDA can overfit, producing over generous classification results based on random errors or noise, and not on the relationships between the recorded spectra. LDA is a supervised technique requiring prior knowledge of the group classes beforehand. LDA then finds the best linear hyperplane between the class groups at each variable dimension, maximising the intergroup variance and minimising the intragroup variance. The validities of the PC-LDA models were estimated using a cross-validation technique, partitioning the data matrix into two sets for training and testing. One third of the original data set was randomly removed while the remaining training set data is used to build the classification model. The classification model is then tested using the removed (blind) test set spectra [32, 33]. The performance of the model is judged on how well the model, built using the training set, predicts the new and previously unseen test set data. PC-LDA model performance was judged through the percentage of test set spectra correctly classified (%CC) and spectra classification sensitivities and specificities based on the 90% inner confidence ellipse. This study investigates three different PC-LDA models.

Results and Discussion

Raman and Immersion Raman spectroscopy comparison

The averaged raw spectra of all tissue samples on CaF_2 using immersion and non-immersion Raman spectroscopy are shown in Figure 2. It is clear that the averaged immersion Raman spectrum has an initial lower spectral background than the averaged non-immersion spectrum in the 600 cm^{-1} to 930 cm^{-1} spectral range, which remains at a relatively stable intensity throughout. Immersion Raman also produces better delineated peaks with larger areas than spectra acquired by non-immersion Raman. Therefore, due to the advantage of better spectral quality, immersion Raman spectroscopy was chosen over standard techniques to analyse its potential as a diagnostic tool for metastatic brain cancer and its ability to identify primary cancer sites of origin.

Spectral Histopathology:

Cancer vs. Metastatic Tumours vs. Non-Cancer

The averaged, vector normalised and background-subtracted immersion Raman spectra of normal (157 spectra from 7 patients), metastatic (668 spectra from 31 patients) and GBM (127 spectra from 5 patients) brain tissue on CaF_2 substrates are displayed in Figure 3. Table 2 lists the assignments of the major peaks present in recorded spectra, based on literature [14, 25, 34-36]. Spectral differences between these tissue types are observed at the various peaks throughout the averaged spectra. The peak at 721cm^{-1}, assigned as the C-N symmetric stretch of choline [34, 35], is prominently observed in GBM tissue. In normal and metastatic tissue this peak is not present, although a 725 cm^{-1} peak, assigned to adenine is present [37] (Figure 3). Choline content has been observed to increase in GBM tissue and all primary tumours through magnetic resonance spectroscopy; thought to be caused by increased cell proliferation and cell membrane turnover, resulting in an increased production of choline transporters and kinase enzymes [38].

GBM and metastatic tissue both exhibit 782 cm^{-1} peaks of a considerable intensity in their spectra, assigned to nucleotide ring breathing [25], whereas the peak in the normal tissue spectrum is much weaker and is not easily identifiable. A peak at ~1081 cm^{-1}, assigned to PO_2^- symmetric stretching (nucleic acids) [39], is observable in the averaged spectrum of normal brain tissue, yet appears absent from the spectra of cancerous tissue (Figure 3).

An amide III peak is present at 1263 cm^{-1} [36] in the normal tissue spectrum but for GBM and metastatic brain tissue spectra, the peak has been shifted to 1250 cm^{-1}, possibly as a result of peak broadening, or changes in hydrogen bonding. As well as the shift, there is a noticeable increase in amide III peak intensity from the GBM and metastatic brain average spectra, when compared to normal tissue spectrum. As the 782 cm^{-1} has previously associated with ring breathing vibrations of nucleic acids [40] the increase in intensity from the 782 cm^{-1} and may suggest an increase in cellular density from the cancerous tissue. The increase in intensity from the amide III band may well be in agreement with this, showing an increase in cellular proteins as a result of increased cell mass in the tissue.

The peak present at approximately 1658 cm^{-1}, attributed to the C=O stretch of amide I [25], appears to be shifted in the average spectrum of metastatic tissue to a higher Raman shift of 1661 cm^{-1}. The intensity of this 1658 cm^{-1} peak is also decreased in both the cancerous spectra. This is consistent with the report of Gniadecka *et al.,* who found reduced intensities of the amide I band of proteins in the spectra of melanoma. The changes in the amide bands between the normal and melanoma specimens are attributed to conformational changes to the protein structures [41].

It can also be observed from the averaged spectrum of normal tissue that the residual paraffin peaks (asterisked peaks) present are of a much lower intensity than those in the spectra of metastatic and GBM tissue. This could be attributed to either structural or

compositional differences between the tissue types showing a tissue effect on the dewaxing efficacy [27].

The mean spectral intensity changes between spectral peaks of normal, GBM and metastatic brain tissue (Figure 3) indicate alterations in specific biomolecule concentrations when the tissue is in different states of health. These observable spectral differences demonstrate the ability of Raman spectroscopy as a potential diagnostic tool for brain cancer.

As a result of there being spectral differences between GBM, Metastatic and normal brain tissue mean spectra (Figure 3), the recorded spectra were investigated using PC-LDA to see if linear separation boundaries could be found between the spectra recorded from the different tissue types in multidimensional space (Figure 4). Figure 4 shows a LDA scores plot (LDA1 vs. LDA2) of spectra recorded from GBM, Metastatic and normal brain tissue using immersion Raman spectroscopy. Although complete separation of the different tissue type spectra is not seen in the LDA scores plot (a degree of overlap seen between the different groups), there is some separation between spectra recorded from all three tissue types. The majority of GBM and normal brain tissue spectra are separated from Metastatic tissue spectra along the LDA 1 axis, and the majority of normal brain tissue spectra can be discriminated from GBM spectra along the LDA 2 axis. However, it can be seen in Figure 1 that all tissue types have some degree of tissue heterogeneity, with some Metastatic tissue types and GBM tissue being very heterogeneous. The heterogeneous nature of GBM tissue is said to be due to areas of necrosis found in the tissue surrounded by pseudopalisades and microvascular hyperplasia, believed to be instrumental in the accelerated growth of GBM [42]. By

investigating different areas on each tissue sample with Raman spectroscopy there is the possibility that non-cancerous tissue chemistry will be captured when investigating very heterogeneous cancerous tissue. This is possible reflected in Figure 4 (a), which shows that both Metastatic and GBM tissue spectra have some overlap with normal tissue spectra. Therefore, the spectra used for Figure 4 (a) was split into smaller groups and averaged (The 20 spectra recorded from each patient tissue sample were split into 4 and averaged). This produced the LDA model displayed in Figure 4 (b), which now shows good separation between all three classes of tissue spectra; producing a classification model performance score of 97%CC. Details of both PC-LDA models created in Figure 4 can be found in Table 3.

Figure 5 displays the LDA loadings plots relating to the LDA scores plot in Figure 4 (b). Table 4 identifies the Raman shifts from the LDA 1 and 2 loading plots which are responsible for distinguishing normal, Metastatic and GBM tissue spectra. The LDA 1 axis separates the Metastatic tissue spectra from the rest. The main peaks in the LDA 1 loading plot included the 718, 925, 1250, 1400 and 1670 cm^{-1} Raman bands associated with symmetric choline C-N stretch (membrane phospholipid head)/adenine, C-C stretching mode of proline, Amide III band, CH$_3$ bending mode due to methyl bond in the membrane and the Amide I C=O stretching mode of β-sheet structural proteins respectively. These spectral differences are in agreement with the mean spectra and show the Metastatic tissue spectra to have reduced intensities from the 718 and 925 cm^{-1} peak's, and increased intensities from the 1250 1400, and 1670 cm^{-1} peaks when compared to the tissue spectra recorded from the GBM and normal tissue. The LDA 2 axis in Figure 4 (b) separates the majority of GBM and Metastatic spectra from normal tissue spectra. The main peaks in the LDA 2 loading plot included the 719, 780/805, 1078, 1234, 1268, 1653, 1672 cm^{-1} Raman bands associated with symmetric choline C-N stretch, cytosine/uracil ring breathing of nucleotide/A-DNA, PO$_2$

symmetric stretching (nucleic acids), antisymmetric phosphate stretching, Amide III: (α-helix, protein), Amide I (C=O stretching mode of proteins, α-helix conformation) and the Amide I C=O stretching mode of β-sheet structural proteins respectively. The loading of the peaks at 1235 and 1268 cm^{-1} are in agreement with the mean spectral differences, which show the Amide III peak in the cancerous spectra to shift from 1268 cm^{-1}, associated with α-helix protein, to a lower frequency associated with β-sheet protein. The LDA 2 loading plot and mean spectral comparisons also show the cancerous spectra to have reduced intensity from the Amide I α-helix peak when compared to the normal tissue spectra, while the cancerous spectra also show increased intensity at ~1670 cm^{-1}, associated with Amide I proteins with β-sheet structural conformations. This may suggest a link between protein secondary structure content and cancerous tissue. Yamada *et al.* also showed β-sheet protein levels to increase significantly when investigating necrotic areas of murine carcinoma using FT-IR microspectroscopy. The increase in β-sheet protein levels coincided with a sharp decrease in α-helix protein levels, in agreement with our results [43]. It can also be seen from Figure 5 (b) that the cancerous tissue spectra have increased intensity from peaks associated with nucleic acids and DNA structures. As previously described, the increase intensity from nucleic acid structures may suggest an increase in cellular density from the cancerous tissue, in agreement with increased cell proliferation.

From the mean spectral differences (Figures 3 and 5) it can be visibly seen that the average normal brain tissue spectrum has a low intensity from the 782 cm^{-1} peak, when compared to the average GBM and Metastatic spectra. It can also be seen that the average GBM mean spectrum has a much greater intensity in the 721 cm^{-1} peak than the average normal tissue and metastatic spectrums. As a result, a 2D scores plot was produced based on the intensity ratios between the 721/782 cm^{-1} peaks and the 620 cm^{-1} peak (620 cm^{-1} peak intensity was relatively uniform in all the recorded spectra) (Figure 6). Figure 6 shows the

721 and 782 cm^{-1} peaks to be significant discrimination markers for GBM, Metastatic and normal tissue spectra, with average GBM, Metastatic and normal tissue spectra clearly grouped in the 2D scores plot (See Table 5 resulting sensitivities and specificities). This shows that the 721 and 782 cm^{-1} peaks could potentially be markers for diseased tissue therefore, the monitoring of these peak intensities through Raman spectroscopy may aid the early diagnosis of disease. However, Figure 6 does show two spectra which cannot be grouped with their representative class, identified as patient 007 and 509, belonging to the normal and metastatic group respectively. Upon visual inspection of the spectra, it was observed that they notably differed from the total averaged spectra for their tissue types. This could potentially be a result of tissue misclassification.

Primary site analysis

Raman spectra were also recorded from tissue representative of primary cancers (breast, melanoma, oesophagus and stomach and colon/rectum) and patient averaged. The primary cancer tissue samples are linked samples (from the same patient) of corresponding metastatic tissue samples in the metastatic tumour model. As before, PC-LDA was performed on the spectra recorded from the tissue samples using immersion Raman spectroscopy. The LDA scores plot (Figure 7) shows that spectra recorded from the colon/rectum (**blue**) and oesophagus & stomach cancerous tissue (**green**) have some good separation from the remaining spectra. Less clear is the separation of the primary breast and melanoma tissue spectra which are shown to be spectrally similar. This may suggest that the tissue architecture of both primary breast and melanoma tissue is similar.

Analysis of site of origin

Figure 7 shows that the identification of spectra recorded from some primary cancers is possible using multivariate classification algorithms. As the origins of metastatic brain tissue

are determined from the tissue architecture observed, it should also be possible to determine metastatic brain tissue origins from the spectra recorded. Therefore, PC-LDA was also performed on the metastatic tissue data with the LDA scores plot displayed in Figure 8. As with the data used in Figure 4 (b), the 20 spectra recorded from each patient tissue sample were split into 4 and averaged. The LDA scores plot in Figure 8 shows that although there is some overlap between the spectra recorded from each type of brain metastases, there is some separation between metastatic breast (**blue**), melanoma (**red**), oesophagus & stomach (**black**) and colon/rectum (**green**)/lung (**cyan**) spectra. As spectra recorded from both metastatic colon/rectum and lung tissue almost overlap in LDA scores space, this may suggest that the different tissue types have similar tissue architecture. The spectral similarities may also be due to insufficient spectra being recorded from the different tissue types, and as a result, spectra recorded from the three different sites of each patient sample have not been able to capture the differing tissue architecture which would discriminate between the two. However, it may also be possible that that the movement of metastatic cancer from primary sites to the brain effects the biochemistry of the cancers, with all metastatic cancer found in the brain becoming biochemically similar due to their new environment. This would suggest that biochemical differences between primary cancers are generally a result of the organ location and not the cancer type.

Spectral analysis of primary site and corresponding brain metastasis

Figure 9 (a) displays the baseline-subtracted averaged spectra of both Metastatic brain and their corresponding primary sites. It is observable from Figure 9 (a) that the Metastatic spectra (**black**) from the different sites of primary origin all look similar, whereas the same cannot be said from the primary cancer spectra (**blue**), with primary Colon/Rectum, Melanoma and Oesophagus and Stomach tissue spectra all showing mean spectral differences

when compared against their corresponding primary sites. This supports the theory that the movement of metastatic cancer from primary sites to the brain affects the biochemistry of the cancers, with all metastatic brain tissue spectra having relatively similar chemistry. However, the average primary breast cancer tissue spectrum and the corresponding metastatic brain tissue spectrum appear relatively similar to one another (Figure 9 (a, 1)). The PCA scores plot also shows no separation between the primary and Metastatic breast spectra therefore, suggesting that there is no real change from primary and Metastatic breast tissue. On the other hand, the spectral differences between oesophagus/stomach tissue and metastatic brain tissue are clear. The most noticeable disparity between the two average spectra is due to the peak at 1621 cm^{-1}, which is much more intense in the primary site spectrum than the metastatic brain spectrum. This coincides with a much reduced peak intensity at ~1658 cm^{-1} for the average primary site spectrum thus, a considerable change in protein structure for the primary oesophagus/stomach tissue during movement to the brain. It can also be seen from the mean spectral comparisons (Figure 9 (a, 4) that the primary site spectrum has a peak at ~800 cm^{-1}, associated with DNA and nucleic acid molecules, whereas the corresponding Metastatic brain spectrum does not. For melanoma tissue, the spectral differences between melanoma primary site and metastatic tissue are primarily due spectral differences seen from the amide I peak, the amide III and peaks present between 850-950 cm^{-1}. However, the spectral differences between Colon/Rectum cancer tissue and the Metastatic brain tissue are not large, and vary mainly through differences in peak intensities. Therefore, only small biochemical changes in the tissue chemistry during Metastasis.

Conclusion

In this study, we have demonstrated the use of immersion Raman spectroscopy to differentiate between GBM, metastatic and normal brain tissue with sensitivities and

specificities of 100% and 94.44% respectively for GBM, 96.55% and 100% respectively for

metastatic brain, and 85.71% and 100% respectively for normal brain tissue, based on the

intensity ratios between the 721/782 cm^{-1} peaks. This shows that the Raman spectral peaks at

721 and 782 cm^{-1}, primarily associated with the symmetric C-N stretch of choline and the

cytosine/uracil ring breathing of nucleotides respectively, may be possible spectral

biomarkers for the discrimination of normal, GBM and Metastatic tissue. As well as this, PC-

LDA was performed on the tissue spectra producing good separation of GBM, Metastatic and

normal tissue spectra, again highlighting the potential of the analytical technique for the

classification of cancerous tissue. As well as investigating GBM, Metastatic and normal

tissue spectra, this study has also been able to distinguish some of the different primary site

cancerous tissue and the different Metastatic brain cancerous tissue, based on the Raman

spectra recorded. Being able to distinguish Metastatic brain tissue origins based on recorded

Raman spectra, coupled with supervised multivariate analysis, would be non-subjective,

therefore, the analytical technique could be used as a qualitative tool alongside

neuropathology for disease diagnosis. For the first time the chemistry of Metastatic brain and

their corresponding primary site tissue have also been investigated in this study by Raman

spectroscopy, with primary Colon/Rectum, primary Melanoma and primary Oesophagus and

Stomach tissue all being separated from their representative Metastatic brain tissue in PCA

scores space, plus, significant mean spectral differences were observed. This suggests that the

migration of cells from primary sites to the brain may cause alterations to their biochemistry,

and result in biochemical conformity to their new environment.

Acknowledgements

Centre for Materials and Science at the University of Central Lancashire, the Focas Institute

at Dublin Institute of Technology, the European Community Action Scheme for the Mobility

of University Students and the Association of British Spectroscopists. The Focas Research

Institute is supported through the National Biophotonics and Imaging Platform (NBIP)

Ireland funded under the Higher Education Authority PRTLI (Programme for Research in

Third Level Institutions) Cycle 4, co-funded by the Irish Government and the European

Union.

References

1. Cancer Research UK. *All cancers combined key facts*. 2012 [15/04/2013]; Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/keyfacts/Allcancerscombined/.
2. Richards, G.M., D. Khuntia, and M.P. Mehta, Therapeutic management of metastatic brain tumors. *Critical Reviews in Oncology/Hematology*, 2007. **61**(1): p. 70-78.
3. Brain Research Trust. *About Brain Tumours*. [cited 2012 08/10]; Available from: http://www.brt.org.uk/brain-tumours.
4. Soffieti, R., P. Cornu, J. Y. Delattre, R. Grant, F. Graus, W. Grisold, J. Heimans, J. Hildebrand, P. Hoskin, M. Kallijo, P. Krauseneck, C. Marosi, T. Siegal and C. Vecht, Brain Metastases. *European Handbook of Neurological Managments, 2nd ed. (eds: N. E. Gilhus, M. P. Barnes and M. Brainin)*, 2011. **1**: p. 437-445.
5. Bergner, N., B.F.M. Romeike, R. Reichart, R. Kalff, C. Krafft, and J. Popp, Raman and FTIR microspectroscopy for detection of brain metastasis. *Clinical and Biomedical Spectroscopy and Imaging II*, 2011: p. 80870X-80870X.
6. Mahadevan-Jansen, A. and R.R. Richards-Kortum, Raman spectroscopy for the detection of cancers and precancers. *Journal of Biomedical Optics*, 1996. **1**(1): p. 31-70.
7. Krafft, C., S.B. Sobottka, G. Schackert, and R. Salzer, Analysis of human brain tissue, brain tumors and tumor cells by infrared spectroscopic mapping. *Analyst*, 2004. **129**(10): p. 921-925.
8. Raab, S.S., D.M. Grzybicki, J.E. Janosky, R.J. Zarbo, F.A. Meier, C. Jensen, and S.J. Geyer, Clinical impact and frequency of anatomic pathology errors in cancer diagnoses. *Cancer*, 2005. **104**(10): p. 2205-13.
9. Ellis, D.I., and R. Goodacre, Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. *Analyst*, 2006. **131**: p. 875-885.
10. Choo-Smith, L.-P., H.G.M. Edwards, H.P. Endtz, J.M. Kros, F. Heule, H. Barr, J.S. Robinson, H.A. Bruining, and G.J. Puppels, Medical applications of Raman spectroscopy: From proof of principle to clinical implementation. *Biopolymers*, 2002. **67**(1): p. 1-9.
11. Moros, J., S. Garrigues, and M. de la Guardia, Vibrational spectroscopy provides a green tool for multi-component analysis. *TrAC Trends in Analytical Chemistry*, 2010. **29**(7): p. 578-591.
12. Gajjar, K., L. D. Heppenstall, W. Pang, K. M. Ashton, J. Trevisan, I. I. Patel, V. Llabjani, H. F. Stringfellow, P. L. Martin-Hirsch, T. Dawson and F. L. Martin, Diagnostic segregation of human brain tumours using Fourier-transform infrared and/or Raman spectroscopy coupled with discriminant analysis. *Analytical Methods*, 2012.

13. Krafft, C., S.B. Sobottka, G. Schackert, and R. Salzer, Raman and infrared spectroscopic mapping of human primary intracranial tumors: a comparative study. *Journal of Raman Spectroscopy*, 2006. **37**(1-3): p. 367-375.

14. Huang, Z., A. McWilliams, H. Lui, D. I. McLean, S. Lam and H. Zeng,, Near-infrared Raman spectroscopy for optical diagnosis of lung cancer. *International Journal of Cancer*, 2003. **107**(6): p. 1047-1052.

15. Teh, S.K., W. Zheng, K.Y. Ho, M. Teh, K.G. Yeoh, and Z. Huang, Near-infrared Raman spectroscopy for early diagnosis and typing of adenocarcinoma in the stomach. *British Journal of Surgery*, 2010. **97**(4): p. 550-557.

16. Bergholt, M.S., W. Zheng, K. Lin, K.Y. Ho, M. Teh, K.G. Yeoh, J.B. Yan So, and Z. Huang, In vivo diagnosis of gastric cancer using Raman endoscopy and ant colony optimization techniques. *International Journal of Cancer*, 2011. **128**(11): p. 2673-80.

17. Lieber, C.A., S.K. Majumder, D.L. Ellis, D.D. Billheimer, and A. Mahadevan-Jansen, In vivo nonmelanoma skin cancer diagnosis using Raman microspectroscopy. *Lasers in Surgery and Medicine*, 2008. **40**(7): p. 461-467.

18. Lin, K., D.L.P. Cheng, and Z. Huang, Optical diagnosis of laryngeal cancer using high wavenumber Raman spectroscopy. *Biosensors and Bioelectronics*, 2012. **35**(1): p. 213-217.

19. Abramczyk, H., J. Surmacki, B. Brożek-Płuska, Z. Morawiec, and M. Tazbir, The hallmarks of breast cancer by Raman spectroscopy. *Journal of Molecular Structure*, 2009. **924–926**(0): p. 175-182.

20. Moreno, M., L. Raniero, E.A.L. Arisawa, A.M. do Espirito Santo, E.A.P. dos Santos, R.A. Bitar, and A.A. Martin, Raman spectroscopy study of breast disease. *Theoretical Chemistry Accounts*, 2009.

21. Kanter, E.M., E. Vargis, S. Majumder, M.D. Keller, E. Woeste, G.G. Rao, and A. Mahadevan-Jansen, Application of Raman spectroscopy for cervical dysplasia diagnosis. *Journal of Biophotonics*, 2009. **2**(1-2): p. 81-90.

22. Lyng, F.M., E.Ó. Faoláin, J. Conroy, A.D. Meade, P. Knief, B. Duffy, M.B. Hunter, J.M. Byrne, P. Kelehan, and H.J. Byrne, Vibrational spectroscopy for cervical cancer pathology, from biochemical analysis to diagnostic tool. *Experimental and Molecular Pathology*, 2007. **82**(2): p. 121-129.

23. Crow, P., A. Molckovsky, N. Stone, J. Uff, B. Wilson, and L.M. WongKeeSong, Assessment of fiberoptic near-infrared raman spectroscopy for diagnosis of bladder and prostate cancer. *Urology*, 2005. **65**(6): p. 1126-1130.

24. Tollefson, M., J. Magera, T. Sebo, J. Cohen, A. Drauch, J. Maier, and I. Frank, Raman spectral imaging of prostate cancer: can Raman molecular imaging be used to augment standard histopathology? *British Journal of Urology International*, 2010. **106**(4): p. 484-488.

25. Stone, N., C. Kendall, J. Smith, P. Crow, and H. Barr, Raman Spectroscopy for identification of epithelial cancers. *Faraday Discussions*, 2004. **126**: p. 141-157.

26. Ó Faoláin, E., M.B. Hunter, J.M. Byrne, P. Kelehan, H.A. Lambkin, H.J. Byrne, and F.M. Lyng, Raman spectroscopic evaluation of efficacy of current paraffin wax section dewaxing agents. *Journal of Histochemistry and Cytochemistry*, 2005. **53**(1): p. 121-9.

27. Fullwood, L.M., D. Griffiths, K. Ashton, T. Dawson, R.W. Lea, C. Davis, F. Bonnier, H.J. Byrne, and M.J. Baker, Effect of substrate choice and tissue type on tissue preparation for spectral histopathology by Raman microspectroscopy. *Analyst*.

28. Bonnier, F., S.M. Ali, P. Knief, H. Lambkin, K. Flynn, V. McDonagh, C. Healy, T.C. Lee, F.M. Lyng, and H.J. Byrne, Analysis of human skin tissue by Raman microspectroscopy: Dealing with the background. *Vibrational Spectroscopy*, 2012. **61**(0): p. 124-132.

29. Bonnier, F., A. Mehmood, P. Knief, A.D. Meade, W. Hornebeck, H. Lambkin, K. Flynn, V. McDonagh, C. Healy, T.C. Lee, F.M. Lyng, and H.J. Byrne, In vitro analysis of immersed human tissues by Raman microspectroscopy. *Journal of Raman Spectroscopy*, 2011. **42**(5): p. 888-896.

30. Hands, J.R., P. Abel, K. Ashton, T. Dawson, C. Davis, R. Lea, A. McIntosh, and M. Baker, Investigating the rapid diagnosis of gliomas from serum samples using infrared spectroscopy and cytokine and angiogenesis factors. *Analytical and Bioanalytical Chemistry*, 2013: p. 1-9.

31. Martin, F.L., M.J. German, E. Wit, T. Fearn, N. Ragavan, and H.M. Pollock, Identifying variables responsible for clustering in discriminant analysis of data from infrared microspectroscopy of a biological sample. *J Comput Biol*, 2007. **14**(9): p. 1176-84.

32. Baker, M.J., E. Gazi, M.D. Brown, J.H. Shanks, P. Gardner, and N.W. Clarke, FTIR-based spectroscopic analysis in the identification of clinically aggressive prostate cancer. *British Journal of Cancer*, 2008. **99**(11): p. 1859-1866.

33. Baker, M.J., E. Gazi, M.D. Brown, N.W. Clarke, J.C. Vickerman, and N.P. Lockyer, ToF-SIMS PC-DFA analysis of prostate cancer cell lines. *Applied Surface Science*, 2008. **255**(4): p. 1084-1087.

34. Huang, Z., X. Chen, Y. Chen, J. Chen, M. Dou, S. Feng, H. Zeng, and R. Chen, Raman spectroscopic characterization and differentiation of seminal plasma. *Journal of Biomedical Optics*, 2011. **16**(11): p. 110501-110501.

35. Sikirzhytski, V., K. Virkler, and I.K. Lednev, Discriminant analysis of Raman spectra for body fluid identification for forensic purposes. *Sensors (Basel)*, 2010. **10**(4): p. 2869-84.

36. Chan, J.W., D.S. Taylor, T. Zwerdling, S.M. Lane, K. Ihara, and T. Huser, Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells. *Biophysical Journal*, 2006. **90**(2): p. 648-56.

37. Feng, S., R. Chen, J. Lin, J. Pan, G. Chen, Y. Li, M. Cheng, Z. Huang, J. Chen, and H. Zeng, Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis. *Biosensors and Bioelectronics*, 2010. **25**(11): p. 2414-2419.

38. Tan, C.H. and E.H. Tan, Post-treatment Assessment of Glioblastoma Multiforme: Imaging with Fluorodeoxyglucose, Sestamibi, and Choline. *World J Nucl Med*, 2012. **11**(1): p. 30-2.

39. Aksoy, C. and F. Severcan, Role of Vibrational Spectroscopy in Stem Cell Research. *Spectroscopy: An International Journal*, 2012. **27**(3).

40. Notingher, I., Raman Spectroscopy Cell-based Biosensors. *Sensors (Basel)*, 2007. **7**(8): p. 1343-1358.

41. Gniadecka, M., P.A. Philipsen, S. Sigurdsson, S. Wessel, O.F. Nielsen, D.H. Christensen, J. Hercogova, K. Rossen, H.K. Thomsen, R. Gniadecki, L.K. Hansen, and H.C. Wulf, Melanoma Diagnosis by Raman Spectroscopy and Neural Networks: Structure Alterations in Proteins and Lipids in Intact Cancer Tissue. *Journal of Investigative Dermatology*, 2004. **122**(2): p. 443-449.

42. Zhang, X., W. Zhang, W.D. Cao, G. Cheng, and Y.Q. Zhang, Glioblastoma multiforme: Molecular characterization and current treatment strategy (Review). *Exp Ther Med*, 2012. **3**(1): p. 9-14.

43. Yamada, T., N. Miyoshi, T. Ogawa, K. Akao, M. Fukuda, T. Ogasawara, Y. Kitagawa, and K. Sano, Observation of molecular changes of a necrotic tissue from a murine carcinoma by Fourier-transform infrared microspectroscopy. *Clin Cancer Res*, 2002. **8**(6): p. 2010-4.

Figure and Table Legends

Table 1 Patient details with histological subtype and metastatic origin.

Table 2 Bio-molecular assignments of main spectral peaks.

Table 3 PC-LDA model details, including spectral numbers and the resulting sensitivities and specificities.

Table 4 Spectral differences causing the separations seen in Figure 4 (b).

Table 5 The resulting sensitivities and specificities from the 2d peak ratio scores plot in Figure 6.

Table 6 PC-LDA model details, including spectral numbers and the resulting sensitivities and specificities.

Figure 1 Microscopic images of H&E stained tissue sections of normal brain, metastatic brain, primary sites and GBM brain WHO grade IV (x 100). Scale bar represents 50 microns.

Figure 2 Immersion Raman average spectrum produced through averaging 1116 immersion acquisitions using a x 50 objective. Non-immersion averages spectrum was produced using 458 non-immersion acquisitions using a x 60 objective.

Figure 3 GBM, Metastatic and normal brain tissue average spectrums, recorded using immersion Raman spectroscopy. The GBM spectrum was created from averaging 127 GBM acquisitions, the Metastatic from 668 metastatic acquisitions and the normal brain spectrum from 157 normal brain acquisitions. Spectra have been vector normalised, the backgrounds have been corrected using a 5th order polynomial fit and 7 points of smoothing. The asterisks correspond to paraffin peaks from residual wax in the tissue.

Figure 4 (a) LDA scores plot showing the separation of GBM (**blue**), metastatic (**green**) and normal brain tissue (**black**) spectra. Training spectra displayed with symbols filled and test spectra un-filled, 95% (outer ellipse) and 90% (inner ellipse) confidence limits are also shown; PC-LDA training model created using the first 12 PCs, which achieved a %CC score of 72% **(b)** LDA scores plot of the same data used in (a) but with the spectra averaged. The 20 spectra recorded from each patient tissue sample were split into 4 and averaged. The training model was created using the first 8 PCs and achieved a % CC score of 97%.

Figure 5 (a) PC-LDA loading plot of the LDA 1 separation boundary produced in Figure 4 (b) (**pink**) (upper curves) and mean spectra (lower curves), GBM (**blue**), metastatic (**green**) and normal tissue spectra (**black**) **(b)** PC-LDA loading plot of the LDA 2 separation boundary produced in Figure 4 (b).

Figure 6 2D score plot of the 620 cm^{-1} to 782 cm^{-1} peak ratio versus the 721 cm^{-1} to 620 cm^{-1} peak ratio. Each data point is the spectral patient average.

Figure 7 LDA scores plot of average spectra recorded from primary breast (**black**), primary colon/rectum (**blue**), primary melanoma (**red**) and primary oesophagus & stomach (**green**) cancerous tissue. Training spectra displayed with symbols filled and test spectra un-filled, 95% (outer ellipse) and 90% (inner ellipse) confidence limits are also shown; training model created using the first 12 PCs and achieved a % CC score of 72%.

Figure 8 LDA scores plot of average spectra recorded from metastatic breast (**blue**), colon/rectum met (**green**), lung met (**cyan**) melanoma met (**red**) and oesophagus & stomach met (**black**). Training spectra displayed with symbols filled and test spectra un-filled, 95% (outer ellipse) and 90% (inner ellipse) confidence limits are also shown; training model created using the first 9 PCs and achieved a % CC score of 63%.

Figure 9 (a) Mean spectral comparisons of primary (**blue**) and Metastatic brain tissue (**black**) (b) PCA comparisons of primary and Metastatic brain tissue recorded spectra (2d PCA scores plot displaying the best separation between Primary and Metastatic tissue). The spectra have been vector normalised, baseline corrected by 5th order polynomial fit and subtraction, and 7 points of smoothing (1) Primary Breast spectra vs. Met Breast tissue spectra (2) Primary Colon/Rectum spectra vs. Met Colon/Rectum tissue spectra (3) Primary Melanoma vs. Met Melanoma tissue spectra (4) Primary Oesophagus and Stomach spectra vs. Met Oesophagus and Stomach tissue spectra.

BTNW no.	Gender	Age	Origin of Metastasis/histological subtype
119	F	50	Breast (met and primary site)
707	F	68	Breast (met and primary site)
756	F	56	Breast (met and primary site)
888	M	65	Colon/rectum (met and primary site)
985	M	80	Melanoma (met and primary site)
988	M	64	Oesophagus & stomach (met and primary site)
1001	F	69	Lung (met only)
690 / 1012	F	53	Breast (met and primary site)
1004	F	57	Breast (met only)
998 / 1010	M	67	Colon (met only)
1020	F	65	Breast (met only)
2	F	76	GBM
3	F	32	GBM
4	M	81	GBM
5	F	55	GBM
7	M	74	Normal Brain
10	F	84	GBM
78	F	58	Lung (met only)
106	F	48	Normal Brain
132	F	70	Normal Brain
136	F	57	Normal Brain
137	M	69	Colon/rectum (met only)
164	M	59	Oesophagus & stomach (met only)
184	F	65	Lung (met only)
215	M	40	Lung (met only)
274	M	73	Colon/rectum (met only)
295	F	77	Lung (met only)
358	M	78	Melanoma (met only)
409	F	88	Colon/rectum (met only)
444	M	35	Normal Brain
509	M	83	Melanoma (met only)
517	F	58	Lung (met only)
521	M	64	Oesophagus & stomach (met only)
562	F	60	Lung (met only)
567	F	66	Melanoma (met only)
668	F	70	Colon/rectum (met only)
678	F	52	Normal Brain
688	F	48	Melanoma (met only)
694	M	62	Lung (met only)
721	F	37	Oesophagus & stomach (met only)
772	M	41	Normal Brain

Table 1 Patient details with histological subtype and metastatic origin.

Raman Shift (cm⁻¹) of peak	Tentative assignment [14, 25, 33-35, 38]
524	S-S stretch in proteins
550	Tryptophan/cytosine, guanine
620	C-C twisting mode of phenylalanine
641	C-C twisting mode of tyrosine
666	C-S stretching mode of cystine
721 (only GBM)	Symmetric choline C-N stretch (membrane phospholipid head)/adenine
725	Adenine/C-S (protein)/C-H$_2$
756	Symmetric breathing of tryptophan
782	Cytosine/uracil ring breathing of nucleotide
805	A-DNA [38]
827	DNA O-P-O, cytosine, uracil, thymine
852	(CCH) ring breathing mode of tyrosine and C-C stretch of proline ring
870	Proline
888	Paraffin wax
925	C-C stretching mode of proline [38]
934	C-C stretching mode of proline and valine and α-helix protein backbone/glycogen
957	Hydroxyapartite (PO$_4^{3-}$ symmetric stretching) /carotenoid/cholesterol
1002	Symmetric ring breathing mode of phenylalanine
1031	C-H in-plane bending mode phenylalanine
1061	Paraffin wax
1081 (only normal)	PO$_2^-$ symmetric stretching (nucleic acids) [38]
1098	C-C/C-O phospholipids
1131	Paraffin wax
1156	C-C carotenoids, C-N stretching of proteins
1171	Paraffin wax
1207	Hydroxyproline, tyrosine/tryptophan, phenylalanine (C-C$_6$H$_5$)
1234	Antisymmetric phosphate stretching [38]
1250	Amide III: (β-sheet, protein) [38]
1263 (only normal)	Amide III: (α-helix, protein) [38]
1294	Paraffin wax
1318	CH$_3$ CH$_2$ twisting mode of collagen/lipids
1340	CH$_3$ CH$_2$ wagging mode of collagen, nucleic acids
1389	CH$_3$ bend
1397	CH$_3$ bending due to methyl bond in the membrane [38]
1417	Paraffin wax
1440	Paraffin wax
1449	CH$_2$ bending mode of proteins/(CH$_3$) (CH$_2$) collagen
1462	Paraffin wax
1554	C=C Tryptophan, porphyrin

1578	Guanadine, adenine, TRP (protein)
1584	Pyrimidine ring (nucleic acids) and heme protein/C=C phenylalanine
1604	C=C in-plane bending mode of phenylalanine and tyrosine, C=C porphyrin
1618	C=C stretching mode of tyrosine and tryptophan
1658	Amide I (C=O stretching mode of proteins, α-helix conformation)/C=C lipid stretch
1670	Amide I (Proteins with β-sheet conformational structures)

Table 2 Bio-molecular assignments of main spectral peaks.

	Train spectra	Test spectra	Confidence level	Histopathological type	Sensitivity (%)	Specificity (%)
GBM vs. norm vs. met Figure 4 (a)	639	316	90%	GBM	90.48	22.26
				Normal	91.89	88.30
				Metastatic	98.08	59.09
GBM vs. norm vs. met Figure 4 (b)	159	78	90%	GBM	90.00	98.15
				Normal	76.92	96.08
				Metastatic	80.49	100.00

Table 3 PC-LDA model details, including spectral numbers and the resulting sensitivities and specificities.

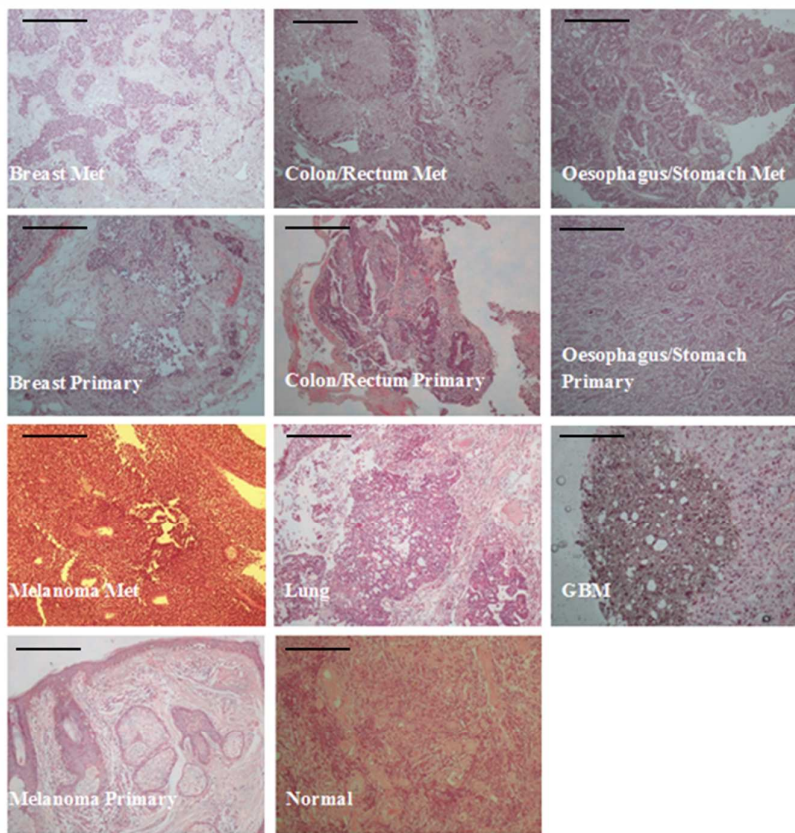| LDA1 loading bands		LDA2 loading bands	
~Negative Loading Band Position (cm^{-1})	~Positive Loading Band Position (cm^{-1})	~Negative Loading Band Position (cm^{-1})	~Positive Loading Band Position (cm^{-1})
524	1010	642	550
642	1250	719	869
718	1402	780	1078
925	1670	805	1268
1200		998	1653
		1234	
		1338	
		1375	
		1578	
		1613	
		1673	

Table 4 Spectral differences causing the separations seen in Figure 4 (b).

	Histopathological type	Sensitivity (%)	Specificity (%)
GBM vs. norm vs. met Figure 6	GBM	100.00	94.44
	Normal	96.55	100.00
	Metastatic	85.71	100.00

Table 5 The resulting sensitivities and specificities from the 2d peak ratio scores plot in Figure 6.

	Train spectra	Test spectra	Confidence level	Histopathological type	Sensitivity (%)	Specificity (%)
Metastatic comparison	448	115	90%	Breast	89.29	85.71
				Lung	60.00	72.22
				Colon/Rectum	92.31	60.78
				Melanoma	80.00	83.33
				Oesophagus & stomach	87.50	54.17
Primary comparison	111	53	90%	Breast	89.29	85.71
				Colon/Rectum	100.00	54.76
				Melanoma	71.43	38.90
				Oesophagus & stomach	85.71	100.00

Table 6 PC-LDA model details, including spectral numbers and the resulting sensitivities and specificities.

Figure 1 Microscopic images of H&E stained tissue sections of normal brain, metastatic brain, primary sites and GBM brain WHO grade IV (x 100). Scale bar represents 50 microns.

Figure 2 Immersion Raman average spectrum produced through averaging 1116 immersion acquisitions using a x 50 objective. Non-immersion averages spectrum was produced using 458 non-immersion acquisitions using a x 60 objective.
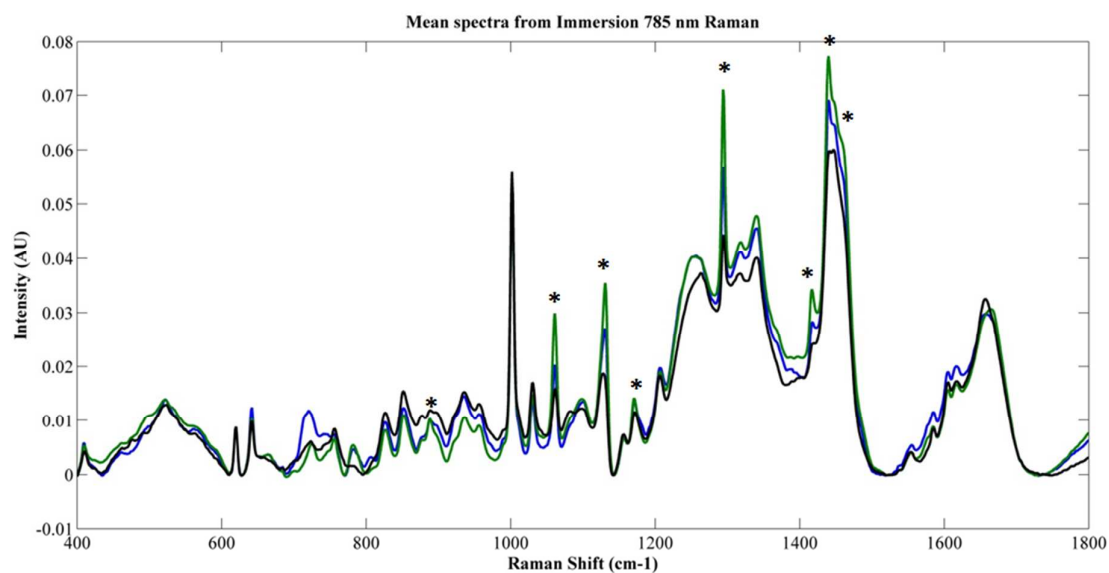
Figure 3 GBM, Metastatic and normal brain tissue average spectrums, recorded using immersion Raman spectroscopy. The GBM spectrum was created from averaging 127 GBM acquisitions, the Metastatic from 668 metastatic acquisitions and the normal brain spectrum from 157 normal brain acquisitions. Spectra have been vector normalised, the backgrounds have been corrected using a 5th order polynomial fit and 7 points of smoothing. The asterisks correspond to paraffin peaks from residual wax in the tissue.

Figure 4 (a) LDA scores plot showing the separation of GBM (**blue**), metastatic (**green**) and normal brain tissue (**black**) spectra. Training spectra displayed with symbols filled and test spectra un-filled, 95% (outer ellipse) and 90% (inner ellipse) confidence limits are also shown; PC-LDA training model created using the first 12 PCs, which achieved a %CC score of 72% **(b)** LDA scores plot of the same data used in (a) but with the spectra averaged. The 20 spectra recorded from each patient tissue sample were split into 4 and averaged. The training model was created using the first 8 PCs and achieved a % CC score of 97%.

Figure 5 (a) PC-LDA loading plot of the LDA 1 separation boundary produced in Figure 4 (b) (**pink**) (upper curves) and mean spectra (lower curves), GBM (**blue**), metastatic (**green**) and normal tissue spectra (**black**) **(b)** PC-LDA loading plot of the LDA 2 separation boundary produced in Figure 4 (b).
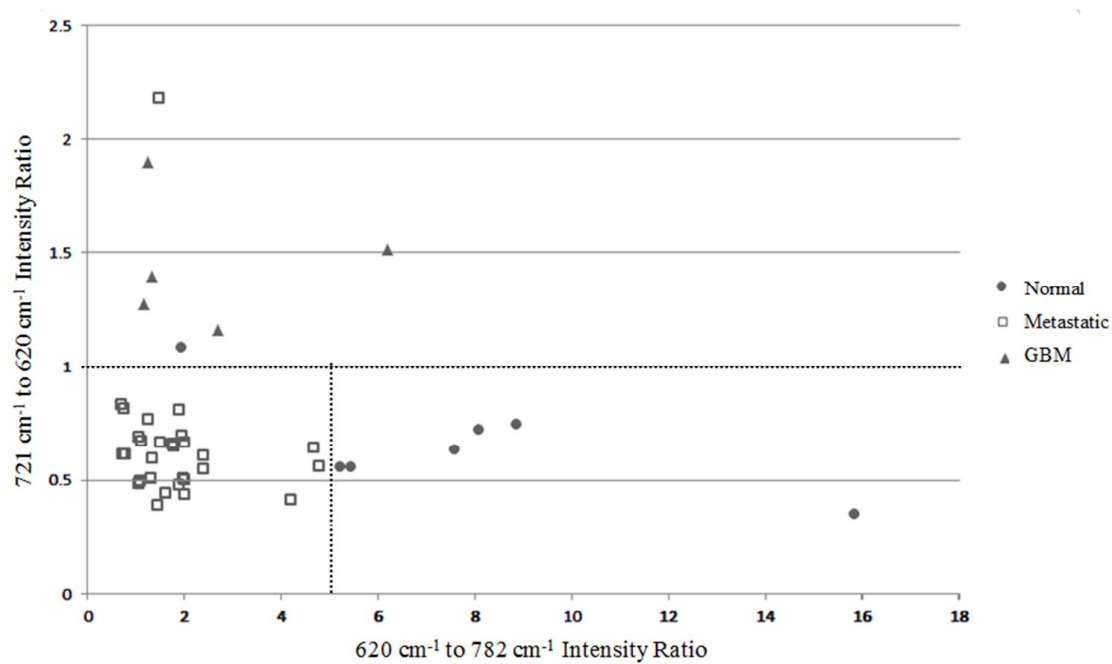
Figure 6 2D score plot of the 620 cm^{-1} to 782 cm^{-1} peak ratio versus the 721 cm^{-1} to 620 cm^{-1} peak ratio. Each data point is the spectral patient average.
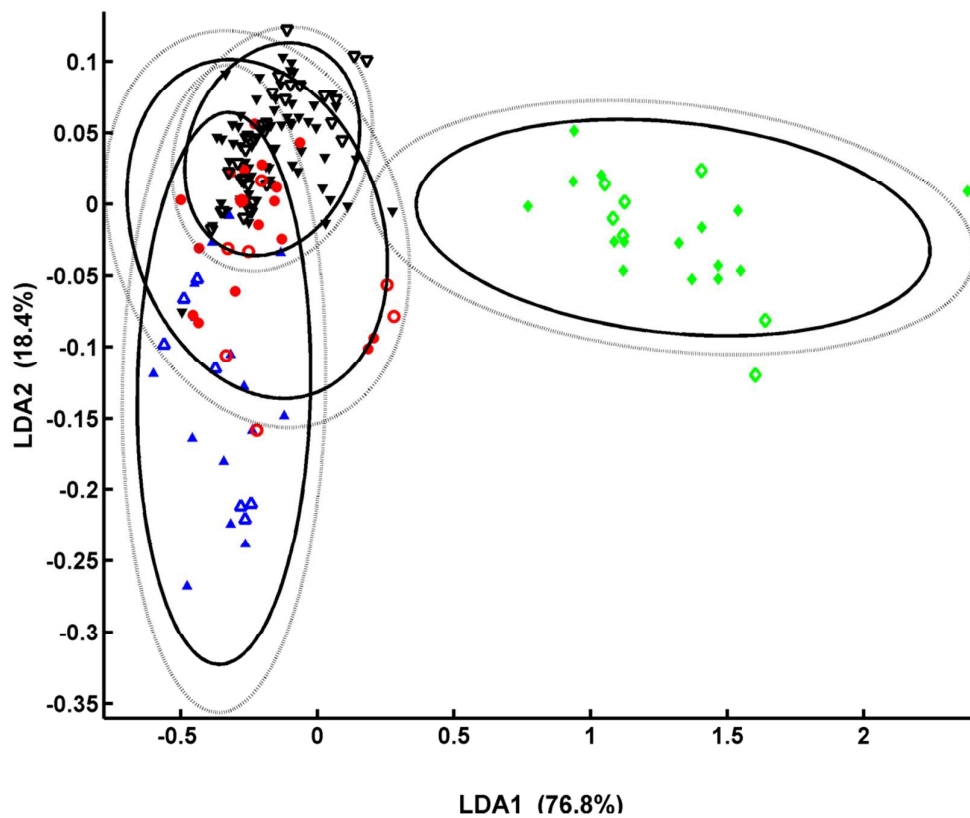
Figure 7 LDA scores plot of average spectra recorded from primary breast (**black**), primary colon/rectum (**blue**), primary melanoma (**red**) and primary oesophagus & stomach (**green**) cancerous tissue. Training spectra displayed with symbols filled and test spectra un-filled, 95% (outer ellipse) and 90% (inner ellipse) confidence limits are also shown; training model created using the first 12 PCs and achieved a % CC score of 72%.
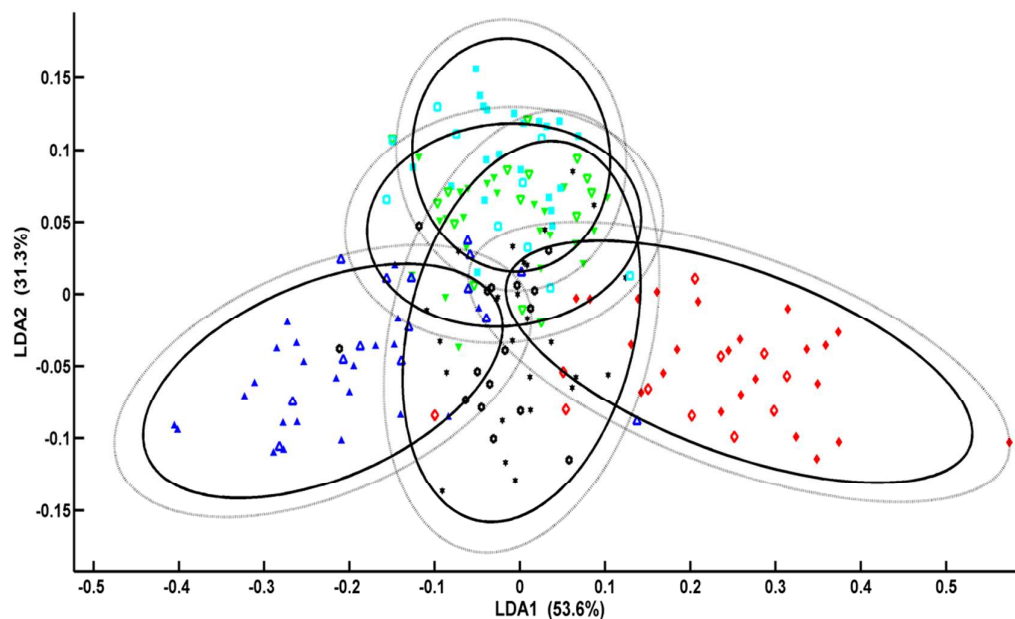
Figure 8 LDA scores plot of average spectra recorded from metastatic breast (**blue**), colon/rectum met (**green**), lung met (**cyan**) melanoma met (**red**) and oesophagus & stomach met (**black**). Training spectra displayed with symbols filled and test spectra un-filled, 95% (outer ellipse) and 90% (inner ellipse) confidence limits are also shown; training model created using the first 9 PCs and achieved a % CC score of 63%.
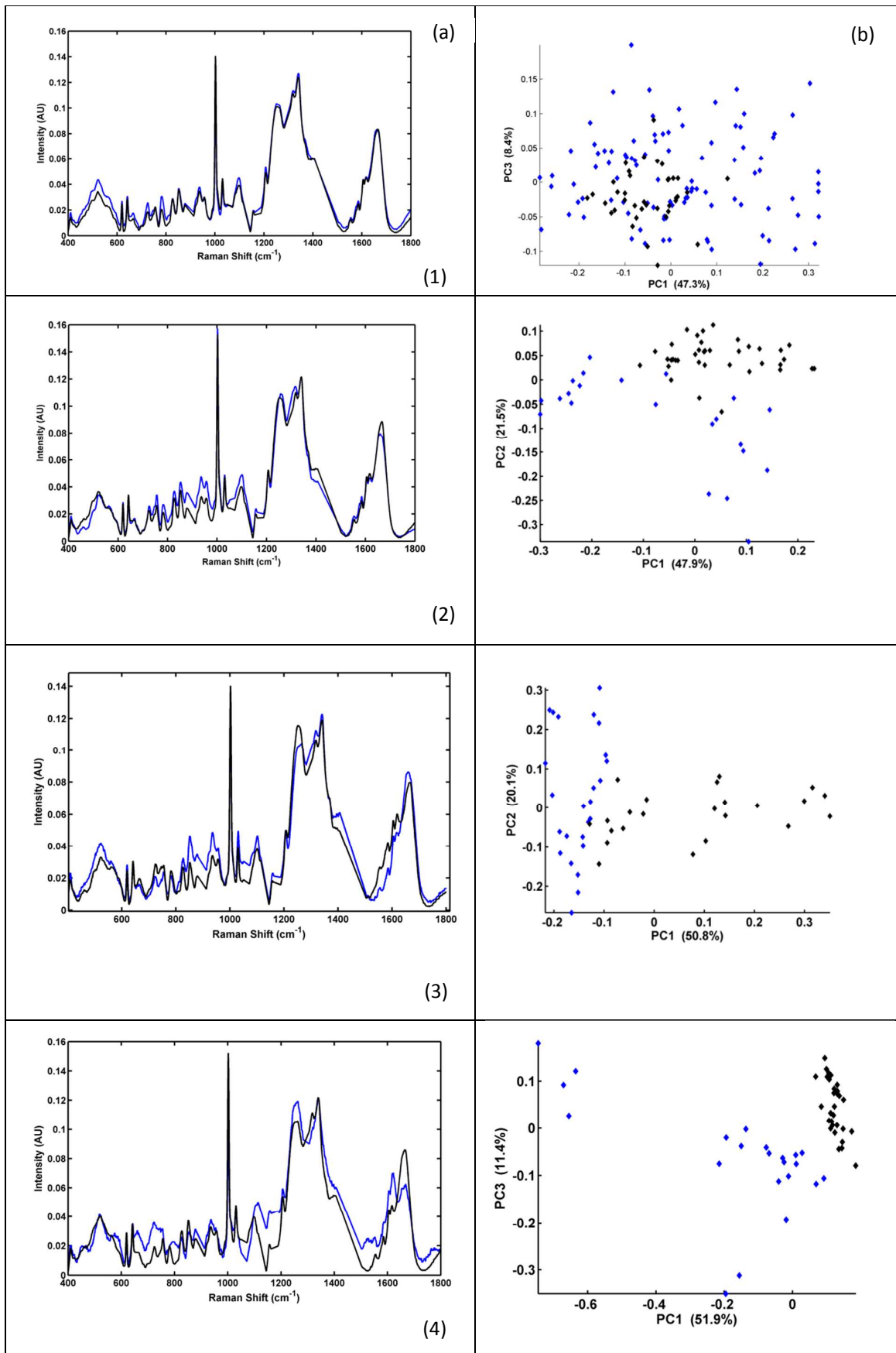
Figure 9 (a) Mean spectral comparisons of primary (**blue**) and Metastatic brain tissue (**black**) (b) PCA comparisons of primary and Metastatic brain tissue recorded spectra (2d PCA scores plot displaying the best separation between Primary and Metastatic tissue). The spectra have been vector normalised, baseline corrected by 5th order polynomial fit and subtraction, and 7 points of smoothing (1) Primary Breast spectra vs. Met Breast tissue spectra (2) Primary Colon/Rectum spectra vs. Met Colon/Rectum tissue spectra (3) Primary Melanoma vs. Met Melanoma tissue spectra (4) Primary Oesophagus and Stomach spectra vs. Met Oesophagus and Stomach tissue spectra.