Published in final edited form as:
Analyst. 2009 June ; 134(6): 1067–1076. doi:10.1039/b821166c.

Detection of breast micro-metastases in axillary lymph nodes by infrared micro-spectral imaging†

Benjamin Bird[a], **Kristi Bedrossian**[b], **Nora Laver**[b], **Miloš Miljković**[a], **Melissa J. Romeo**[a], and **Max Diem**[a]

Benjamin Bird: benjamin_bird@hotmail.com; Kristi Bedrossian: ; Nora Laver: nlaver@tuftsmedicalcenter.org; Miloš Miljković: ; Melissa J. Romeo: ; Max Diem:

[a]Laboratory for Spectral Diagnosis (LSpD), Department of Chemistry and Chemical Biology, Northeastern University, 360 Huntington Avenue, Boston, USA. Fax: +1-617-373-7773; Tel: +1-617-373-5099

[b]Department of Pathology, Tufts Medical Center, Tuft University, Boston, USA. Fax: +1-617-636-8302; Tel: +1-617-636-1035

Abstract

We report the ability of infrared micro-spectral imaging, coupled with completely unsupervised methods of multivariate statistical analysis, to accurately reproduce the histological architecture of axillary lymph nodes and detect metastatic breast cancer cells. The acquisition of spectral data from tissue embedded in paraffin provided spectra free of dispersive artefacts that may be observed for infrared microscopic measurements using a 'reflection/absorption' methodology. As a consequence, superior tissue classification and identification of cellular abnormality unattainable for deparaffinised tissue was achieved.

Introduction

The detection of breast micro-metastases (metastases measuring <2 mm in size) or isolated tumour cells (a tissue area <0.2 mm)[1,2] in axillary lymph nodes *via* standard histopathology is a formidable task, since the identification of a small cluster of cancerous cells is extremely difficult owing to their size and distinct lack of regular structure within lymph node tissue.[3] Yet the detection of these breast micro-metastases is of prime importance to correctly stage the spread of disease. If the sentinel lymph node, which has direct lymphatic connection to the primary site of cancer in the breast, is found to be free of metastatic breast cancer cells, it is thought that the spread of disease has been contained. Thus, a clinician is less likely to prescribe the extensive dissection of the axilla and remove all remaining lymph nodes, since this may lead to other serious side effects, including arm lymphodema and shoulder dysfunction.[4] However, if a micro-metastasis in a lymph node is missed, the prognosis for the patient will be incorrect, and a more aggressive treatment should have been indicated, which may include chemotherapy or in some cases radiotherapy.

Currently there are few diagnostic techniques available to clinicians that provide both accurate and effective diagnosis of axillary lymph node metastases. Methods such as ultrasound

---

†This paper is part of an *Analyst* themed issue on Optical Diagnosis. The issue includes work which was presented at SPEC 2008 Shedding Light on Disease: Optical Diagnosis for the New Millennium, which was held in São José dos Campos, São Paulo, Brazil, October 25–29, 2008.

Correspondence to: Benjamin Bird, `benjamin_bird@hotmail.com`.

imaging, mammography and lymphoscintigraphy, which do not require an invasive procedure, can only provide a tentative diagnosis. Thus, a more accurate assessment of the axilla is routinely performed during a lumpectomy or mastectomy, with lymph node sampling or clearance. More recently, sentinel lymph node biopsy is of prime focus, where it has been shown that patients who do not display a metastasis in the sentinel lymph node are unlikely to reveal a metastasis further down the chain of nodes.[5] Biopsy materials collected during these surgical procedures are subsequently scrutinised using traditional histological techniques.[6] However, many concerns remain with the true accuracy of routine histopathology. The majority of hospital pathology departments do not have the resources required to scrutinise every section that is cut from a node, and it is accepted that many micro-metastases could be missed. A more detailed study that re-examined previously diagnosed axillary lymph node tissues revealed that 7–33% of lymph nodes previously diagnosed as being negative were in fact harbouring metastases.[3] The problems associated with the detection of micro-metastases using standard histopathology can be visually gauged by inspection of Fig. 1. Panel 1(a) displays a photomicrograph captured from an entire axillary lymph node that has been H&E stained. This node is *ca*. 5 mm × 5 mm in size and displays all the normal architectural features of a healthy lymph node, including the sinuses (1) and medullary cords (2) in the medulla of the node, and the follicles in the cortex of the node (3). Panel 1(b) displays the area highlighted by the red square in panel 1(a) at 20 × magnification. At this resolution, it is easy to visualise two follicles in the cortex, with distinctly lighter stained germinal centres (1), and what would appear to be histiocytes in both the sinuses (2) and the subcapsular sinus (3). However, if we take a closer look at the subcapsular sinus region highlighted by the red square, captured at 40 × magnification in panel 1(c), we can see a very small breast micro-metastasis that has distinctly abnormal but classic breast cancer morphology. This micro-metastasis is *ca*. 10 × 2 cells or 250 μm × 50 mmin size, and lies within a healthy tissue mass of *ca*. 5 mm × 5 mm in size. Taking these statistics into account, it is clearly apparent why some small micro-metastases are initially missed by trained histopathologists.

Immunohistochemical studies of tissue, which employ protein-specific antibodies, such as antikeratin and antimucin, employ a molecular targeting strategy that can allow increased sensitivity for cellular abnormality.[3] However, this procedure is more time consuming and costly, and still relies on human microscopic inspection of the stained tissue for detection. Alternative techniques have been employed to facilitate faster intra-operative diagnosis of sentinel nodes, including imprint cytology,[7,8] and frozen section analysis.[9,10] The processing of samples is accelerated for these techniques, involving an analysis time of approximately 30–60 minutes. Yet, both approaches report wide variation in their sensitivity to detect cancerous lesions: detection levels as low as 44% and as high as 93% when compared with conventional histology.[7–13] These variations indicate that such methodologies do not solve the problems associated with screening lymph nodes.

Previously, we have reported the successful application of infrared micro-spectral imaging, coupled with Hierarchical Cluster Analysis (HCA), to characterise tissue pathology and identify large metastases within lymph nodes.[14,15] During theses investigations, spectral images were recorded from entire lymph nodes, where each pixel sampled a 25 μm × 25 μm area at the sample plane. Acquisition of spectral data at this spatial resolution provided sensitivity to the regular anatomical features of a lymph node, including the capsule, medullary cords, T-lymphocytes in the paracortex, and the activated and non-activated B-lymphocytes in the primary and secondary follicles of the cortex. In addition, different types of breast and colon metastatic cancers have been identified using this technique, which are spectrally discernable from each other, allowing the primary site of cancerous invasion to be identified before architectural features of that organ are replicated in the lymph node. More recently, we have demonstrated the application of diagnostic algorithms, trained by use of Artificial Neural Networks (ANNs), to rapidly classify spectral data recorded from frozen and deparaffinised axillary lymph nodes.[15–17] This was achieved by correlating spectral data acquired from unstained tissue to morphological interpretations made by histopathologists of stained tissue. Most prominently, the algorithms were able to classify spectral data recorded from lymph nodes that were not used to train the neural nets, and provided strong evidence that automated diagnosis by means of infrared micro-spectral imaging is possible.

However, the application of this technology to identify micro-metastases within lymph nodes provides more of a challenge. A typical metastatic cancer cell measures *ca*. 25 μm across, and small micro-metastases may only contain a few of these abnormal cells. Thus, in order to detect these malignancies, spectral data need to be recorded at a 5–10 μm pixel resolution. In collaboration with the Pathology Department at Tufts Medical Center, we are now directly investigating lymph node tissues that contain only very small breast micro-metastases in an effort to establish automated routines to detect such abnormalities in tissue. Since cancerous cells migrate to the lymph nodes *via* the lymphatic drainage of the primary tumour, they tend to settle in the subcapsular sinus, and sometimes in the capsule itself. Thus, micro-metastases are most likely found in the periphery of the lymph node. We have therefore recorded consecutive 1 mm × 1 mm infrared images from the periphery of large lymph nodes suspected of harbouring micro-metastases using a 6.25 μm pixel resolution. We shall report current developments and some novel results highlighting the diagnostic abilities of infrared spectroscopic imaging in lymph node cancer research.

Experimental

Tissue selection and preparation

Sample sections were cut from formalin-fixed paraffin-embedded tissue blocks of axillary lymph nodes obtained from the archival tissue depository at Tufts Medical Center. The investigation of tissue that has been chemically treated is unavoidable in this scenario, and is likely to have effects upon the biochemistry of the tissue. For example, during the wax embedding procedure, tissue sections are subject to a series of solvents with decreasing polarity (water, ethanol/water, xylene, paraffin). Deparaffinisation conversely involves a series of baths with increasing polarity (xylene, ethanol/water, water). This can dissolve lipids and consequently remove them from the tissue sections. Therefore the application of this sample preparation would not be recommended for the investigation of tissue morphology inside adipose-rich tissues where lipid biochemistry is vital for tissue characterisation or classification. However, despite these chemical treatments, potent biochemical information is retained as highlighted by conventional immunohistochemical studies. For example, IR spectroscopic investigations on frozen[18] and deparaffinised[19] brain tissues excised from rats displayed similar sensitivities in their ability to identify anatomical features of both healthy and diseased tissues. Tissue blocks were chosen based on the criterion that previous histological analysis had identified a breast micro-metastasis, and in some cases displayed abnormal regions of tissue with only a few metastatic cells. Specimens were cut by use of a microtome to provide tissue sections of *ca*. 5 μm in thickness and were subsequently mounted onto ‘low-e microscope slides’ (Kevley Technologies, Chesterland, Ohio, USA). These reflective substrates consist of a glass slide that on one surface contains a thin Ag coating and a transparent SnO_2 overcoat to protect the silver layer. Such substrates are almost completely transparent in the visible spectral range, but totally reflecting in the infrared. Thus infrared micro-spectroscopic measurements can be made using a transflection methodology, whereby the probing infrared beam passes the sample, is reflected by the substrate, and passes the sample again experiencing twice the attenuation of a single pass. The infrared beam is subsequently collected by the Cassegrain objective that was used to focus the beam, and is analysed in reflection geometry of the infrared microscope. Since these substrates are also transparent in the visible spectral range, visual images can be easily obtained from the same sample after spectroscopic investigation and subsequent hematoxylin and eosin (H&E) staining. However, the overwhelming advantage of these substrates is their cost (~$1 USD per slide compared with ~$200 USD for BaF_2 or CaF_2), their robustness, and also their ability to be used without modification in conventional automated instruments that prepare, stain and coverslip tissue and exfoliated cellular samples (*i.e*. cervical smear/fine needle biopsies. *etc*..) that are presently used in pathology laboratories worldwide.

In total, 20 tissue sections, each from a different patient, were collected for spectroscopic analysis. To date, spectroscopic imaging data have been collected from 12 lymph nodes, recorded whilst embedded in paraffin and also directly after deparaffinisation. This was achieved by etching a cross mark on the sample substrate using a diamond tipped pen, which allowed the same tissue regions to be scrutinised both before and after deparaffinisation of the tissue.

Spectroscopic data collection

Small micro-metastases that are making their first steps towards a major invasion of the lymph node are prevalent in the periphery of the organ. As a consequence, the periphery of each lymph node was scanned as consecutive 1 mm × 1 mm infrared images. Stained histopathological slides from the original diagnosis were also used to re-orientate and localise anatomical features in the tissue.

Infrared micro-spectral images were recorded in transflection (transmission/reflection) mode using one of two Perkin Elmer (Shelton, Connecticut, USA) Spectrum 1/Spotlight 400 instruments that incorporate an 8 × 2 element linear array detector system. Each of the photoconductive HgCdTe (MCT) detector elements in the array measures 25 μm × 25 μm in size and operates at liquid nitrogen temperature. Infrared images are acquired by rastering the sample through the focal point of the infrared microscope using an automated stage, that is synchronised to the scan speed of the interferometer and operates at 12.5 Hz. In this investigation a pixel resolution of 6.25 μm was utilised for data acquisition in an effort to identify more detailed anatomical structures within the tissue. In this mode, the area sampled by the microscope is de-magnified by a factor of 4 such that the nominal coverage of a detector is 6.25 μm × 6.25 μm, and the microscope stage is moved in increments of 6.25 μm. The spatial resolution of the infrared microscope was calibrated using a special air force resolution target deposited on a low-e slide over a layer of polystyrene, and was found to be 12 μm at 1600 cm^{-1}. Infrared images were recorded by co-adding 8 interferograms at each pixel position, using a 4 cm^{-1} spectral resolution, before Norton–Beer apodisation and Fourier transformation. An appropriate background spectrum was collected outside the sample area to ratio against the single beam spectra. The resulting ratioed spectra were then converted to absorbance. Each 1 mm × 1 mm infrared image contained 160 × 160, or 25 600 spectra, and took *ca*. 70 minutes to be acquired.

The entire instrument, including the microscope sample chamber, was continuously purged with dry air, keeping the relative humidity constant and below 5% to eliminate water vapour contribution to the observed spectra.

Data processing and analysis

The raw infrared imaging data sets were imported into and processed using software written in Matlab (version 2007b, Mathworks, Natick, MA, USA) or the CytoSpec[20] software package. Each data set consists of 25 600 individual infrared spectra, containing 1650 intensity values within the spectral range 4000–700 cm^{-1}, and occupies *ca*. 400 Mbytes of memory. As mentioned previously, infrared images were initially acquired from a tissue section whilst it was still embedded in paraffin, and later data acquisition was repeated on the same regions after the tissue had been deparaffinised. Pre-treatment of the data was therefore tailored to suit the tissue spectra type. Spectral data recorded from tissue embedded in paraffin display strong contributions from paraffin-specific bands that obscure the C–H stretching (3000–2800 cm^{-1}) and C–H deformation (1490–1350 cm^{-1}) spectral regions. Thus, data sets recorded from paraffinised tissue were initially imported into Matlab and the intensity values recorded above 1800 cm^{-1}, between 1496 and 1350 cm^{-1}, and below 900 cm^{-1} were removed. This procedure provides spectral vectors that are only composed of intensity values recorded between the spectral ranges 1800–1496 cm^{-1} and 1350–900 cm^{-1}, and consist of 379 data points. Spectral data recorded from deparaffinised tissue were imported into Matlab software and intensity values recorded above 1800 cm^{-1} and below 900 cm^{-1} removed to provide spectral vectors that are composed of 451 data points describing the spectral region 1800–900 cm^{-1}. The C–H deformation (1490–1350 cm^{-1}) spectral region was not removed for the deparaffinised tissue since close monitoring of the bands in this region did not reveal any residual paraffin for the tissues examined, and we also wished to assess whether the loss of spectral information could hinder tissue classification in paraffinised tissue. After this initial data reduction procedure, all data sets were then imported into the Cyto-Spec software and uniformly pre-processed. Initially a spectral quality test was conducted to remove all spectra that were recorded from areas where no tissue existed, or displayed poor signal to noise. All spectra that pass the test were then converted to 1st derivative spectra (Savitzky–Golay algorithm, 9 smoothing points) and vector normalised across the entire spectral vector. After normalisation the same Savitzky–Golay algorithm was passed across the data to produce 2nd derivative spectra.

Processed data sets were then subject to hierarchical cluster analysis (HCA), using the Euclidean distance to define spectral similarity,[21] and Ward's algorithm for clustering.[22] Pseudo-colour 'cluster images', that describe pixel cluster membership, were then assembled and compared directly with H&E images captured from the same sample. HCA images of between 2 and 15 clusters, which describe different clustering structures, were assembled by cutting the calculated HCA dendrogram at different levels. These cluster images were then provided to collaborating pathologists who confirmed the clustering structure that best replicated the morphological interpretations they made upon the H&E stained tissue.

Results and discussion

The results presented in Fig. 2 display examples of spectral histopathology. Panel 2(b) displays a 6-cluster image constructed from HCA analysis of a spectral data set recorded from a deparaffinised tissue section, using the spectral range 1800–900 cm^{-1}. Panel 2(a) alternatively displays the H&E image of the same tissue region stained after data acquisition. A direct comparison between the spectral characterisation of the tissue and the morphological interpretations made by a histopathologist can therefore be made. Panels 2(a) and 2(b) reveal that the 6-cluster image accurately describes the main features of the tissue region identified by a histopathologist. The red colour in the HCA image displays the edge of a significant breast micro-metastasis (5), whereas the cyan and dark blue colours describe regions of B-lymphophytes that exist both in the germinal centre of the follicle and in several other small clusters within the tissue area (2). The grey and green colours correlate with T-lymphocyte (3) and vascular regions of tissue (4), respectively, whereas the capsule is clearly demarked as yellow in the image (1). Fig. 2(c) depicts a spectral image obtained from a 15-cluster endpoint of the same HCA analysis (note that the colours assigned to clusters are not related to those of 2(b)). At this cluster endpoint a greater amount of heterogeneity is revealed within the tissue region, with the light brown cluster now assignable to the subcapsular sinus, the green cluster correlating to the activated B-lymphocytes in the germinal centre alone, and the breast metastasis forming two discrete clusters (coral and grey colours). However, more interestingly, the regions of B- and T-lymphocytes display multiple clusters. It is hard to ascertain whether these additional clusters describe the different subtypes of lymphocytes that can exist within a node, or whether they simply describe a broad natural variation of lymphocyte spectra. Such conclusions could only be drawn from a more detailed immunohistochemical study of the tissue section and its adjacent sections by use of protein-specific antibodies. However, this example clearly shows the potency of a spectral diagnosis.

Unfortunately, for some deparaffinised tissues the excellent correlation between tissue pathology and the spectral methods described previously are not always evident. Fig. 3 displays results obtained from a different lymph node that contained a smaller micro-metastasis. Panel 3(a) displays the H&E image captured from the tissue region examined. This area of tissue exhibits a large heterogeneity, and contains a small breast micro-metastasis that has taken root in the subcapsular sinus and is beginning to infiltrate the cortex of the lymph node (1). Again we can visualise the capsule (2), regions of T-lymphocytes (3) and B-lymphocytes (4), but also some scattered regions of histiocytes (5) that spread from the subcapsular sinus and surround the micro-metastasis. Interestingly, there is also a very small cluster of breast cancer cells, *ca*. 3 × 8 cells in size, which have nestled in the capsule itself (6). Panel 3(b) displays a 4-cluster image constructed from the HCA analysis of the spectral data set recorded after the tissue was deparaffinised, using the spectral range 1800–900 cm^{-1}. At this early cluster endpoint, the capsule (red and grey colours), and the interior of the node (dark green) are clearly discernable. It is also apparent that a cluster has formed which solely describes regions of sparse tissue cellularity (dark blue colour). When cutting the HCA dendrogram at different points to describe larger numbers of clusters, as shown previously in Fig. 1, it is common for all tissue types present to eventually be discerned within distinct tissue boundaries, albeit with multiple clusters in some cases. However, this was not true for this spectral data set. Panel 3(c) displays a 14-cluster image of the same HCA analysis. It was not until this cluster endpoint that any kind of contrast could be made between the micro-metastasis and the surrounding lymph node tissue. The capsule is again clearly discernable from the interior of the node, but is now described by many different clusters of spectra. Regions of T-lymphocytes are also fairly well described by the cornflower blue and dark pink colours in the image. However, only very poor correlations with histology can be made for the breast micro-metastasis as the light brown colour, B-lymphocytes as the white colour, and histiocytes as the grey colour. The dark blue cluster on the other hand correlates very well with regions upon the tissue that have very sparse tissue cellularity. Fig. 4(a) displays the mean absorbance spectra calculated for the dark blue (sparse tissue region; blue spectrum) and cornflower blue (highly cellular region of T-lymphocytes; red spectrum) clusters in the image. This spectral plot clearly reveals that the mean absorbance spectrum of the sparse tissue region is heavily contaminated with a dispersive artefact.[23] When comparing this spectrum to the mean absorbance spectrum of the T-lymphocytes, a region of high cellularity, the spectral distortions made to the sparse tissue spectrum can be more easily gauged. A large decrease to the amide I/amide II peak height ratio is observed, and the peak maximum of the amide I band is shifted to lower wavenumber. The downward-pointing feature at 1700 cm^{-1} is due to a dispersive line shape of the reflective contribution, and because of its negative intensity, shifts the amide I peak and decreases its intensity. Therefore, in this example, the HCA analysis has become confounded by dispersive contributions to the recorded tissue spectra, especially in regions of sparse cellularity, and provides many clusters within the image that are purely descriptive of the amount of dispersive contamination.

An alternative method to extract meaningful information from the recorded spectral data is to utilise only the low wavenumber region of the spectrum, between 1350 and 900 cm^{-1}. This spectral region contains several vibrations of the phosphodiester linkage found in phospholipids, DNA and RNA, and does not bear gross spectral distortions as described previously for the amide region of the spectrum. Panel 3(d) displays a 6-cluster image constructed from HCA analysis of the same spectral data set, but using the spectral range 1350–900 cm^{-1}. As can be seen in the image, the breast micro-metastasis in the capsule and subcapsular sinus is now clearly discernable from the surrounding tissues as the dark blue colour, and provides far superior correlation with histology than in the previous examples that used the spectral range 1800–900 cm^{-1}. Again the capsule of the node is clearly discerned, albeit by multiple clusters (cyan, green and pink colours) and the interior of the node is described by the red colouration. The grey cluster of pixels does bear some correlation with regions of histiocytes, but not in all locations, and is likely to be more characteristic of sparse tissue cellularity. It is also apparent that a large number of pixels have been misclassified and grouped into the cluster describing the micro-metastasis, especially within the region of B-lymphocytes, and in some areas of distinct histiocyte territory. When increasing the number of clusters in the HCA analysis it was hoped that more tissue differentiation would be accomplished. However, as shown in the HCA image of panel 3(e) that describes a 14-cluster endpoint, no further tissue differentiation was made, and the T-lymphocytes, B-lymphocytes and histiocytes were not provided any contrast, even if cluster images of up to 20 members were studied. This result indicates that although the low wavenumber region between 1350 and 900 cm^{-1} is diagnostic for cellular abnormality, highly proliferating cells such as the B-lymphocytes do bear similar spectral features in this region. It is also clear that the loss of the protein-specific bands (due to the dispersion artefact) adversely affects the ability to classify the different types of lymphocytes and histiocytes in the tissue.

As detailed previously, spectral data sets were also recorded from each tissue section whilst they were still embedded in paraffin. This additional data acquisition step was employed in an effort to reveal whether any advantages could be realised using this method. Fig. 5 displays the HCA results that were attained from analysing the spectral data set recorded from the same tissue area shown in Fig. 3, while still embedded in paraffin. The first image displayed in panel 5(a) was constructed after HCA analysis of the processed data set using the spectral ranges 1800–1496 cm^{-1} and 1350–900 cm^{-1}, and depicts a 4-cluster endpoint. By direct comparison of the H&E stained image displayed in Fig. 3(a), it can be clearly seen that even with a small number of clusters a sensitivity to the majority of the breast micro-metastatic cells is observed, shown as the dark blue colour in the image. The main body and outer areas of the capsule are depicted by the dark green and grey colours respectively, whereas the interior of the node is coloured red. The second image displayed in panel 5(b) alternatively displays a 13-cluster endpoint from the same HCA analysis. This clustering structure correlates very well with the tissue structures identified by histopathology. The main body of the capsule is characterised by the yellow pixels, whereas the light blue, white, and dark green pixels correlate very well with the B-lymphocytes, T-lymphocytes and histiocytes, respectively. The inside edge of the capsule and the beginning of the subcapsular sinus is highlighted by the cyan colour, whereas the micro-metastasis is clearly marked by the red pixels. It is also very interesting to notice there is now a second, smaller cluster of pixels, coloured dark blue, that lies on areas of the micro-metastasis that have a high cytoplasmic ratio. The remaining cluster members lie upon different areas of the outer capsule and are non-diagnostic, but nevertheless describe pixels of high spectral variance. It must also be noted that the shape and membership of the main micro-metastasis cluster, now displayed in red, did not change between Fig. 5(a) and 5(b), indicating that the spectral differences between cancerous and normal tissue are larger than the spectral differences that appear in the normal areas of the lymph node at 13 clusters. The third image, displayed in Fig. 5(c), was alternatively constructed after HCA analysis of the same data set but using the spectral range 1350–900 cm^{-1} and displays a 13-cluster endpoint. In this spectral region a variety of ‘phosphate’-carrying compounds and carbohydrates are observed, including the backbone residues of DNA and RNA. This image is very similar to the one previously shown in Fig. 5(b), and would indicate that there are concomitant spectral changes occurring in this region of the infrared spectrum for the tissue types present.

When studying the mean absorbance spectra calculated for the HCA image shown in Fig. 5 (b), it becomes apparent why the analysis of the paraffinised spectral data has provided superior results. Fig. 4(b) displays the mean absorbance spectrum calculated for the T-lymphocytes (highly cellular region of tissue; black spectrum) and the histiocytes (mostly a sparse cellularity; green spectrum). The histiocyte mean spectrum was chosen since this cluster had a large number of pixels located on regions of the tissue area that were previously identified to contain strong dispersive artefacts in the deparaffinised data set. When comparing the two spectra, it is clearly evident that the dispersive contributions seen previously are no longer apparent in the spectrum. The amide I and amide II bands in the paraffinised tissue spectra do provide a different profile to those previously shown for the deparaffinised tissue. However, no distortion to the spectra, *i.e*. a shift to lower wavenumber of the amide I peak position, or a gross decrease in the amide I/amide II band ratio, is observed for the histiocyte spectrum. This loss in spectral contamination has undoubtedly provided a greater sensitivity in tissue classification.

Fig. 5(d) displays mean 2nd derivative spectra calculated from the HCA image displayed in Fig. 5(b) that used the 1800–1496 cm^{-1} and 1350–900 cm^{-1} spectral regions of the data. The capsule of the node, displayed as the orange spectrum in the plot, and corresponding to the yellow pixels in the image, exhibits very characteristic spectral features. These include a small but noticeable shift of the amide I and amide II bands to higher wavenumber, and also a triad of peaks in the amide III region (indicated by the asterisks) that are distinctive of pure collagen, the main biochemical constituent of fibrocollagenous tissues. As previously described, a number of additional clusters were formed that depict the outer regions of the capsule tissue. These have not been displayed in the spectral plot since they merely describe the natural spectral variation observed for collagen concentration, and are characterised by variations in the peak intensity of the triad of spectral bands mentioned, and their ratio to the amide I- and amide II-specific bands. The normal tissue types within the interior of the lymph node are described by the black, royal blue and green spectra, and represent the B-lymphocytes, T-lymphocytes and histiocytes, respectively. The lymphocyte spectra display very similar spectral profiles and have some distinctive features in the spectrum. The most prominent features are broadened and more intense shoulder peaks for the symmetric phosphate band at 1086 cm^{-1}, and more notably 1060 cm^{-1}. They also display a large anti-symmetric/symmetric phosphate band intensity ratio and the most intense peaks at 965 cm^{-1}. However, small but noticeable spectral variations can be observed between the two lymphocyte types. These include an increase in the intensity of anti-symmetric and symmetric phosphate bands for T-lymphocytes, and also a distinct change to the shape and intensity of the low wavenumber shoulder of the amide II band in B-lymphocytes. The histiocyte spectrum displays the least intense phosphate-specific bands, but in contrast displays the most intense low wavenumber shoulder for the amide II band. The abnormal tissues are displayed in the lower part of the plot as the red and dark blue spectra, and depict areas identified as containing breast cancer cells. The red spectrum, which holds the greatest abnormal pixel membership, has the smallest amide I/amide II peak height ratio, the least intense low wavenumber shoulder of the amide I band, and a very small but identifiable 2 cm^{-1} upshift in the amide I peak position, from 1656 cm^{-1} seen in all non-fibrocollagenous tissue. The dark blue spectrum has a very similar spectral profile, but has a larger amide I/amide II peak height ratio, a broadening of the low wavenumber shoulder on the amide II spectral band, and a more regular amide I peak position at 1656 cm^{-1}. In summary, the abnormal tissues display a gross change to their protein composition, most notably characterised by changes to the amide I and amide II spectral bands. Increased intensities were observed for the anti-symmetric phosphate band at *ca*. 1235 cm^{-1}, which can be characteristic of increased DNA content in the absence of lipids peaks at 1735 cm^{-1}. However, the T-lymphocytes displayed the most intense anti-symmetric phosphate band, and both the B- and T-lymphocytes provided broad and more intense bands for a triad of peaks observed at 1086, 1080 and 1060 cm^{-1}, and a band at 965 cm^{-1}. As mentioned previously, a variety of 'phosphate'-carrying compounds and carbohydrates have overlapping bands in this spectral region making the assignment of each band to a specific compound hard. However, previous investigations upon tissues that were treated with both RNase and DNase indicate that the increase in intensity of such bands are observed and correlated to cells that are highly proliferating or neoplastic in nature[24,25]

Fig. 6 displays the mean 2nd derivative spectra of the capsule, T-lymphocytes and breast cancer tissues calculated from the deparaffinised and paraffinised tissue region described in Fig. 3–Fig 5. These tissues were chosen for direct comparison since they characterise the most cellular regions of the tissue, and provided spectra that appeared free of dispersive artefacts in the deparaffinised tissue. Pre-processing of the reduced data sets included vector normalisation across the entire spectral vector (1800–900 cm^{-1} for deparaffinised tissue, 1800–1496 cm^{-1} and 1350–900 cm^{-1} for paraffinised tissue). Thus, a fair comparison of the fully processed spectra can not be made since the vectors of the different tissue preparations are different lengths (379 and 451 data points, for paraffinised and deparaffinised tissue respectively) and would display intensity variations because of these changes in vector length. For that reason, the mean 2nd derivative spectra shown in Fig. 6 were normalised using a different protocol. Initially the mean derivative spectra were cut into two spectral regions, describing the wavenumber ranges 1750–1500 cm^{-1} and 1350–900 cm^{-1}. These reduced spectra were then independently vector normalised within these ranges to allow a fair comparison to be made. Fig. 6(a) and 6(b) display the resulting mean 2nd derivative spectra calculated for the 1750–1500 cm^{-1} and 1350–900 cm^{-1} spectral ranges respectively. For the capsule, the deparaffinised and paraffinised spectral profiles are identical and display only small intensity differences. However, the T-lymphocyte and breast cancer spectra display large spectral differences after deparaffinisation. These include a significant shift in the position and shape of the amide I and amide II bands, with more prominent low wavenumber shoulders for both bands, and a small shift to lower wavenumber for all bands in the 1350–900 cm^{-1} spectral range. There is also a significant loss in the intensity of the spectral band located at *ca*. 1125 cm^{-1} for both tissues, among some small, less notable variations for the breast cancer tissue. Considering the deparaffinisation process, with both xylene and alcohol washes, we postulate that these spectral variations could be caused by a change in hydration of the tissue, or even a change to the secondary structure of proteins.

In total, about 50 infrared images have been recorded from 12 lymph nodes, collected whilst in paraffin, and then repeated after deparaffinisation. Only 5 of these lymph nodes contained breast micro-metastases. This is, however, quite understandable since our tissue sections were cut some distance from the original sections scrutinised for patient diagnosis. Despite this drawback, 15 of the infrared images recorded contained small breast micro-metastases (<300 μm in size), some with small clusters of only a few abnormal cells. HCA results to date have shown very similar results as detailed previously in this manuscript, with paraffinised spectral data sets providing the same or superior tissue classification when compared to the deparaffinised data sets, and can provide a sensitivity to detect isolated cancer cells. Some examples of the paraffinised paradigm are shown in Fig. 7, with panels 7(a)–7(c) illustrating H&E stained images of 3 tissue areas examined, and panels 7(d)–7(f) displaying the same images but with the regions identified as being abnormal by HCA analysis overlayed as a semi-transparent red colour. Fig. 8 alternatively displays the mean 2nd derivative spectra from 5 different breast micro-metastases. As can be seen in the plot, the 5 mean spectra are almost completely superimposable and would indicate a spectral homogeneity not often observed for deparaffinised tissue, which can display large spectral variations for the amide I spectral band.[26,27] Similar studies that detected basal cell carcinoma within paraffin-embedded tissue were recently reported[28] that provide additional evidence that the acquisition of spectral data before deparaffinisation may provide some distinct advantages. These include a reduction in tissue preparation, stable samples less likely to degrade, and a loss of spectral artefacts as documented in this manuscript.

The application of Artificial Neural Nets (ANNs) to diagnose tissue spectra acquired from lymph node sections in a supervised fashion have been documented previously,[15–17] and displayed reliable diagnostic ability. Large breast and colon metastases were detected and contrast provided for healthy and abnormal tissue. This technique has now been extended to the diagnosis of micro-metastases and isolated tumour cells within popliteal lymph nodes of mice, and is documented in the paper by Romeo *et al*. in this special issue of *Analyst.*[27] This work is presently being extended to the data described in this manuscript for human axillary lymph nodes. The application of an ANN to classify infrared spectral data sets captured from a 1 mm × 1 mm tissue area, using a 6.25 μm × 6.25 μm pixel size, would take *ca*. 10 seconds, providing a more realistic time frame for tissue classification than HCA.

Conclusions

In this report, we document the application of infrared microspectral imaging, coupled with multivariate statistical analysis, to detect breast micro-metastases and isolated tumour cells in human axillary lymph nodes. Acquisition of spectroscopic data from tissue embedded in paraffin provided spectra that were not contaminated by dispersive artefacts, and provided superior sensitivity for cellular classification. Given the present problems associated with detecting micro-metastases by standard histopathology, and the limited molecular information available with immunohistochemical staining, it is envisaged that spectral histopathology will provide a diagnostic and potentially prognostic tool essential for clinicians.

Acknowledgements

Funding of this research from CA-111330 (to M. D.) is gratefully acknowledged.

References

1. Green, FL. AJCC Cancer Staging Manual. New York: Springer; 2002.
2. Cserni G, Bianchi S, Boecker W, Decker T, Lacerda M, Rank F, Wells CA. P. European Working Grp. Breast Screening, Cancer 2005;103:358–367.
3. Chatterjee, SJ.; Hawes, D.; Taylor, CR.; Neville, AM.; Cote, RJ. Surgical Oncology: Multidisciplinary Approach to Difficult Problems. Siberman, H.; Siberman, AW., editors. London: Hodder Education Group; 2002. p. 69-102.
4. Kissin MW, Querci della Rovere G, Easton D, Westbury G. Br. J. Surg 1986;73:580–584. [PubMed: 3730795]
5. Turner RR, Ollila DW, Krasne DL, Giuliano AE. Ann. Surg 1997;226:271–276. [PubMed: 9339933] discussion 276–278.
6. Luna, LG. Manual of Histologic Staining Methods of the Armed Forces Institute of Pathology. New York: McGraw-Hill; 1968.
7. Salem AA, Douglas-Jones AG, Sweetland HM, Mansel RE. Eur. J. Surg. Oncol 2003;29:25–28. [PubMed: 12559072]
8. Salem AA, Douglas-Jones AG, Sweetland HM, Mansel RE. Eur. J. Surg. Oncol 2006;32:484–487. [PubMed: 16569493]
9. Grabau DA, Rank F, Friis E. APMIS 2005;113:7–12. [PubMed: 15676009]
10. Esserman L, Weidner N. Cancer J. Sci. Am 1997;3:266–267. [PubMed: 9327147]
11. Gulec SA, Su J, O'Leary JP, Stolier A. Am. Surg 2001;67:529–532. [PubMed: 11409799]
12. Van Diest PJ, Torrenga H, Borgstein PJ, Pijpers R, Bleichrodt RP, Rahusen FD, Meijer S. Histopathology 1999;35:14–18. [PubMed: 10383709]
13. Galimberti V, Veronesi P, Arnone P, De Cicco C, Renne G, Intra M, Zurrida S, Sacchini V, Gennari R, Vento A, Luini A, Veronesi U. Ann. Surg. Oncol 2002;9:924–928. [PubMed: 12417517]
14. Romeo MJ, Diem M. Vibrational Spectrosc 2005;38:115–119.

15. Romeo, MJ.; Dukor, RK.; Diem, M. Vibrational Spectroscopy for Medical Diagnosis. Diem, M.; Chalmers, JM.; Griffiths, PR., editors. Chichester: John Wiley & Sons; 2008. p. 1-25.
16. Romeo, MJ.; Boydston-White, S.; Matthäus, C.; Miljković, M.; Bird, B.; Chernenko, T.; Diem, M. Biomedical Vibrational Spectroscopy. Lasch, P.; Kneipp, J., editors. Hoboken, New Jersey: John Wiley & Sons; 2008. p. 121-152.
17. Bird B, Miljković M, Romeo MJ, Smith J, Stone N, George MW, Diem M. BMC Clin. Pathol 2008;8:8. [PubMed: 18759967]
18. Amharref N, Bejebbar A, Dukic S, Venteo L, Schneider L, Pluot M, Vistelle R, Manfait M. Biochimica Et Biophysica Acta-Biomembranes 2006;1758:892–899.
19. Bambery KR, Schultke E, Wood BR, MacDonald STR, Ataelmannan K, Griebel RW, Juurlink BHJ, McNaughton D. Biochimica Et Biophysica Acta-Biomembranes 2006;1758:900–907.
20. Lasch, P. www.cytospec.com.
21. Adams, MJ. Chemometrics in Analytical Spectroscopy. Vol. 2nd edn. Cambridge: Royal Society of Chemistry; 2004.
22. Ward JH. J. Amer. Stat. Assoc 1963;58:236–244.
23. Romeo M, Diem M. Vibrational Spectroscopy 2005;38:129–132.
24. Chiriboga L, Yee H, Diem M. Applied Spectroscopy 2000;54:480–485.
25. Romeo, MJ.; Boydston-White, S.; Matthaeus, C.; Miljković, M.; Bird, B.; Chernenko, T.; LP; Diem, M. Vibrational Spectroscopy for Medical Diagnosis. Diem, M.; Griffiths, PR.; Chalmers, JM., editors. Chichester: John Wiley & Sons Ltd; 2008. p. 27-70.
26. Bird B, Romeo MJ, Laver N, Diem M. J. Biophotonics 2009;2:37–46. [PubMed: 19343684]
27. Romeo MJ, Bird B, Kuhn L, Natarajan E, Diem M. Analyst. 2008submitted.
28. Ly E, Piot O, Wolthuis R, Durlach A, Bernard P, Manfait M. Analyst 2008;133:197–205. [PubMed: 18227942]

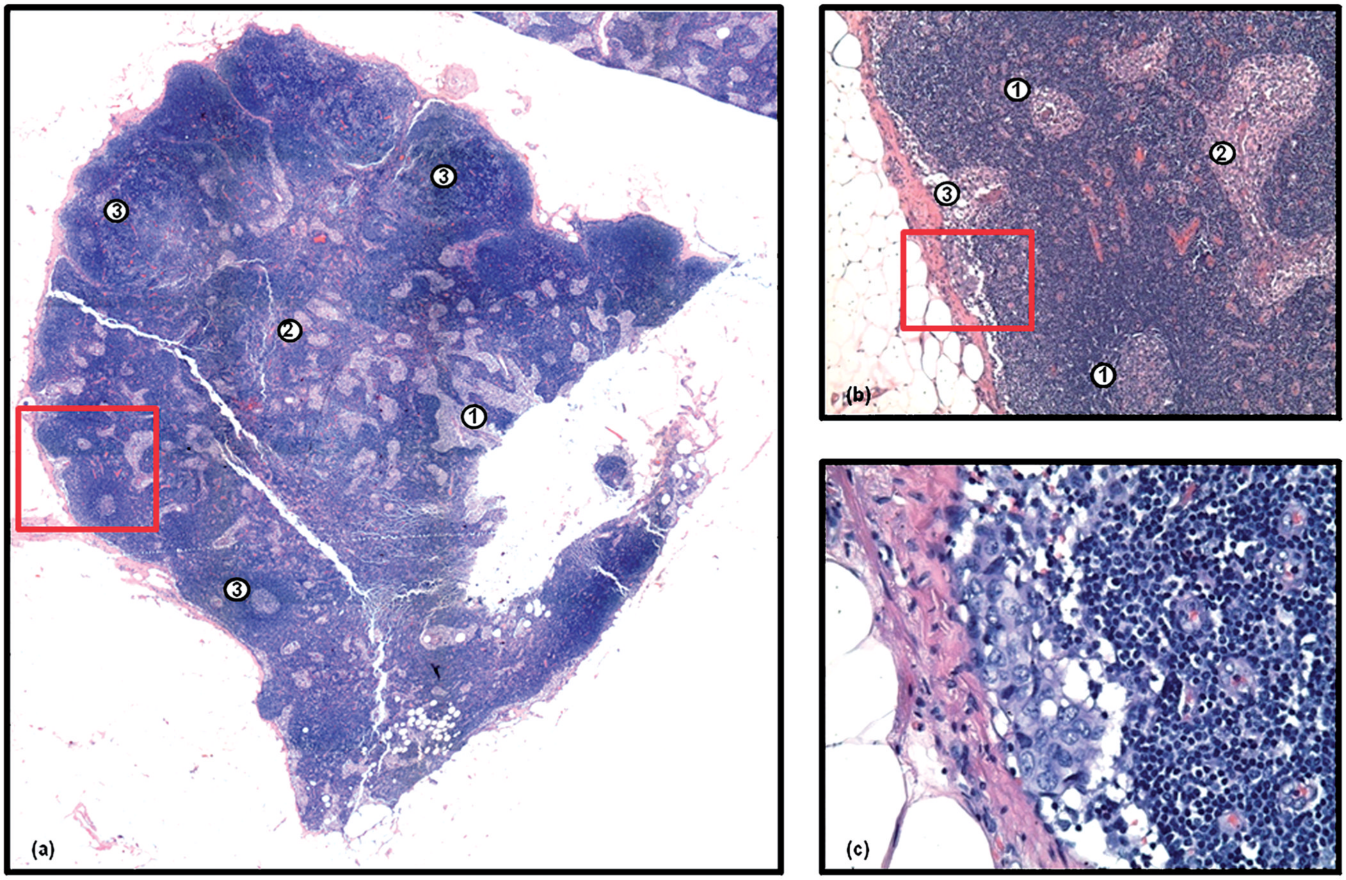

Fig. 1.
Panel (a): photomicrograph of a H&E stained axillary lymph node (*ca*. 5 mm × 5 mm in size) containing a small breast micro-metastasis. The sinuses (1), medullary cords (2) and follicles (3) of the lymph node are indicated. Panel (b): microscopic view at 20× magnification of the area indicated by the red square in panel (a). At this resolution, it is easy to visualise two follicles in the cortex, with distinctly lighter stained germinal centres (1), and what would appear to be histiocytes in both the sinuses and the subcapsular sinus (2). Panel (c): microscopic view at 40× magnification of the area indicated by the red square in panel (b). Note the appearance of a small breast micro-metastasis under the capsule of the lymph node.

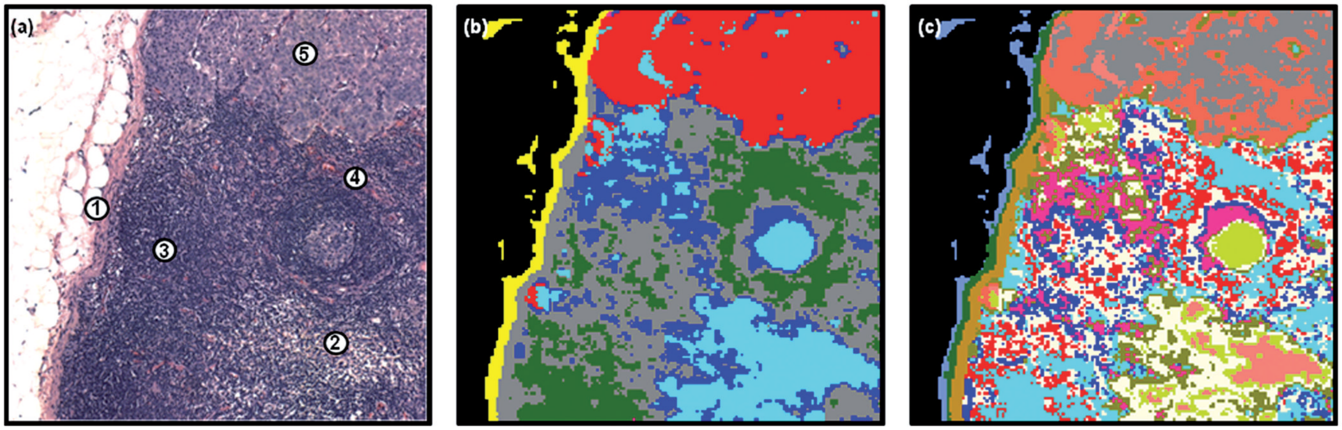

Fig. 2.

Panel (a): microscopic view of a 1 mm × 1 mm tissue area analysed by infrared micro-spectral imaging. The capsular (1), B-lymphocyte (2), T-lymphocyte (3), vascular (4) and breast micro-metastatic (5) regions are indicated. Panel (b): 6-cluster spectral image, constructed *via* HCA in the 1800–900 cm^{-1} spectral range, of the unstained and deparaffinised tissue area shown in panel (a). The red colour co-localises with the breast micro-metastasis apparent in panel (a). Panel (c): 15-cluster spectral image obtained from the same HCA.

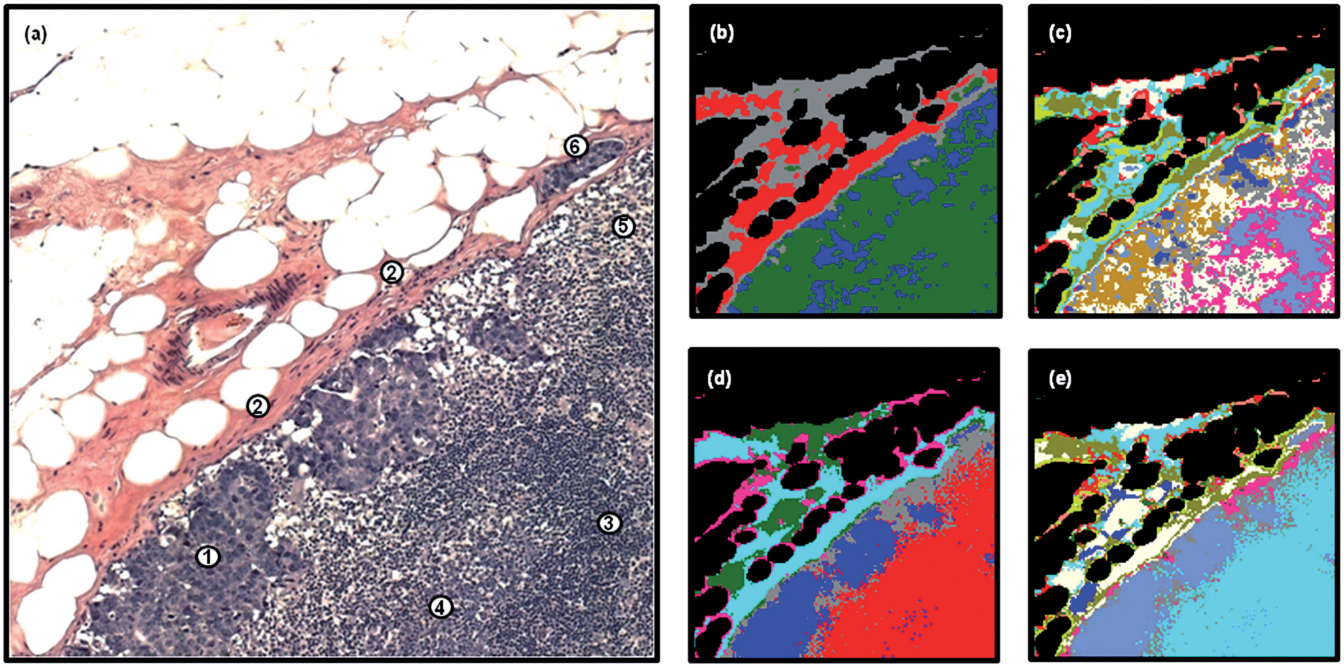

Fig. 3.

Panel (a): microscopic view of a 1 mm × 1 mm tissue area examined by infrared micro-spectral imaging. The capsular (2), B-lymphocyte (4), T-lymphocyte (3), histiocyte (5), and breast micro-metastatic (1) regions are indicated. Panel (b): 4-cluster spectral image, constructed via HCA in the 1800–900 cm−1 spectral range, of the unstained and deparaffinised tissue area shown in panel (a). Panel (c): 14-cluster spectral image obtained from the same HCA described in panel (a). Panel (d): 6-cluster spectral image, constructed *via* HCA in the 1350–900 cm^{-1} spectral range. The dark blue colour co-localises with the breast micro-metastasis apparent in panel (a). Panel (e): 14-cluster spectral image obtained from the same HCA described in panel (d).

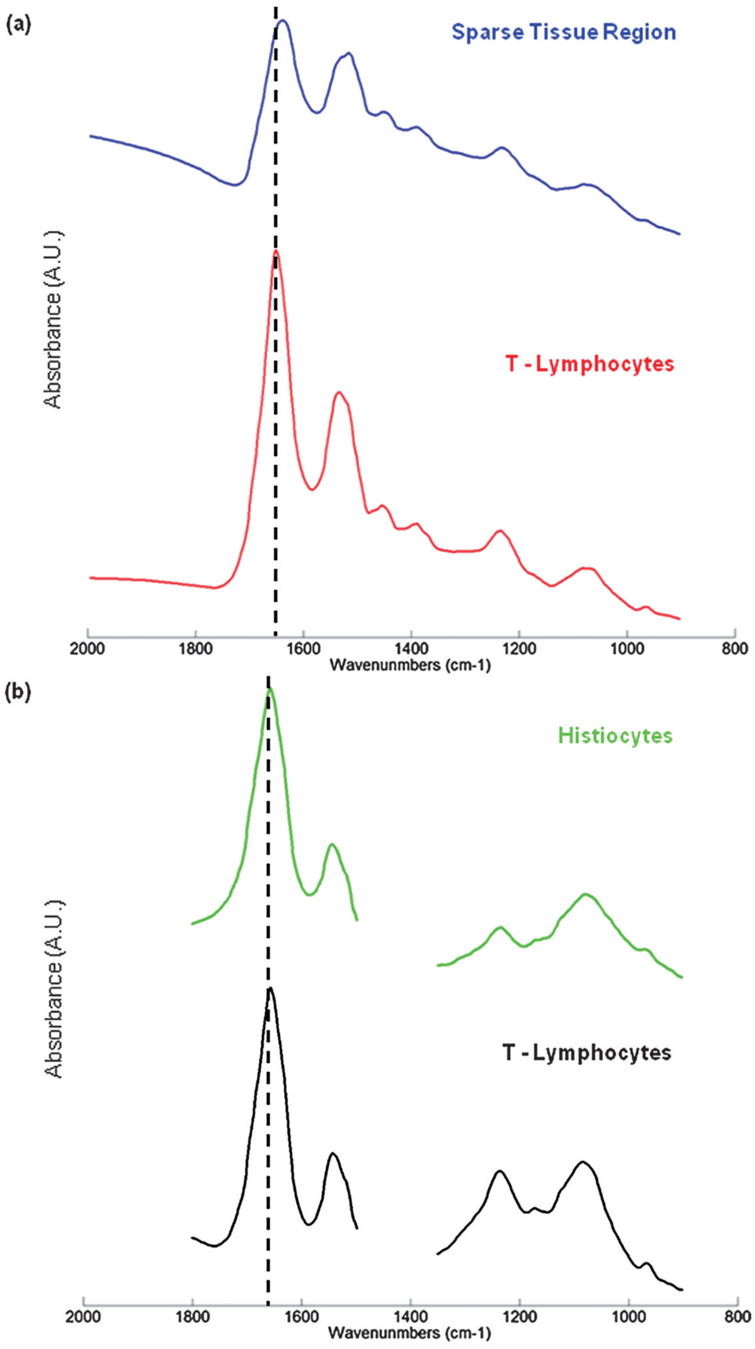

Fig. 4.

Panel (a): mean absorbance spectra calculated for the dark blue (sparse tissue) and cornflower blue (T-lymphocytes) regions of Fig. 3(c). Panel (b): mean absorbance spectra calculated for the dark green (histiocytes) and white (T-lymphocytes) regions of Fig. 5(b).

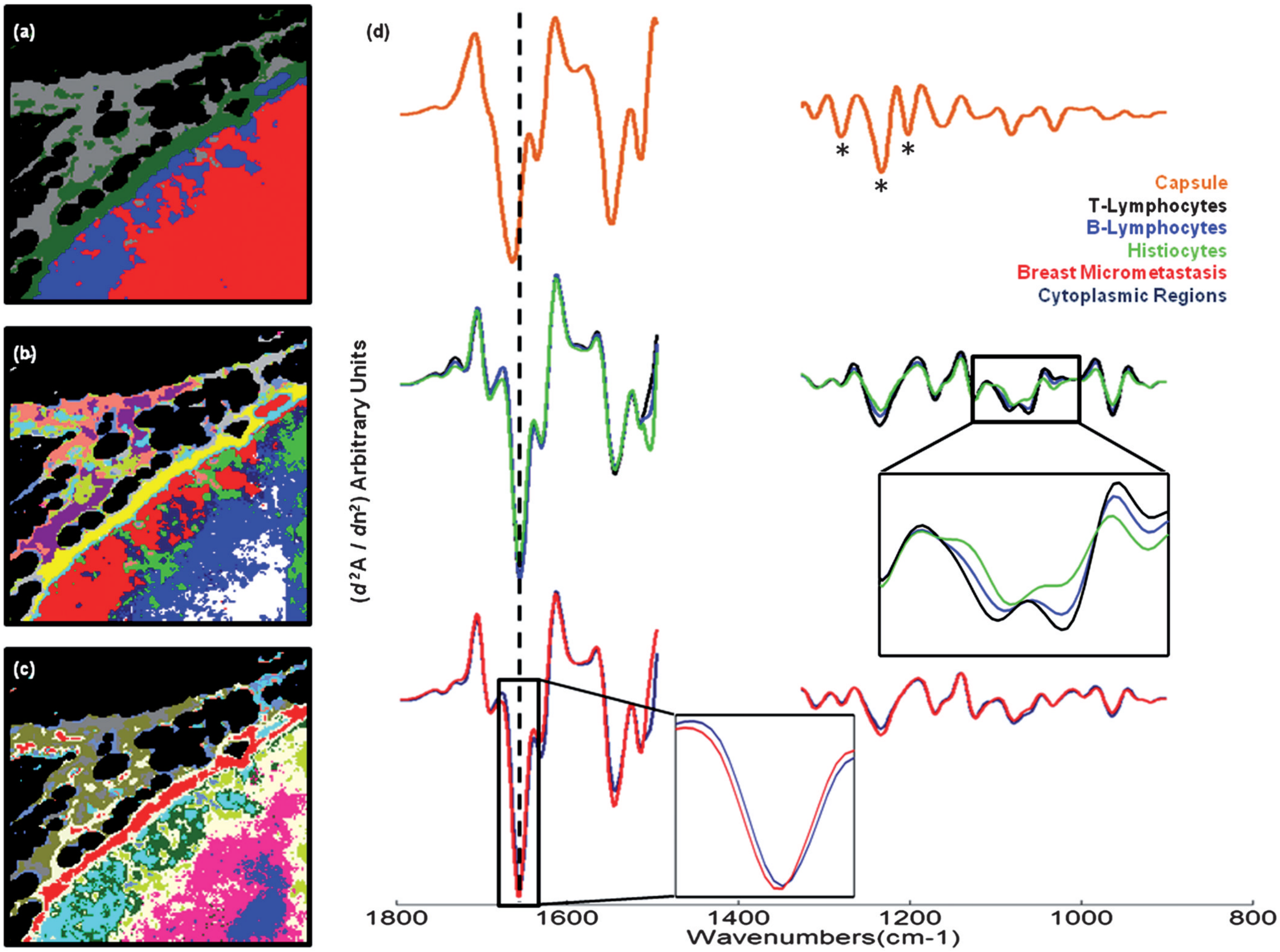

Fig. 5.

Panel (a): 4-cluster spectral image, constructed via HCA in the 1800–1496 cm^{-1} and 1350–900 cm^{-1} spectral ranges, of the unstained and paraffinised tissue area shown in Fig. 3(a). Panel (b): 13-cluster spectral image obtained from the same HCA described in panel (a). Panel (c): 13-cluster spectral image, constructed *via* HCA in the 1350–900 cm^{-1} spectral range. Panel (d): mean 2nd derivative spectra calculated for the yellow (capsule), white (T-lymphocyte), royal blue (B-lymphocyte), green (histiocyte), dark blue (breast micro-metastasis, cytoplasmic region) and red (breast micro-metastasis, highly nucleated region) regions of panel (b).

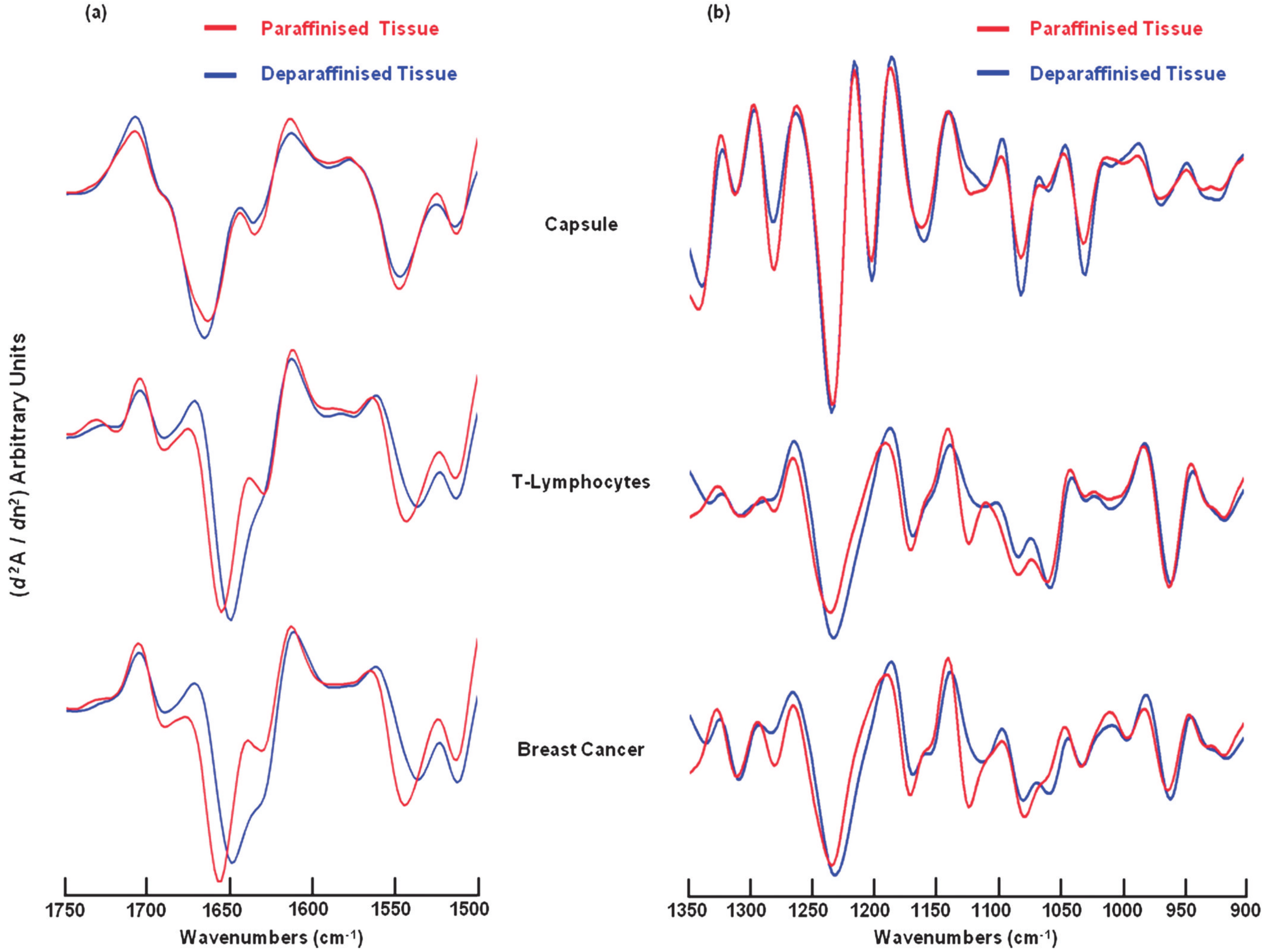

Fig. 6.

Mean 2nd derivative spectra calculated for the capsule, T-lymphocytes, and breast micro-metastasis shown in Fig. 3(c) (deparaffinised tissue) and Fig. 5(b) (paraffinised tissue). Panel (a): spectral range 1750–1500 cm^{-1}. Panel (b): spectral range 1350–900 cm^{-1}.

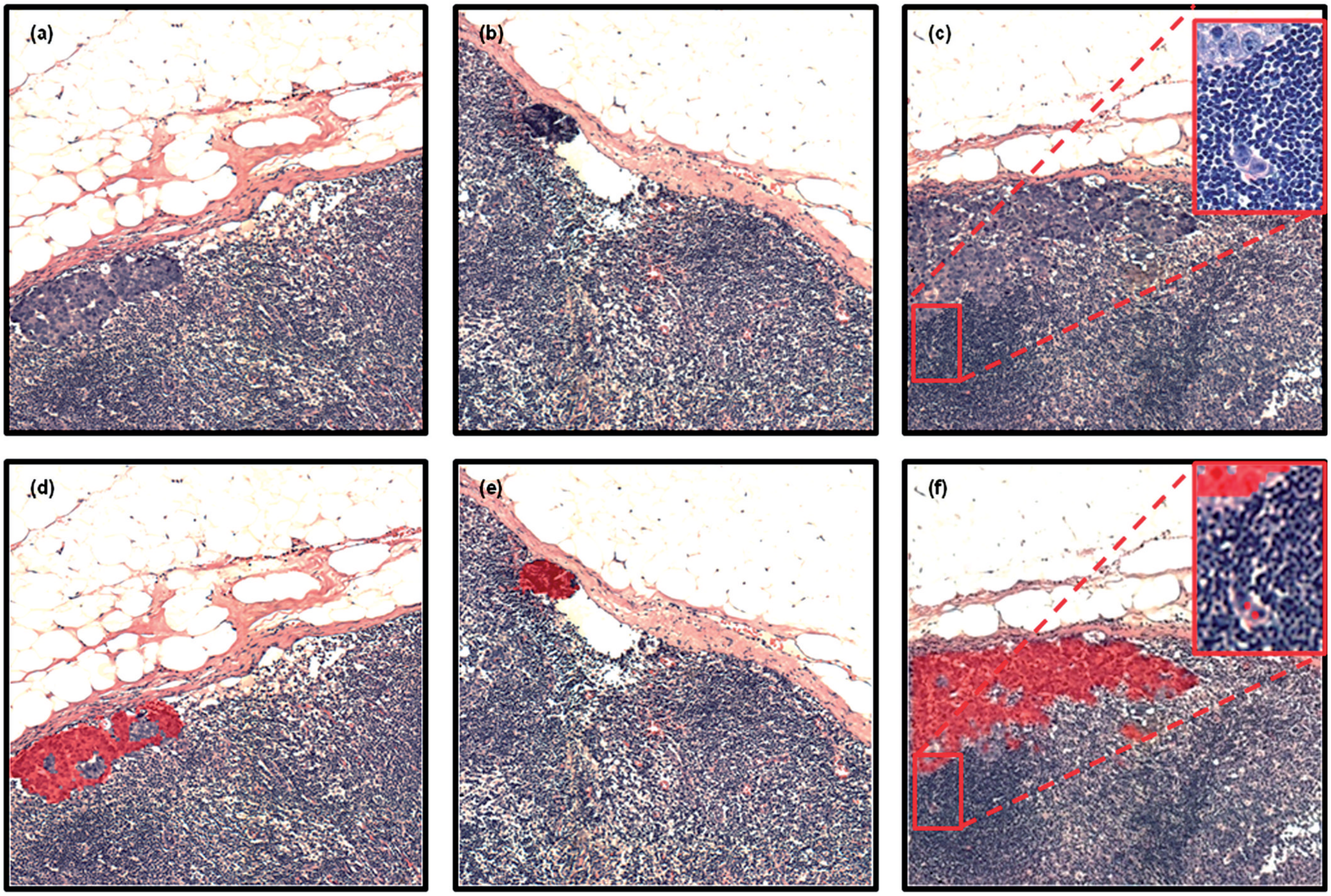

Fig. 7.
Panels (a)–(c): H&E stained images captured from 3 tissue regions (1 mm × 1 mm areas) scrutinised by infrared micro-spectral imaging. Panels (d)–(f): semi-transparent (red colour) overlays of areas identified by HCA to contain abnormal tissue. These overlays co-localise extremely well with changes in cell morphology detectable by microscopic inspection. Note the isolated tumour cells identified in panel (c) and detected by HCA in panel (f).

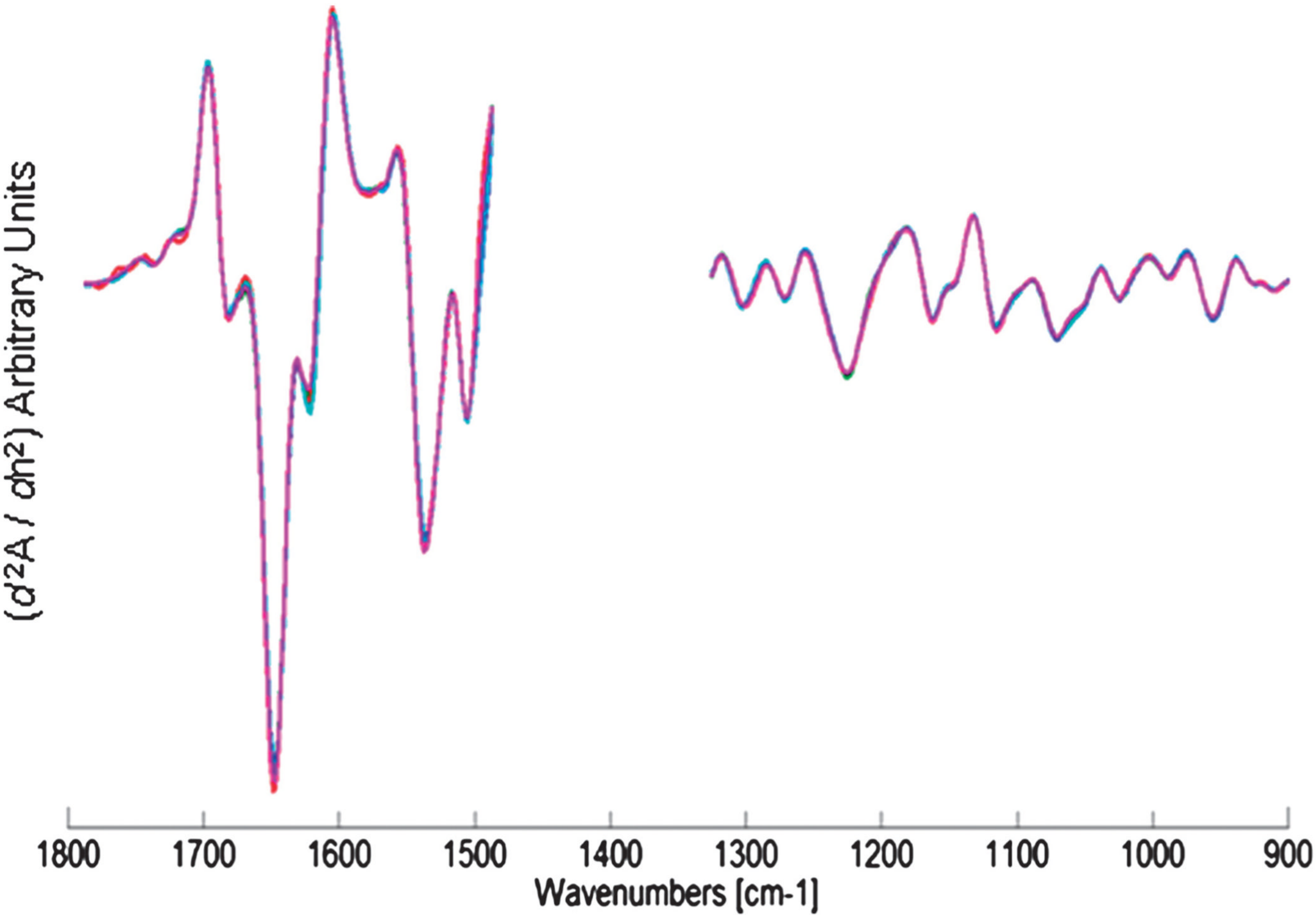

Fig. 8.

Mean 2nd derivative cluster spectra of 5 different breast micro-metastases. Spectral data were acquired whilst tissue was still embedded in paraffin. Note the spectral homogeneity.