

Novel concepts in infrared spectral imaging as a cancer diagnostic tool

Jayakrupakar Nallala,^{a,#} Olivier Piot,^a Marie-Danièle Diebold,^{a,b} Cyril Gobinet,^a Olivier Bouché,^{a,c} Michel Manfait^a and Ganesh Dhruvananda Sockalingum^{a,*}

^aMéDIAN Biophotonique et Technologies pour la Santé, Université de Reims Champagne-Ardenne, FRE CNRS 3481 MEDyC, UFR de Pharmacie, SFR Cap Santé, 51 rue Cognacq-Jay, 51096 Reims Cedex, France.

E-mail: ganesh.sockalingum@univ-reims.fr

^bLaboratoire d'Anatomie et Cytologie Pathologiques, CHU Robert Debré, Avenue du Général Koenig, 51092 Reims Cedex, France

^cService d'Hépatogastroentérologie et de Cancérologie Digestive, CHU Reims, Avenue du Général Koenig, 51092 Reims Cedex, France

Introduction

Novel methods are continuously being sought as potential diagnostic tools which can provide bio-molecular markers in a rapid and objective manner. This is extremely important, especially in biomedical research dealing with cancers in which accurate diagnosis determines the course of the treatment and, hence, the disease itself and the survival outcome. Infrared (IR) spectral imaging has shown promising results in this field due to its capacity to provide bio-molecular information from cells and tissues in a non-destructive and label-free manner.¹ Several recent studies have shown its capability to differentiate between normal and tumoural cells and tissues in different types of cancers; these have led to new concepts of spectral histopathology and cytology.^{2,3}

IR spectral histopathology involves the imaging of tissues and the construction of spectral images from which spectral information can be obtained and correlated to the bio-molecular profiles of the tissue. Extraction of this informa-

tion is facilitated with the use of various multivariate data pre-processing and processing methods that can be implemented to provide valuable diagnostic information.

One of the recent developments in the field of IR spectral imaging is the concept of IR spectral barcoding, which has the potential to provide clinically significant information for cancer diagnosis and monitoring and could enhance and provide new insights into various applications of IR spectral imaging.⁴ This concept has been developed by combining IR spectral imaging with statistical tools wherein the spectral barcodes represent discriminant spectral features between the different tissue types analysed. The main objective of this review article is to summarise the potentials of this concept to a variety of applications concerned with cancer diagnosis.

Instrumentation and image acquisition

IR spectral micro-imaging has evolved rapidly over recent decades in parallel with several advances in various instrumental features. The current commercially available bench-top IR imaging instruments consist of a microscope coupled to a Fourier transform infra-

red (FT-IR) spectrometer, which allows one to visualise and acquire spectral images from tissue sections. Using a motorised stage, a region of interest can be selected using a white light illumination source while the IR radiation is projected onto the same region using an IR source such as a Globar. Images can be obtained using a pixel size as small as $6.25\ \mu\text{m} \times 6.25\ \mu\text{m}$ for a linear, or even $ca\ 3\ \mu\text{m} \times 3\ \mu\text{m}$ using a focal plane array (FPA) 64×64 multi-element detector. With the advent of these multi-element array detectors, larger maps can be obtained in smaller time frames and with increased pixel resolution. A typical IR spectrum at any pixel element consists of wavenumbers represented on the x-axis and the intensity of absorption (or % transmission) at each wavenumber on the y-axis.

Figure 1 illustrates a methodology employed for IR spectral histopathology of colon tissue specimens procured in the form of paraffinised tissue arrays that were stabilised in an agarose matrix. In this procedure, a single spot of the tissue array block section of $10\ \mu\text{m}$ thickness, placed on a calcium fluoride (CaF_2) substrate, corresponding to a non-tumoural colon tissue, is directly imaged using

*Current address: Biomedical Imaging and Biosensing, University of Exeter, Biomedical Physics, School of Physics, EX4 4QL, UK.

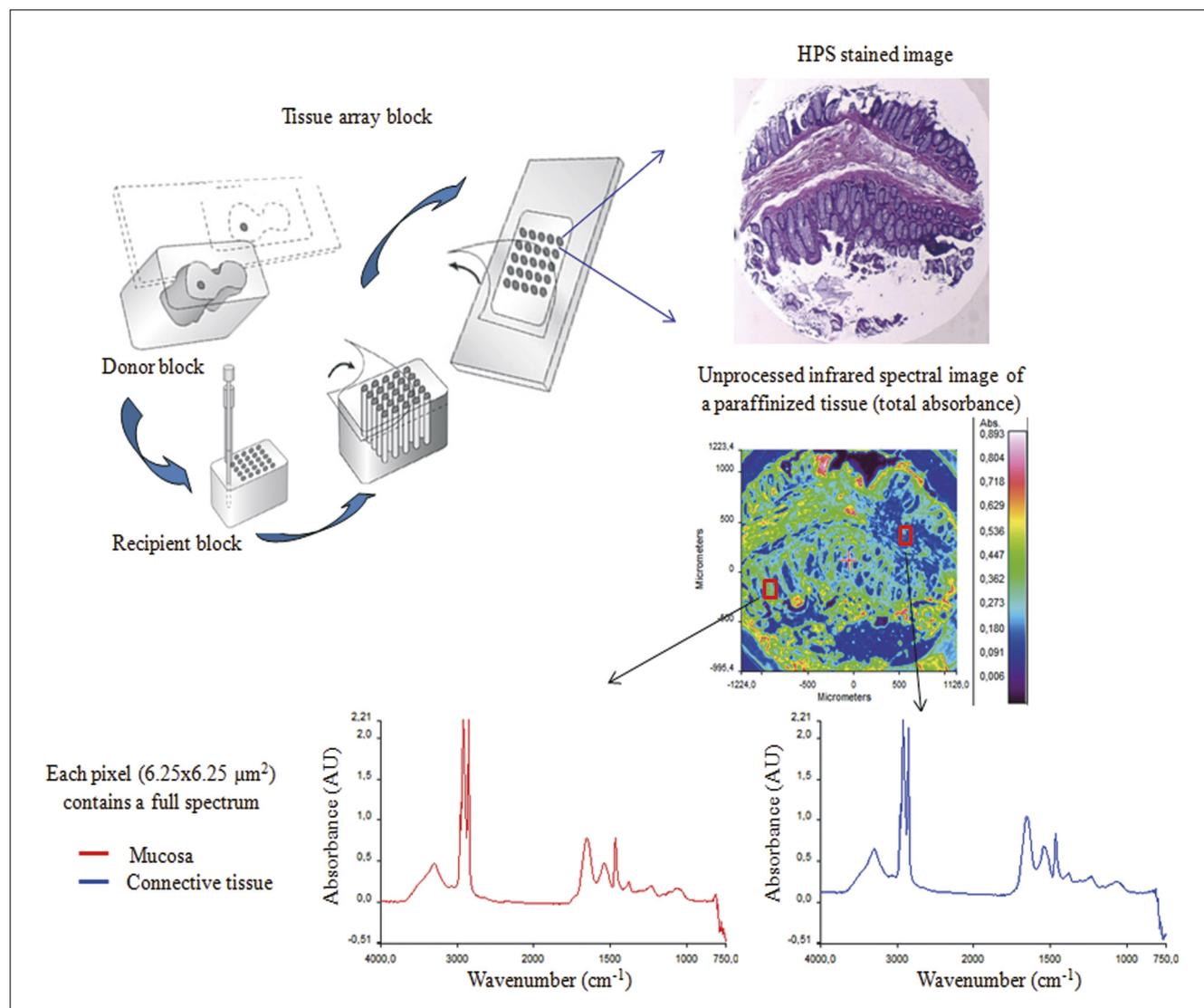


Figure 1. Infrared spectral imaging methodology of colon tissue arrays. A paraffinised tissue array core is imaged directly by an infrared imaging system that constitutes the unprocessed infrared spectral image, which harbours a full spectrum at each pixel size of 6.25 μm × 6.25 μm, using a conventionally stained image as a morphological reference. Reproduced with permission from the paper by J. Nallala, C. Gobinet, M.D. Diebold, V. Untereiner, O. Bouché, M. Manfait, G.D. Sockalingum and O. Piot, "Infrared spectral imaging as a novel approach for histopathological recognition in colon cancer diagnosis", *J. Biomed. Opt.* **17(11)**, 116013, (2012).

IR spectral imaging without any prior chemical deparaffinisation protocols. The resulting pseudo colour-coded IR spectral image (total absorbance) is represented corresponding to its reference histological image (adjacent section) stained with haematoxylin, phloxine and saffron (HPS). The spectra originating from mucosa (red) and connective tissue (blue) of the colon tissue represented in the image shown below the HPS-stained image correspond to a single pixel element

of 6.25 μm × 6.25 μm of a nitrogen-cooled 16-element mercury, cadmium and telluride (MCT) linear array detector. These spectra were acquired at a spectral resolution of 4 cm⁻¹ in the mid-IR range of 750–4000 cm⁻¹.

Pre-processing

Before the spectra can be analysed, a pre-processing step is essential in order to remove various interfering effects occurring during image acquisition due to the inherent properties of light and

its interaction with different materials. Procedures for baseline correction, normalisation and smoothing are the most commonly employed pre-processing steps. Alternative methods such as extended multiplicative signal correction (EMSC) are specific pre-processing methods that permit one to neutralise the paraffin interferences mathematically and to exploit only the bio-molecular information from the tissue. Rather than direct subtraction of spectral signatures of paraffin, the EMSC

algorithm employs a modelling procedure to neutralise the influence of the spectral variability arising from paraffin. Therefore, in the image analysis, only the spectral variability originating from the tissue biochemical features is taken into account rather than the addition of those from paraffin features. This approach, therefore, allows direct imaging of paraffinised tissues without the need to use any chemical deparaffinisation procedures, thereby reducing analysis time, costs and avoiding toxic chemical treatments. EMSC allows for the correction of the raw IR data for various spectral interferences at the same time, including baseline and normalisation as well as mathematical deparaffinisation.

Processing and data analysis

Different data processing methods are available to analyse the large spec-

tral data sets generated in IR imaging. One such method is cluster analysis, which allows one to partition IR spectral data into their constituent bio-molecular features based on their spectral distance. Using these results, pseudo-colour images can then be reconstructed, depicting the original histology of the tissues and be correlated to gold-standard histological images en-route towards spectral histopathology.

As an example, Figure 2 shows cluster analysis image results for a non-tumoural and a tumoural frozen colon tissue section (middle panels) in comparison to their reference haematoxylin and eosin (HE) stained sections (left panels). The IR spectral images were subjected to an unsupervised *K*-means clustering analysis using 8 and 12 cluster numbers, respectively, in order to partition various tissue features based on the bio-molecular information. It can be observed that the cluster analysis of

the non-tumoural colon tissue reveals the prominent histological features of a normal colon tissue such as normal epithelium (Clusters 1 and 8) corresponding to the outer nuclear part and the inner cytoplasmic part of the crypts, the lamina propria (Cluster 6) which is the loose connective tissue surrounding the crypts, the submucosa (Clusters 4 and 7), the muscularis mucosa (Cluster 2) and the secreted mucus (Cluster 3). Cluster 5 may be associated with the outer part of the crypt. Conversely, partitioning of the tumoural tissue needed 12 classes, of which Clusters 2 and 5 could be attributed to the malignant epithelium (moderately differentiated adenocarcinoma) and Cluster 10 to the tumour-associated stroma. The remaining clusters appeared to be associated with the surrounding stromal tissue. It should be noted that in the tumoural tissue, the normal histological features were no longer discernible in either the

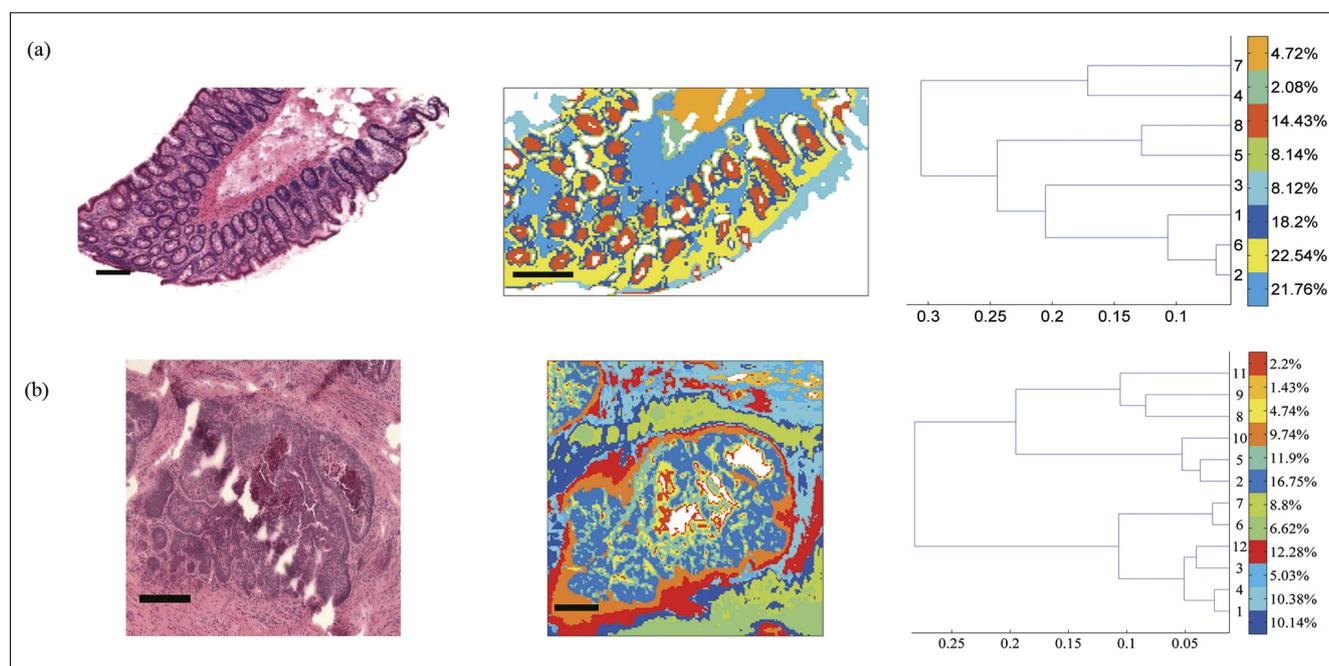


Figure 2. *K*-means clustering of non-tumoural and tumoural colonic FT-IR spectral images (middle panels) with the respective dendrograms (right panels) compared to the HE-stained sections (left panels). (a) Non-tumoural colonic tissue section clustered using eight clusters representing the major normal colonic tissue features by random pseudo-colours. The representation is as follows: Clusters 1 and 8, normal epithelium (peripheral and central parts of the crypts); Cluster 2, muscularis mucosa; Cluster 3, secreted mucous; Clusters 4 and 7, submucosa; Cluster 6, lamina propria; and Cluster 5, associated with the peripheral parts of the crypts. (b) A moderately differentiated adenocarcinoma of a colon tissue section classified using 12 clusters representing the important tumoural tissue features by random pseudo-colours. The representation is as follows: Clusters 2 and 5, tumour epithelial component; Cluster 10, stromal tissue. Remaining clusters are not attributed to any histological class. Scale bar=200 μ m. (Reproduced with permission, from the paper by J. Nallala, O. Piot, M.D. Diebold, C. Gobinet, O. Bouché, M. Manfait and G.D. Sockalingum, *Cytometry Part A* **83A**, 294300 (2013).

histological or the IR cluster images. In addition to this information, the spectral distance between different clusters can be visualised in the form of dendrograms shown on the right of Figure 2. As an example, the class “tumour-associated stroma” can be seen to be closely related to the tumour, indicating their close spectral characteristics.

The cluster analysis results can be combined with statistical tests in order

to obtain specific bio-molecular information from the IR spectral signatures of the tissue. One such recent and novel application using this procedure was the exploitation of the cluster analysis results in combination with the Mann–Whitney U test to develop the concept of IR spectral barcoding. IR spectra from two different tissue types can be compared using the statistical test and their discriminant features (wave-

numbers) can be identified. A gradient can be obtained for these discriminant features ranging from highly discriminant to less discriminant based on the significance levels. These features can then be colour-coded and represented in the form of a barcode which reflects the bio-molecular changes associated with the compared classes and gives a direct and easy-to-interpret profile of the tissues examined.

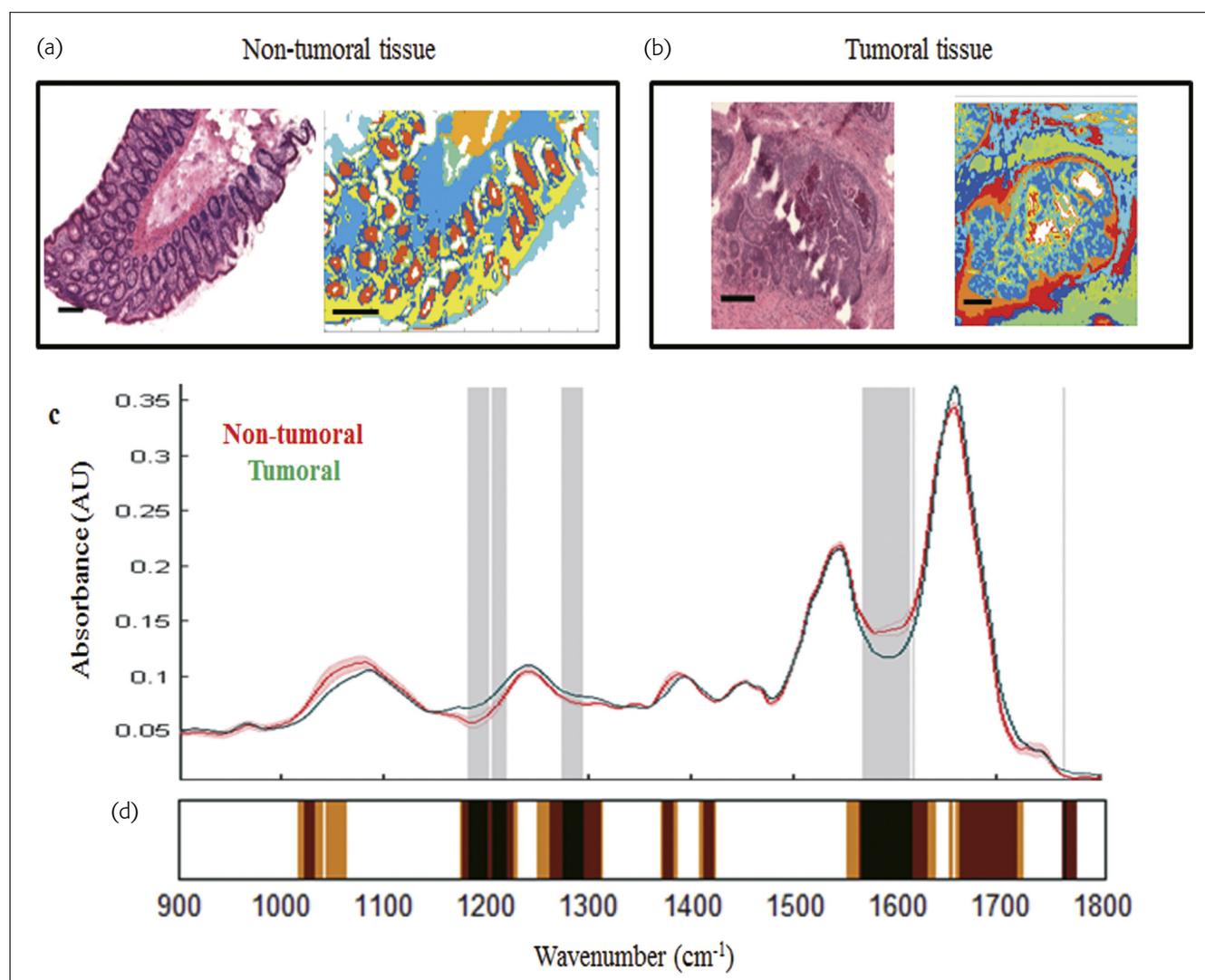


Figure 3. Construction of spectral barcode. IR spectra (using the same colour code for the IR spectra as shown in Figure 2) corresponding to non-tumoural and tumoural epithelial components from (a) and (b) are retrieved from the K -means cluster images and compared using a statistical test with different p -values of significance, to find out the significant discriminant wavenumbers (c). As an example, the discriminant wavenumbers at $p \leq 0.00016$ are shown as grey bars in (c), which are later represented, in the form of spectral barcodes (black colour) in (d). At the same time, the discriminant wavenumbers using other p -values of significance ($p \leq 0.001$ and $p \leq 0.01$) are also represented by different colours (brown and orange, respectively). Finally, all the discriminant wavenumbers using different p -values of significance are sorted out in a gradient represented by different colours which constitute the spectral barcode (d) representing the discriminant biomolecular features. Scale bar = 200 μm . Reproduced with permission, from the paper by J. Nallala, O. Piot, M.D. Diebold, C. Gobinet, O. Bouché, M. Manfait and G.D. Sockalingum, *Cytometry Part A* **83A**, 294300, (2013).

Figure 3 shows a schematic representation for identification of discriminant features and construction of IR spectral barcodes between a non-tumoural and a tumoural tissue over the IR spectral range of 900–1800 cm^{-1} . In this example, the clustering results (Figure 2) of a non-tumoural and a tumoural sample are exploited to construct the barcode. For this, IR spectra from the non-tumoural epithelial component (crypts) are compared to the IR spectra from the tumoural epithelial component (moderately differentiated adenocarcinoma) from the same patient using a Mann–Whitney U test using different p -values of significance. The discriminant features thus identified are represented as: the most discriminant in black ($p=0.00016$), the intermediately discriminant features in brown ($p=0.001$) and the least discriminant features in orange ($p=0.01$).

From the barcode, a correlation between the occurrence of the discriminant wavenumbers and the bio-molecular alterations in cancerous condition can be obtained. As an example, the most discriminant wavenumbers in black can be correlated to the alterations due to some of the important biomolecules such as proteins involving the amide I and the amide II vibrational bands and carbohydrates. This simple and rapid approach of spectral barcodes holds important prospects in providing valuable information with diagnostic significance especially in cancers.

Conclusions and perspectives

Currently and as shown in Figure 3, the concept of spectral barcode has only been applied on non-tumoural and tumoural epithelial counterparts of a colon tissue to highlight the bio-molecular and, hence, the spectral variations. However, this approach can be extended to other tissues and tissue associated features such as tumour- and tumour-associated stroma, tumours with inflammatory components etc. It would be interesting to see how the barcodes vary in these conditions and how they may be helpful in understanding how these

interactions change the tumoural environment. One of the other interesting potential applications of this approach would be to construct barcodes for different stages of tumour progression such as normal, adenoma and adenocarcinoma, where-in the bio-molecular changes associated with each of these stages can be observed. At the same time, barcodes for different tumour grades that increase with the complexity of the cancer progression could provide important information. Furthermore, spectral signatures in the form of barcodes can be constructed for primary and secondary tumours with the same normal counterpart that can reveal valuable information on the metastatic properties of the tumour. The spectral information obtained in this easy-to-interpret manner could help pathologists to complement the current gold-standard histopathology by combining the most commonly used morphological visualisation of tissues in histopathology with the bio-molecular spectral imaging information. Using this perspective, it thus becomes essential in the long term to exploit such information in order to develop algorithms and prediction models that could be used to objectively diagnose cancers and also to automate diagnosis. However, to be able to achieve this, several new studies on IR spectral imaging involving different tissue types and on a much larger scale need to be undertaken in order to construct large spectral data banks and archive them. One such way would be to construct spectral barcodes from paraffinised tissues through access to large tissue banks.

There are a few important challenges remaining which need to be considered in developing the concept of spectral barcodes such as the inherent heterogeneity arising from different patients and different tumours according to their genotypes, localisation etc. which can introduce differences in the spectral profiles. Taking these challenges into consideration, it could be envisaged that IR imaging-based spectral barcodes could provide valuable molecular information complementary to histo-

pathology for diagnostic purposes. It is important, at this point, to also mention that spectral barcodes hold important prospects for IR spectral studies at the single cell level. To mention a few, identification of spectral markers for disease characterisation, cell typing, looking at cell function and interaction with drugs which can be carried out in an easy and rapid manner. The concept of spectral barcodes could also be enriched using Raman spectroscopy or using other modes of IR spectroscopy such as attenuated total reflection (ATR) which can provide complementary information to IR barcodes and which can also be extended to *in vivo* analysis of tissues.

References

1. R. Bhargava, "Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology", *Anal. Bioanal. Chem.* **389(4)**, 1155 (2007). doi: [10.1007/s00216-007-1511-9](https://doi.org/10.1007/s00216-007-1511-9)
2. J. Nallala, C. Gobinet, M.D. Diebold, V. Untereiner, O. Bouché, M. Manfait, G.D. Sockalingum and O. Piot, "Infrared spectral imaging as a novel approach for histopathological recognition in colon cancer diagnosis", *J. Biomed. Opt.* **17(11)**, 116013 (2012). doi: [10.1117/1.JBO.17.11.116013](https://doi.org/10.1117/1.JBO.17.11.116013)
3. J.M. Schubert, B. Bird, K. Papamarkakis, M. Miljković, K. Bedrossian, N. Laver and M. Diem, "Genitourinary & reproductive systems: spectral cytopathology of cervical samples: detecting cellular abnormalities in cytologically normal cells", *Lab. Invest.* **90(7)**, 1068 (2010). doi: [10.1038/labinvest.2010.72](https://doi.org/10.1038/labinvest.2010.72)
4. J. Nallala, O. Piot, M.D. Diebold, C. Gobinet, O. Bouché, M. Manfait and G.D. Sockalingum, "Infrared imaging as a cancer diagnostic tool: Introducing a new concept of spectral barcodes for identifying molecular changes in colon tumors", *Cytometry Part A*, **83(3)**, 294 (2013). doi: [10.1002/cyto.a.22249](https://doi.org/10.1002/cyto.a.22249)