Analyst

Spectral histopathology of colon cancer tissue sections by Raman imaging with 532 nm excitation provides label free annotation of lymphocytes, erythrocytes and proliferating nuclei of cancer cells

Laven Mavarani,[†][a] Dennis Petersen,[†][a] Samir F. El-Mashtoly,[‡][a] Axel Mosig,[a] Andrea Tannapfel,[b] Carsten Kötting[a] and Klaus Gerwert[*][a]

Spectral histopathology (SHP) is an emerging tool for label free annotation of tissue. While FTIR based SHP provides fast annotation of larger tissue sections, Raman based SHP is slower but achieves a 10 times higher spatial resolution as compared to FTIR. Usually NIR excitation is used for Raman measurements on biological samples. Here, for the first time 532 nm excitation is used to annotate colon tissue by Raman SHP. Excellent data quality is obtained, which resolves for example erythrocytes and lymphocytes. In addition to Raman scattering auto-fluorescence is observed. We found that this auto-fluorescence overlaps spatially with the fluorescence of antibodies against p53 used in routine immunohistochemistry in surgical pathology. This fluorescence indicates nuclei of cancer cells with mutated p53 and allows new label free assignment of cancer cells. These results open new avenues for optical diagnosis by Raman spectroscopy and autofluorescence.

Introduction

More than 1 million individuals worldwide develop colorectal cancer each year.[1] Most colon cancers start as small benign polyps which can be endoscopically resected. Thus simple and reliable detection methods for early detection are desirable. Usually cancer is identified by visual inspection during colonoscopy and by a pathologist performing microscopy of stained tissue slices from a biopsy. The pathologist has to use as marker hematoxylin–eosin (H & E) staining and in certain cases immunohistochemistry (IHC) to get the correct diagnosis. Many efforts were undertaken in the recent past to apply optical methods for label free and automatic detection of colon cancer and preneoplastic lesions (adenomas). These include auto-fluorescence, Fourier transform infrared (FTIR) and Raman spectroscopic methods.[2–4]

Tryptophan autofluorescence was successfully used to distinguish between isolated cells from human adenocarcinoma and normal mucosa due to the increased tryptophan fluorescence in the cancerous cells. However, autofluorescence

failed to distinguish surgical specimens of colonic neoplasms from normal mucosa due to optical absorption.[5] The ratio of autofluorescence upon 365 nm excitation and 405 nm excitation was used for autofluorescence imaging (AFI) of adenomas. A change in the NADH concentration due to the metabolic change in the tumour was suspected to influence the fluorescence ratio.[6] Near infrared excitation at 785 nm with auto-fluorescence detection between 810 nm and 1000 nm was used to distinguish between cancer, precancer and benign colonic tissue *in vivo* during endoscopy.[7] Besides the changes of light penetration due to morphological changes in the cancer and pre-neoplastic lesions, changes in concentrations of the fluorophores NADH, collagen, flavins and hematoporphyrin were discussed.

A vibrational spectrum contains many more spectral biomarkers than the fluorescence ratio. Thus an improved classification of the biochemical status of a tissue section can be obtained using spectral histopathology (SHP) by mapping the vibrational spectra of a specimen by FTIR or by Raman imaging. SHP has been applied successfully to many tissue types and several reviews are found in the recent literature.[2,4,8] This includes the application of Raman microspectroscopy to colon tissue by several groups.[9–12] In all cases normal and tumorous tissues were successfully distinguished. The excitation wavelength used in these studies was either 785 nm or 830 nm. Shorter wavelength Raman microspectroscopy is usually avoided for biological samples, because larger fluorescence background is expected.

[a] Department of Biophysics and Protein Research Unit Europe (PURE), Ruhr University Bochum, ND/04 Nord, 44780 Bochum, Germany. E-mail: gerwert@bph.ruhr-uni-bochum.de; Fax: +49 234 3214238; Tel: +49 234 3224461

[b] Institute of Pathology, Klinikum Bergmannsheil, Ruhr University Bochum, Germany

† These authors have equally contributed to this work.

‡ Current address: The Center for Aging and Associated Diseases, Helmy Institute of Biomedical Sciences, Zewail City of Science and Technology, 6th of October City, Egypt.

Here, we present a study using Raman imaging with 532 nm excitation. The obtained data quality is excellent. The shorter wavelength compared to the standard far red excitation lead to higher spatial resolution. Further, the signal/noise ratio is better due to better scattering and detector efficiency. We measured the same colon tissue samples and applied the same bioinformatics as before for FTIR based SHP.[13] In the training stage, hierarchical cluster analyses (HCA) of the Raman spectroscopic images were performed and index colour images were obtained. These were annotated by comparison with H & E staining and immunohistochemical staining. With these data a random forest (RF) classifier was trained. This classifier was used to identify colon tissue types. The higher spatial resolution compared to infrared allows for the detection of smaller structures like lymphocytes. Interestingly, we had no problems with large background fluorescence, but we observed strong auto-fluorescence from p53 expressing areas, which can be used for label free assignment of tumour and pre-tumour cells.

Experimental

Sample preparation

Collected datasets were gathered from formalin-fixed, paraffin-embedded tissue sections. They were obtained from the Institute of Pathology of Ruhr-University Bochum. The tissue sections were mounted on reflectively silver coated microscope slides (low-emissivity slides [Kevley Technologies, Chesterland, OH]) and de-paraffinized before measurements. For de-paraffinization a standard histological protocol was used.

Data acquisition

Raman hyperspectral datasets were acquired using a confocal Raman microscope (Alpha300AR, WITec Inc., Ulm, Germany) coupled to a frequency doubled solid state laser operating at 532 nm (WITec, Nd:YAG, max. 40 mW). A 25 μm diameter single-mode optical fiber was used to couple the laser radiation into the microscope. The incident laser beam was collimated *via* an achromatic lens and passed through a holographic band-pass filter before being focused into the sample through a 100×/0.9 NA objective (Olympus, Japan). The Raman scattered light is collected with the same objective and passed through a holographic edge filter onto a multi-mode optical fiber (50 μm diameter) to a spectrometer equipped with a back-illuminated deep-depletion charge coupled device (CCD) camera (1600 × 200 px.). Raman datasets were obtained with a pixel size of 0.9–1.0 μm for regions between 80 and 150 μm × 80–150 μm and 7 s exposure time per pixel was used for all measurements. The laser intensity was fixed to 1.5 mW.

Data analysis

The raw data were processed in Matlab Version 7.14 along with the Image Processing and Statistics toolboxes (The Mathworks, Inc., Mass., USA) and algorithms developed in-house. Cosmic spikes were removed using an impulse noise filter[14] and the spectra were interpolated to a reference wavenumber scale. HCA[15] was performed on the normalized data in the region between 700 cm^{-1} and 1800 cm^{-1} and between 2600 cm^{-1} and 3100 cm^{-1}. Pseudo-colour images generated from the clustering of the spectra were compared to the annotation of a pathologist and immunohistochemical staining. Spectra for training of a supervised learning algorithm (random forest[13,16]) were extracted from these datasets. Spectra with high autofluorescence were separated beforehand by setting a threshold on the signal intensity of the raw data. For further classification with a RF classifier a fifth order polynomial was fitted to each spectrum to remove the residual spectral baseline. Supporting points were selected by applying a sweep algorithm on the wavelet-denoised spectrum (Daubechies wavelet D4).[17]

The RF was trained for the following classes/tissue components (number of spectra): carcinoma, including pre-carcinoma (1007), connective tissue close to carcinoma (620), connective tissue (510), muscle (1272), inflammation (315), erythrocytes (47), crypts (goblet cells) (428), crypts (lumen) (85), lymphocytes (593) and background (1516). The total number of 6393 training spectra has to be expanded for additional colon tissue sections and for the enhancement of the stability of the classification. Therefore the results of the RF classification are preliminary and a full account will be published elsewhere.

H & E staining

After data acquisition and immunohistochemical staining the tissue slides were stained with hematoxylin–eosin (H & E). Staining of the cell nucleus and endoplasmatic reticulum was achieved by incubation of the tissue with hematoxylin for 15 minutes. After washing and stopping the hematoxylin reaction with H_2O the cytosol was stained with eosin for 3 minutes. The tissue slides were washed with H_2O and dehydrated in an EtOH gradient.

The H & E stained tissue slides were evaluated by a pathologist and compared to the HCA results of selected Raman data in order to get spectra for the training of the RF classifier.

Immunohistochemical staining

Fluorophores and antibodies were coupled according to the application details of the Fluorescent Dye Labelling Kit (antibodies-online GmbH). The immunohistochemical staining was performed after collecting the Raman datasets on the same slides. Therefore the slides were digested with boiling citrate buffer (10 mM, pH 6) for 25 minutes. After washing with PBS buffer (pH 7.4) the slides were blocked with 10% BSA for 20 minutes at room temperature. The slides were washed again with PBS and incubated at room temperature with a fluorophore-labelled monoclonal antibody Anti-phospho-p53 (EP155Y, Millipore, Germany). A fluorescein (FITC) tag was used as the fluorophore. In order to remove excessive antibodies the slides were washed thoroughly with PBS. Fluorescence images were obtained by using an Olympus microscope.

Results and discussion

In an earlier study, we reported on the infrared SHP of a large (16.5 mm × 15.1 mm) colon tissue section.[13] Assignment was

done by H & E staining and by immunohistochemistry using anti-desmin for muscle, anti-Col III for connective tissue and anti-mucin for submucosa, anti-p53 and anti-Ki-67 for cancerous regions. Here, an adjacent slice of the same biopsy was investigated by Raman microspectroscopy with an excitation wavelength of 532 nm. Thus the extensive information from the previous work can be used for the evaluation of the Raman data. While the sampling time for Raman microspectroscopy is substantially increased, the spatial resolution is improved. The measurement time is 7 s for each pixel which sums up to about 19 h for a typical image with about 10 000 pixels. The spatial resolution in FTIR is diffraction limited to about 10 μm, while we obtain about 1 μm in our Raman setup. Thus in a 100 μm^2 tissue section we have 10 000 independent pixels in the Raman while FTIR only resolves about 100 pixels for the same region.

In Fig. 1(A) a region of interest, crypts in the *Lamina propria mucosae*, is shown. This region responds to the anti-p53 and the anti-Ki-67 markers and was also picked up as cancerous by the FTIR SHP.[13] p53 and Ki 67 are markers for mutated and also proliferating cells, which are hallmarks for cancer cells. A section of this region was measured by Raman microspectroscopy and evaluated by the RF classifier. Spectra of pixels with high fluorescence background (>3000 counts) were sorted out and allocated to a separate class. A very detailed annotation emerges (Fig. 1(D)). The section is dominated by carcinoma or pre-carcinoma (red), and small areas characterized by the strong fluorescence (brown) are seen. In Fig. 2, the comparison of the Raman pseudo-colour image with the anti-p53 IHC image is shown. There is an excellent overlap of the auto-fluorescence-cluster obtained from the unstained tissue upon 532 nm excitation with the fluorescence from the anti-p53 stained sample. This suggests that 532 nm excitation can be used for a fast and label free annotation of p53 active tissue.

As a negative control we measured crypts from healthy tissue areas (Fig. 3). In this case, no fluorescence above the threshold is observed in the Raman measurement.

The higher spatial resolution of Raman compared to FTIR imaging allows resolving lymphocytes and erythrocytes. In Fig. 4 it is shown how the presence of lymphocytes changes,
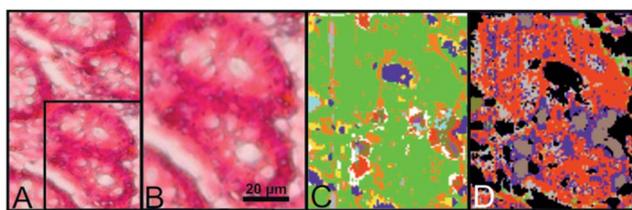

Fig. 2 (A) Raman pseudo-colour image (90 × 110 px.) of the crypts shown in (1D), constructed from the trained RF classifier. Red: pre-carcinoma, brown: p53 active areas, blue: adjacent tissue, green: co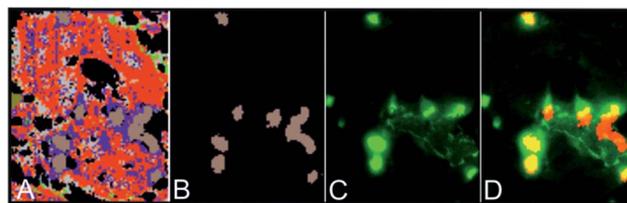nnective tissue, and olive: erythrocytes. (B) Raman-based p53 active areas from Raman imaging datasets. (C) Anti-p53 fluorescence staining, showing the p53 active areas. (D) Overlay of (B) in red and (C) in green. Overlapping regions appear in yellow.
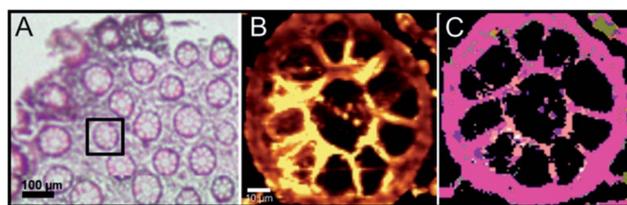

Fig. 3 (A) H & E stained colon tissue section. Crypts in the *Lamina propria mucosae* are shown in a healthy tissue area. For the region within the rectangle Raman imaging was performed. (B) Raman image (90 × 90 px.) of a crypt, showing the integrated Raman intensities of the CH-stretching mode from 2850 to 3050 cm^{-1}. (C) Raman pseudo-colour image of the crypt shown in (B), constructed from the trained RF classifier. Pink: goblet cells, purple: lymphocytes, rose: lumen of the crypts, and olive: erythrocytes.

depending on the grade of infiltration by tumour in the respective tissue area. In the healthy crypt shown in Fig. 4(A) and (D) the purple cluster of the RF classifier representing the lymphocytes is displayed occasionally. However, in Fig. 4(B) and (E), where a crypt that lies in an adjacent area to a tumour region is shown, the RF classifier recognizes a more frequent appearance of lymphocytes. The presence of a small p53 active area in this crypt shows an infiltration from adjacent cancer. A further increase of p53 active area, corresponding to increasing abnormal p53 expression and cancer activity, reduces the appearance of lymphocytes in the affected area as shown in the measured crypt in Fig. 2. In tissue regions within the carcinoma the presence of lymphocytes or active p53 areas is hardly visible due to mostly apoptotic cells. An exemplary tumour area measured by Raman spectroscopy is shown in Fig. 4(C) and (F).

Fig. 5 shows the mean spectra of selected classes of the RF classifier and the area of strong fluorescence: lymphocytes, erythrocytes, carcinoma, connective tissue and p53 active areas. The maximum of the fluorescence in the p53 active area (black in Fig. 5) is around 605 nm. Most likely porphyrin derivates are the source. Protoporphyrin IX was found in colorectal cancer.[18] Further, hematoporphyrin, the hydrolysis product of haemoglobin, is known to accumulate in cancer.[7] This accumulation of porphyrins is used in photodynamic therapy.[19] Interestingly protoporphyrin IX can directly bind to p53.[20]

Further structures, which were not resolved by FTIR, are lymphocytes. The spectrum of lymphocytes shows the

Fig. 1 (A) H & E stained colon tissue section. The same tissue was investigated by infrared imaging (see Fig. 5 in the literature[13]). Crypts in the *Lamina propria mucosae* are shown. For the region within the rectangle Raman imaging was performed. (B) Detail of the crypts marked by rectangle in (A). (C) Raman pseudo-colour image (90 × 110 px.) of the crypts shown in (B), obtained by unsupervised classification by HCA (D): Raman pseudo-colour image (90 × 110 px.) of the crypts shown in (B), constructed from the trained RF classifier. Red: pre-carcinoma, brown: p53 active areas, blue: adjacent tissue, green: connective tissue, and olive: erythrocytes.
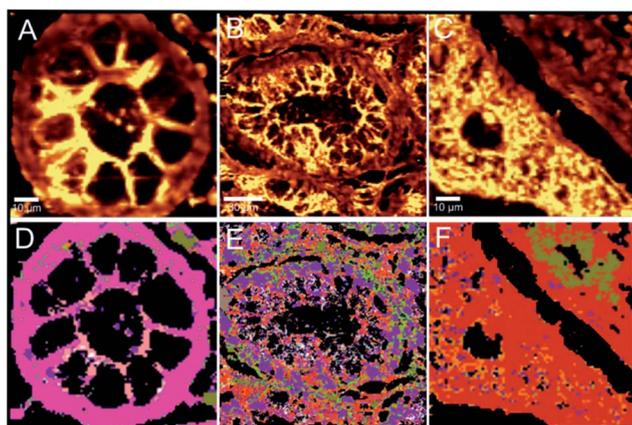
Fig. 4 (A) Raman image (90 × 90 px.) of a crypt in a healthy tissue area, showing the integrated Raman intensities of the CH-stretching mode from 2850 to 3050 cm^{-1}. (B) Raman image (180 × 180 px.) of a crypt in the adjacent area of a tumour region, showing the integrated Raman intensities of the CH-stretching mode from 2850 to 3050 cm^{-1}. (C) Raman image (90 × 90 px.) of a tumour region, showing the integrated Raman intensities of the CH-stretching mode from 2850 to 3050 cm^{-1}. (D)–(F) Raman pseudo-colour image of the tissue shown in (A), (B) and (C), respectively, constructed from the trained RF classifier. Pink: goblet cells, purple: lymphocytes, rose: lumen of the crypts, olive: erythrocytes, green: connective tissue, orange: inflammation, and red: carcinoma. The p53 active areas were added to the pseudo-colour image in brown.
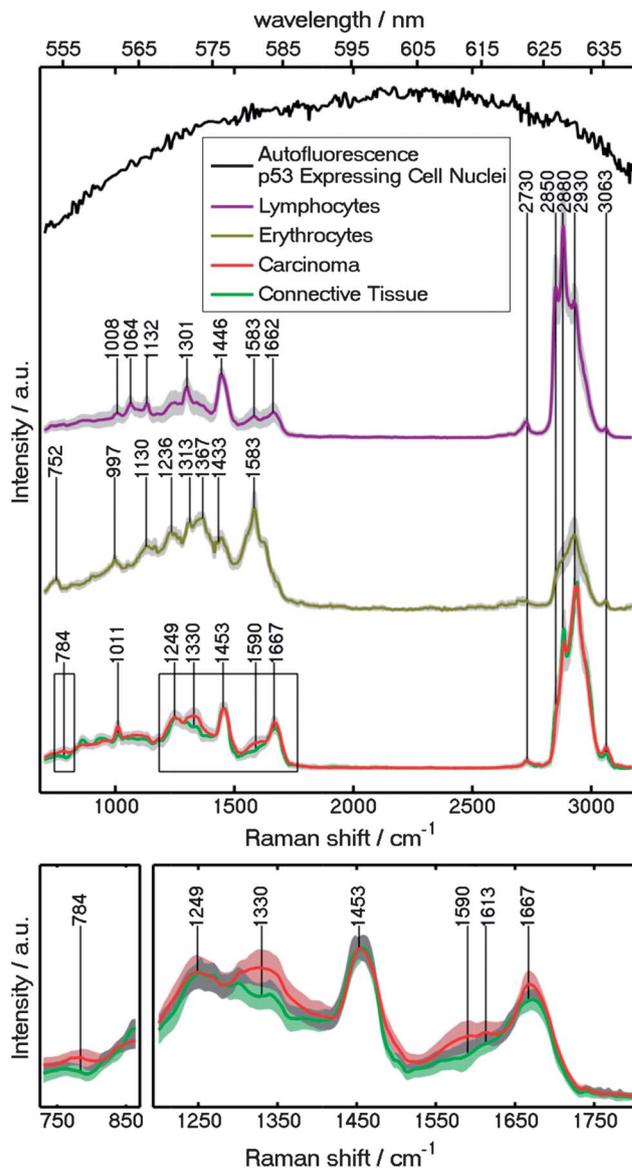

Fig. 5 Mean spectra of tissue components and an autofluorescence spectrum of the p53 active area, wavelet-denoised. The standard deviation is marked in grey. In the enlarged window the standard deviation is adjusted to the line colour. The overlapping standard deviation is marked in grey.

characteristic pattern of the large lipid content, with a strong band around 1450 cm^{-1} assigned to the CH$_2$ bending but only a small amide I band.[21] Further studies could try to differentiate lymphocytes in different stages. Small but significant differences in the Raman spectrum of activated and non-activated T-lymphocytes are reported.[22]

The spectrum of the erythrocytes is very different from the other components. It is neither dominated by the typical lipid nor the typical protein signals. Instead it is dominated by heme bands. These bands are significantly enhanced due to the resonance Raman excitation into the Q-band of the heme,[23,24] e.g. the marker band at 752 cm^{-1} can be assigned to the pyr-breathing mode. Note that the enhancement is not seen in Fig. 5 due to the normalization. Our spectrum resembles nicely most features of a spectrum assigned to hemoglobins obtained from a vertex component analysis in colon tissue.[12]

The spectra of connective tissue and carcinoma are similar, but they are separated by the classifier. At some band positions, the spectra of carcinoma are outside of the standard deviation of the connective tissue. Larger signals for the carcinoma are found at 1330 cm^{-1} and 1667 cm^{-1} indicating higher protein content. Further, a slightly higher content in DNA is indicated by a small band at 784 cm^{-1}. The larger content of protein and DNA was also found by a Raman imaging study on gastric cancer[25] and in a fiberoptic approach for colon cancer.[26]

The disadvantage of Raman imaging is the weak signal which leads to very long measurement times. There are several ways to reduce the measurement time. One is undersampling, e.g. by measuring with a step size of 10 or 100 μm.[9] However, in that case FTIR imaging might be better. Another possibility is the use of non-linear techniques like Coherent Antistokes

Raman Scattering (CARS).[27] With the results presented here another way might open up: cancerous regions can be searched for by auto fluorescence measurements and those regions can be investigated in detail with the same instrument by Raman. This approach is also feasible for fiberoptic setups.[28]

Conclusions

Due to the higher spatial resolution, Raman SHP can resolve smaller structures like erythrocytes and lymphocytes. Especially the information on the lymphocytes can be used for diagnosis. Surprisingly, 532 nm excitation yields good quality Raman spectra with autofluorescence specific to p53 active areas. This

finding might explain why simple autofluorescence imaging is capable of tumour recognition.

Acknowledgements

Notes and references

1 D. Cunningham, W. Atkin, H.-J. Lenz, H. T. Lynch, B. Minsky, B. Nordlinger and N. Starling, *J.-Lancet*, 2010, **375**, 1030–1047.

2 M. Diem, P. R. Griffiths and J. M. Chalmers, *Vibrational Spectroscopy for Medical Diagnosis*, Wiley-VCH, Weinheim, 2008.

3 J. Popp, *Handbook of biophotonics*, Wiley-VCH, Weinheim, 2011.

4 R. Salzer and H. W. Siesler, *Infrared and Raman spectroscopic imaging*, Wiley-VCH, Weinheim, 2009.

5 B. Banerjee, T. Renkoski, L. R. Graves, N. S. Rial, V. L. Tsikitis, V. Nfonsom, J. Pugh, P. Tiwari, H. Gavini and U. Utzinger, *J. Biomed. Opt.*, 2012, **17**, 016003.

6 K. Imaizumi, Y. Harada, N. Wakabayashi, Y. Yamaoka, P. Dai, H. Tanaka and T. Takamatsu, ed. D. L. Farkas, D. V. Nicolau and R. C. Leif, *Proc. SPIE 7902, Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues IX,* 2011, pp. 79020E.

7 X. Shao, W. Zheng and Z. Huang, *J. Biomed. Opt.*, 2011, **16**, 067005.

8 M. Diem, M. Miljković, B. Bird, T. Chernenko, J. Schubert, E. Marcsisin, A. Mazur, E. Kingston, E. Zuser, K. Papamarkakis and N. Laver, *Spectrosc. Int. J.*, 2012, **27**, 463–496.

9 A. Beljebbar, O. Bouché, M. D. Diébold, P. J. Guillou, J. P. Palot, D. Eudes and M. Manfait, *Critical Reviews in Oncology/Hematology*, 2009, **72**, 255–264.

10 C. Krafft, D. Codrich, G. Pelizzo and V. Sergo, *J. Biophotonics*, 2008, **1**, 154–169.

11 G. R. Lloyd, J. Wood, C. Kendall, T. Cook, N. Shepherd and N. Stone, *Vib. Spectrosc.*, 2012, **60**, 43–49.

12 C. Krafft, B. Dietzek, M. Schmitt and J. Popp, *J. Biomed. Opt.*, 2012, **17**, 040801.

13 A. Kallenbach-Thieltges, F. Großerüschkamp, A. Mosig, M. Diem, A. Tannapfel and K. Gerwert, *J. Biophotonics*, 2013, **6**, 88–100.

14 G. Judith.M.C and N. Kumarasabapathy, *Signal & Image Processing: An International Journal*, 2011, **2**, 82–92.

15 M. Miljković, T. Chernenko, M. J. Romeo, B. Bird, C. Matthäus and M. Diem, *Analyst*, 2010, **135**, 2002.

16 L. Breiman, *Int. J. Mach. Learn. Cybern.*, 2001, **45**, 5–32.

17 D. L. Donoho, *IEEE Trans. Inf. Theory*, 1995, **41**, 613–627.

18 K. T. Moesta, B. Ebert, T. Handke, D. Nolte, C. Nowak, W. E. Haensch, R. K. Pandey, T. J. Dougherty, H. Rinneberg and P. M. Schlag, *Cancer Res.*, 2001, **61**, 991–999.

19 D. E. J. G. J. Dolmans, D. Fukumura and R. K. Jain, *Nat. Rev. Cancer*, 2003, **3**, 380–387.

20 A. Sznarkowska, K. Maleńczyk, L. Kadziński, K. P. Bielawski, B. Banecki and J. Zawacka-Pankau, *FEBS Lett.*, 2011, **585**, 255–260.

21 A. I. Mazur, J. L. Monahan, M. Miljković, N. Laver, M. Diem and B. Bird, *J. Biophotonics*, 2013, **6**, 101–109.

22 K. L. Brown, O. Y. Palyvoda, J. S. Thakur, S. L. Nehlsen-Cannarella, O. R. Fagoaga, S. A. Gruber and G. W. Auner, *J. Immunol. Methods*, 2009, **340**, 48–54.

23 G. Rusciano, *Physica Medica*, 2010, **26**, 233–239.

24 M. Asghari-Khiavi, A. Mechler, K. R. Bambery, D. McNaughton and B. R. Wood, *J. Raman Spectrosc.*, 2009, **40**, 1668–1674.

25 M. S. Bergholt, W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh, J. B. Y. So, A. Shabbir and Z. Huang, *J. Biophotonics*, 2013, **6**, 49–59.

26 M. V. P. Chowdary, K. K. Kumar, K. Thakur, A. Anand, J. Kurien, C. M. Krishna and S. Mathew, *Photomed. Laser Surg.*, 2007, **25**, 269–274.

27 C. Krafft, A. A. Ramoji, C. Bielecki, N. Vogler, T. Meyer, D. Akimov, P. Rösch, M. Schmitt, B. Dietzek, I. Petersen, A. Stallmach and J. Popp, *J. Biophotonics*, 2009, **2**, 303–312.

28 M. S. Bergholt, W. Zheng, K. Lin, K. Y. Ho, M. Teh, K. G. Yeoh, J. B. Y. So and Z. Huang, *Biosens. Bioelectron.*, 2011, **26**, 4104–4110.