

FULL ARTICLE

Immunohistochemistry, histopathology and infrared spectral histopathology of colon cancer tissue sections

Angela Kallenbach-Thieltges^{#,1}, Frederik Großerüschkamp^{#,1}, Axel Mosig¹, Max Diem^{1,2}, Andrea Tannapfel³, and Klaus Gerwert^{*,1}

¹ Department of Biophysics and Protein Research Unit Europe (PURE), Ruhr University Bochum, Gebäude ND/04 Nord, 44780 Bochum, Germany

² Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA, USA

³ Department of Pathology of the Bergmann's Heil Hospital, Ruhr University Bochum, Germany

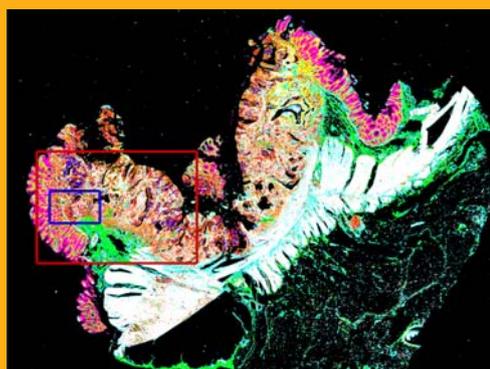
Received 17 July 2012, revised 26 October 2012, accepted 12 November 2012

Published online 10 December 2012

Key words: infrared spectral imaging, spectral diagnosis, hierarchical cluster analysis, random forest, colon cancer, fluorescence, immunohistochemistry, classification

➔ **Supporting information** for this article is available free of charge under <http://dx.doi.org/10.1002/jbio.201200132>

During the past years, many studies have shown that infrared spectral histopathology (SHP) can distinguish different tissue types and disease types independently of morphological criteria. In this manuscript, we report a comparison of immunohistochemical (IHC), histopathological and spectral histopathological results for colon cancer tissue sections. A supervised algorithm, based on the “random forest” methodology, was trained using classical histopathology, and used to automatically identify colon tissue types, and areas of colon adenocarcinoma. The SHP images subsequently were compared to IHC-based images. This comparison revealed excellent agreement between the methods, and demonstrated that label-free SHP detects compositional changes in tissue that are the basis of the sensitivity of IHC.



Infrared pseudo-color image of colon tissue section, constructed from the trained RF classifier.

1. Introduction

Colon cancer, with an incidence of over 140,000 cases and a mortality of over 50,000 in the US alone (2009) [1] is – after lung cancer – the second leading cause of death in cancers that affect both men and women. The worldwide incidence is estimated to be at least 10 fold higher than the US numbers.

Although regular colon cancer screening *via* colonoscopy can significantly reduce the mortality from colorectal cancer, its high incidence and the high cost of colonoscopy (over \$ 1,000 in the US) present an enormous burden to the health care system.

Most colon cancers start as small benign adenomas (polyps) which may turn malignant with time. Risk factors for developing colon cancer include the

* Corresponding author: e-mail: gerwert@bph.rub.de; Tel.: +49-(0)234-3224461, Fax: +49-(0)234-3214238

Both authors contributed equally to this work.

hereditary condition referred to as “familial adenomatous polyposis” (FAP), [2, 3] a condition in which the colonic mucosa develops hundreds to thousands of polyps, and which carries a near 100% risk of turning malignant. Another hereditary risk factor leads to a condition known as hereditary non-polyposis colon cancer (HNPCC) [2]. Non-hereditary risk factors include smoking, obesity, and lack of physical exercise. In most colon cancers, several genes are mutated, among them the APC gene, which leads to increased levels of β -catenin, which can activate stem-cell like traits in cells and lead to cancer [4]. Furthermore, p53 and/or other apoptotic proteins often are found to have acquired mutations in colon adenocarcinomas, and certain oncogenes, such as KRAS, are overexpressed in them. The presence or absence of some of these cancer markers form the basis of presently used immunohistochemical methods to diagnose and classify colon cancers. Thus, there is ample evidence that suggests that several mutations must occur for colon cancer to develop, and that the presence of a number of these markers may be useful for typing and classifying colon cancers.

The detection and classification of colon cancers is presently carried out, at the most basic level, by visual inspection of the suspicious lesion or a polyp during colonoscopy. Indeed, experienced physicians can predict whether or not a polyp is benign or malignant with reasonably high accuracy. The next level of diagnostic and classification decisions lie with the pathologist who inspects a biopsied polyp or lesion, after staining, by visual microscopy. Although this step is reasonably accurate in detecting and diagnosing cancer and cancer stage, it cannot reveal information on the origin of the disease in terms of mutations and cancer markers. The ultimate test for the detection of mutations is next generation gene sequencing [5], which directly can detect mutations in excised tumor tissue. However, even with the enormous increase in speed and reduced cost of this methodology, it is presently not a viable option for routine diagnostic work. Information on the presence of certain cancer markers is available only from multi-panel immunohistochemical (IHC) analysis in which tissue sections are incubated with specific antibodies, linked to labels which can be detected by their UV/visible absorption or emission spectra. However, IHC has its own set of experimental uncertainties and difficulties. First, cancers with the same pathological diagnosis may present different response to the panel of IHC-based markers, since the different highly coupled signalling pathways give different individual responses in each patient. In addition, not all markers may be detectable in a given tissue section [6, 7]. Thus, the results from a multi-panel IHC analysis may need multivariate methods to link specific marker response to a bio-molecular

description of the cancer. Furthermore, in IHC one needs to preselect markers one expects to find – it is not an unsupervised method that detects abnormalities in general. Finally, IHC is a method that requires several adjacent tissue sections, each of which must be incubated with its specific antibody or marker, linked to a dye or reporter moiety. A multi-panel IHC analysis may take several days of work, and run in the thousands of dollars.

The IHC methodology can be complemented with recently developed spectroscopic and spectrometric methods. Among the latter, proteomic approaches, based on mass spectrometry (MS) or MS-based imaging methods, such as MALDI, ToF-SIMS or DESI [41, 42] present the researcher with a list of all proteins that are up-regulated, down-regulated or at normal levels. This presents the advantage that prior knowledge of expected changes is not necessary. However, MS-based methods are quite work-intensive as well, require a substantial amount of material, and analysis of the entire proteome of a cancer biopsy is too complicated for routine diagnostics. Furthermore, MS-based methods (with the exception of MALDI-imaging) do not present spatial information, which is needed in medical diagnostics. Among the spectroscopic methods to analyse tissue biopsies or individual cells, those based on vibrational (infrared and Raman) spectroscopy have proven particularly useful. These techniques possess high sensitivity, but the signatures of all bio-molecular components of a tissue sample are superimposed. This offers advantages as well as disadvantages. Since all signatures of biomolecules fall into a reasonable narrow spectral range, they all can be detected simultaneously in a routine experiment. In this sense, vibrational microspectroscopy is a “top-down” technique that simultaneously, and spatially resolved, integrates the entire proteome, genome and metabolome of a cell or tissue. Since all these signals are superimposed, they can be detected rapidly and with a high signal-to-noise (S/N) ratio; however, pattern analysis of the spectral data needs to be employed to detect small spectral changes between healthy and diseased tissue. Subsequently, these differences have to be related to specific diseases with high sensitivity and specificity. In this context, the emphasis is that only the spectral differences between disease states need to be determined and interpreted. This technique, commonly referred to as spectral histopathology (SHP), is a combination of microscopic infrared-spectral data acquisition from pixel elements about 10 μm on edge (determined by the diffraction limit of infrared radiation) and multivariate analysis to detect spectral changes from different tissue areas, and relate them to the morphology of the tissue and specific disease states and stages. Over the past decade, SHP has been applied by a number of research groups worldwide to

first identify spectral differences between different (normal) tissue types (for a recent review, see Refs. [8, 9]), between normal and diseased tissues [10–15], between tissues with different disease types [16], and most recently, as a method to elucidate prognostic information [17, 18]. In addition, it can be seen as a major advantage of infrared microspectroscopy as compared to MS based methods that it works on formalin-fixed paraffinized samples, since such samples are available in large numbers and keep stable even over long periods of time. In agreement with the early work by Lasch [19, 22] we observe, that patient-to-patient variations in spectral patterns are smaller than those caused by disease, or even due to different tissue types. Thus, these pioneering studies realized and implemented the possibility of training supervised machine-learning algorithms for the diagnosis of colon cancer [23]. This latter study also indicated that the changes in spectral patterns caused by disease are reproducible between different methods of selecting training spectra. Modern infrared imaging micro-spectrometers are now commercially available which allow the acquisition of spectral images many mm^2 in size in short acquisition times, as are computational facilities that permit multivariate analysis to be carried out for such data sets. Thus, it is timely and highly important to investigate the limits of sensitivity and specificity of SHP in terms of its ability for classification of cancerous tissue, and establish its prognostic values. Several researchers have now collected data sets with 70 to 300 patient databases [16, 24, 25], and have demonstrated that the distinction of cancer types is possible with high specificity and sensitivity [16]. This is the first step in establishing SHP as a therapeutic tool, since proper diagnosis, in certain cancer, determines the next treatment steps. This approach might open avenues into personalized medicine by more precise medical diagnostics.

The (infrared) spectral signatures of disease, i.e., the spectral differences between normal and diseased tissues which are used by the diagnostic algorithms for detection and classification of disease have largely been ignored, although some attempts were made to interpret the changes observed between normal and diseased spectral classes [26]. Recently, however, some literature references indicate that methods of vibrational spectroscopy can detect markers of abnormality [27, 28]. To further verify such claims, and to gain a further understanding of the features that are used as classifiers by the diagnostic algorithms, datasets are needed that contain, aside from the infrared spectra, additional information from immunohistochemistry and/or fluorescence staining. In this paper, a first correlation of these additional cancer biomarkers with infrared spectral features will be presented. Thus, the goal of this study is to pave the way of correlating specific mar-

kers, such as p53 or EGFR, to infrared spectral biomarkers. The patient number in this study is too small to arrive at any generalization toward a larger pool of samples or at a generally applicable diagnostic algorithm. Rather, the aim of this study is to demonstrate the superiority of SHP in detecting all cellular and tissue components simultaneously, therefore it is a more universal tool than other spectral diagnostic methods, and to demonstrate the advantages of this label-free and spatially resolved method.

To this end, this study presents results comparing hematoxylin/eosin (H&E) stained images, infrared spectral images, and images based on fluorophore-labelled antibodies to mucin, collagen, muscle tissue, p53 and Ki-67 in order to establish whether or not SHP can use spectral biomarkers to characterise diseased colon tissue with high specificity and sensitivity. The results presented here demonstrate the colocalization of these markers with specific spectral signatures, and suggest that SHP – in one single measurement – presents multiplexed data presently available only from several separate staining experiments. This aspect is particularly important since SHP may be able to provide, in addition, a therapeutic pointer: if, for example, alterations induced by overexpressed mutated KRAS can be detected in a cancer biopsy by SHP, it may save the patient from being treated with drugs that inhibit EGFR, such as cetuximab or panitumumab, which are nearly ineffective in tumors that over-express mutated KRAS. Furthermore, this study sheds light on the sensitivity of SHP, and helps establish the possibility of data mining of infrared hyperspectral data sets. Our study utilizes two different FT-IR microscope platforms, one for obtaining training data and the other for validation on a large tissue sample, and thus validates the robustness of SHP against utilizing such different platforms.

Finally, this paper introduces two novel contributions on the computational analysis of FT-IR spectral images, one in the context of unsupervised segmentation during image annotation, and one regarding supervised classification of image spectra. Regarding unsupervised segmentation, we introduce a novel cross-validation scheme that allows to quantitatively assessing the accuracy of segmentation obtainable from hierarchical clustering, compared to qualitative studies that have been conducted previously. On supervised classification of image spectra, we assess the utilization of Random Forest (RF) classifiers. Introduced by Breiman in the early 2000s [37], Random Forests combine a number of decision tree classifiers into a so-called bagging classifier. While each decision tree by itself is relatively weak, the majority vote over the individual decision trees typically and in certain cases also provably [37] yields a highly accurate classification. A number of favourable proper-

ties have contributed to the popularity of using RFs rather than support vector machines (SVMs), discriminant analysis, or artificial neural networks (ANNs) in numerous applications [38, 40] and have propelled their rapid rise within essentially a decade [43]. Beside an accuracy generally comparable to other established approaches, RFs are robust against overfitting both in theory and practice [37] and computationally efficient both in training and validation. As an RF is constituted by nothing more than a collection of decision trees, RFs can be considered easy to use in practice, since they require neither choosing a network topology like ANNs nor a kernel function, as required by SVMs. In fact, the only parameters required to train an RF are the number of decision trees and the number of features drawn at random for each node of the decision tree. Rather than requiring a feature selection prior to training like ANNs or SVMs [23], RFs can be and have been utilized for the purpose of feature selection [38].

2. Materials and methods

The procedures underlying the research presented here are conceptualized in Figure 1, which details the steps in sample preparation, data acquisition, data segmentation and training of classifier. In the following section, these steps are discussed in more detail.

2.1 Sample preparation

Formalin-fixed, paraffin-embedded tissue sections were obtained from the Department of Pathology of the Bergmann's Heil Hospital in Bochum. Several

adjacent 5 μm thick tissue sections were cut from the tissue blocks via a microtome. The sample to be used for spectral analysis was mounted on a reflectively silver coated microscope slide (low-emissivity or "low-e" slide [Kevley Technologies, Chesterland, OH]). All sections were de-paraffinized using standard procedures [49] and left without a coverslip. Fluorescence and immunohistochemical staining was performed as described below. The size of the tissue section shown in Figure 2 was ca. 16.5 mm \times 15.1 mm.

2.2 Immunohistochemical staining

Primary antibodies were coupled with fluorophores of the Fluorescent Dye Labeling Kits (antibodies-online GmbH), as described in the application details [29]. The slide used previously for IR data acquisition was used for fluorescence studies. First, the tissue was digested with 0.1% pepsin (0.01 M HCl) for 30 minutes at room temperature. After being washed for three times with PBS (pH 7.4) it was blocked with 10% BSA for 20 min at room temperature. The tissue was again washed with PBS and incubated for 30 minutes with fluorophore-labelled antibodies at room temperature. Three antibodies against tissue-structures were applied in one step: monoclonal Mucin 2 (B306.1, Santa Cruz Biotechnology, USA) with fluorescein (FITC), monoclonal Anti-desmin (D33, antibodies-online GmbH) with rhodamine and monoclonal Anti-Collagen type III (FH-7A, Sigma-Aldrich, USA) with AMCA. Excess antibodies were removed by washing three times with PBS. Fluorescence images were recorded automatically with an Olympus microscope. Subsequently, the three antibodies were removed by soaking the tis-

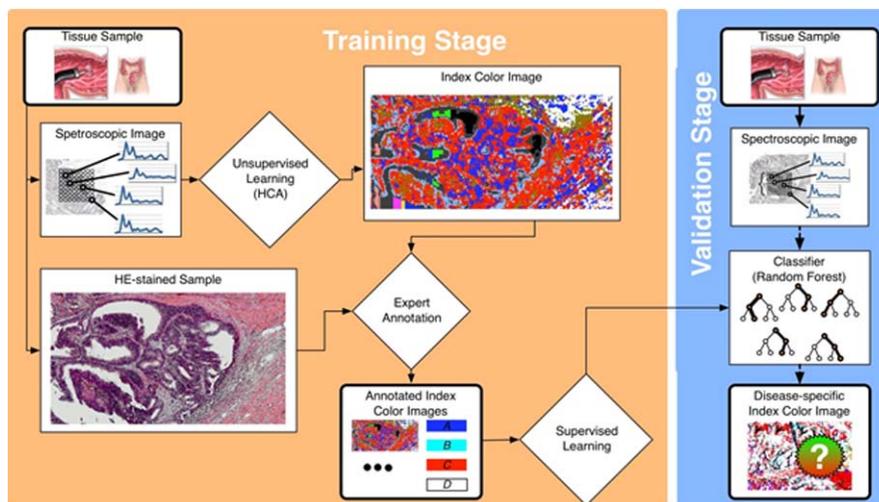


Figure 1 (online color at: www.biophotonics-journal.org) Procedure for obtaining spectral image classifiers. The supervised classifier for image spectra (right hand side) is trained on spectra obtained from spectral images annotated during the training stage (left hand side). Annotation is based on overlaying an H&E stained image with a pre-segmented spectral image of the same region.

sue in PBS (pH 11) overnight. Next, the tissue was incubated with antibodies for specific cancer markers. This procedure started again with a digestion step (0.1% pepsin in 0.01 M HCl) for 30 minutes at room temperature. The antibodies used were monoclonal Anti-phospho-p53 (EP155Y, Millipore, Germany) with a fluorescein (FITC) tag, and monoclonal Anti-Ki-67 (Ki-S5, Millipore, Germany) with AMCA.

2.3 Data Acquisition

Infrared hyperspectral datasets were collected in transfection (reflection-absorption) mode using a Bruker (Ettlingen, Germany) Model Hyperion 3000 infrared microscope coupled to a Vertex 70 FTIR bench for the training stage, and an Agilent (Santa Clara, California, USA) Model Cary 620 infrared microscope coupled to a Cary 670 bench for the validation stage. The Bruker system, henceforth referred to as the “Hyperion”, employs a liquid nitrogen-cooled 64×64 detector element focal plane array with rapid scan (rather than step-scan) readout. Each 64×64 element data acquisition provided a field of view (FOV) of approximately $172 \mu\text{m} \times 172 \mu\text{m}$. The instrument and the microscope chamber are continuously purged with dry air to reduce atmospheric water spectral contributions.

The Agilent system, henceforth referred to as the “Cary”, employs a liquid nitrogen-cooled 128×128 detector element focal plane array with rapid scan readout. Each 128×128 element data acquisition provided a field of view (FOV) of approximately $715 \mu\text{m} \times 715 \mu\text{m}$. The instrument and the microscope chamber of both are continuously purged with dry air to reduce atmospheric water spectral contributions. For the Cary instrument, a 24/7 liquid nitrogen cooling supply (Norhof, Maarssen, Netherlands) was added.

For each FOV of the Hyperion, a background acquisition of 32 co-added interferograms was collected, and Fourier transformed using power phase correction and Blackman-Harris 3-term apodization. 32 co-added sample interferograms were processed in the same way and ratioed against the background spectra. 16 individual tiles, each measuring 64×64 pixels were stitched automatically, before a new data collection area was selected. Thus, each stitched dataset covered an area of approximately $700 \mu\text{m} \times 700 \mu\text{m}$, and consisting of 65536 individual spectral vectors. Each raw spectral vector consisted of 1362 data points (resolution 4 cm^{-1} , zero-filling 4, upper folding limit 3949 cm^{-1} , acquisition range $2700\text{--}950 \text{ cm}^{-1}$), with a data point spacing of 1.286 cm^{-1} .

For each FOV of the Cary, a background acquisition of 64 co-added interferograms was collected,

and Fourier transformed using Mertz phase correction and Blackman-Harris 4-term apodization. 64 co-added sample interferograms were processed in the same way and ratioed against the background spectra. Individual mosaic tiles, each measuring 128×128 pixels, were stitched automatically. Each raw spectral vector consisted of 1428 data points (resolution 4 cm^{-1} , zero-filling 4, upper folding limit 5266 cm^{-1} , acquisition range $3700\text{--}950 \text{ cm}^{-1}$), with a data point spacing of 1.928 cm^{-1} .

2.4 Spectral pre-processing

Raw data sets were pre-processed as follows. In early attempts, 16 pixel spectra (collected at a pixel size of $2.7 \mu\text{m}$ on edge via the Hyperion) were co-added (binned) to produce one pixel spectrum with improved S/N ratio and a spatial resolution approximately equal to the diffraction limited focal spot, which is about $6 \mu\text{m}$ at 1600 cm^{-1} and $10 \mu\text{m}$ at 1000 cm^{-1} . This reduced the size of the dataset 16-fold. The use of this datasets in the training stage (see below) showed that it was possible to work with the oversampled dataset collected at $2.7 \mu\text{m}$ (Hyperion) or $5.5 \mu\text{m}$ (Agilent) spatial resolution. In fact, all images presented here are from the original dataset of un-binned data (RF based images with binned and un-binned data are shown in the Supplementary Figure 1).

Pixel spectra from the vicinity of voids or cracks in the tissue, or pixel spectra from highly spherical, small cells such as lymphocytes, may exhibit strong scattering artifacts that have been attributed to “resonance Mie” scattering [30, 39]. In order to eliminate such confounded spectra, a random forest classifier was trained on a set of spectra selected manually from representative confounded tissue regions to reject these spectra. This pre-screening was performed in the $1800\text{--}950 \text{ cm}^{-1}$ wavenumber range. Subsequently, all spectra were subjected to an EMSC-based Mie and resonance-Mie scattering correction [31], $2700\text{--}950 \text{ cm}^{-1}$, one iteration step. Finally, the spectral range outside the fingerprint ($1000\text{--}1800 \text{ cm}^{-1}$) region was rejected and spectra were smoothed using a 9 point Savitzky-Golay filter [32]. All spectral data and image processing was implemented using Matlab Version 12 along with the Image Processing and Statistics toolboxes (The Mathworks, Inc., Mass., USA).

2.5 HCA validation

For quantifying the reliability of HCA for spectral image annotation, we employed a novel cross-valida-

tion scheme for hierarchical clustering. Here, a training data set with k types of tissue components is represented by sets of spectra C^1, \dots, C^k , where C^i is the set of all training spectra annotated as class i . HCA is performed on the combined set of all spectra $C = C^1 \cup C^2 \cup \dots \cup C^k$ resulting in a dendrogram Z . Here, the symbol \cup denotes the combination (union) of all training data. "Cross validation" now can be interpreted as identifying the vertices v representing each class i , $i = 1, \dots, k$. Now, for each vertex v in Z we can determine the percentage of spectra belonging to S^i , formally denoted by $n^i(v)$. A segmentation into k classes formally is equivalent to selecting k vertices v^1, \dots, v^k in such a way that each spectrum $x \in C$ is represented by at most one vertex v^i . Thus, the best possible segmentation supported by Z consists of those vertices v^1, \dots, v^k that maximize $\sum n^i(v^i)$, while requiring each spectrum to be covered by precisely one of the v^i . This optimal set can be computed using an integer linear program [33], yielding the maximum number of correctly assigned spectra and thus the optimal accuracy of annotation that can be achieved by Z . The validation was implemented in Matlab using a Matlab-interface to *lpsolve* Version 5.5 to solve the integer linear programs.

Accuracy of HCA Classification. Using the novel cross-validation scheme for HCA analysis described above, ten-fold Monte Carlo cross validations of Ward's clustering (using 2/3 of the data as training data and the remaining 1/3 for validation in each of the ten runs) in combination with correlation distance yielded an accuracy of 68.03% ($\pm 0.013\%$) of HCA on our training data.

Accuracy of Random Forest Classification. Performing ten Monte Carlo cross validation using a random forest classifier without prior feature selection (using the *randomforest-matlab* implementation [48]) on the training data set yielded an accuracy of 95.6%, which can be achieved by neural networks or support vector machines only if an appropriate feature selection, network topology and kernel function, respectively, are chosen (Supplementary Figure 2). While it has been observed previously on a qualitative basis that unsupervised classification in the training stage is less accurate than the subsequent supervised classification, our approach allows putting this observation on a quantitative ground. The classification accuracy is further demonstrated by a validation on 46 Biomax sample spots randomly chosen from a set of 96 different patients. 36 of which are associated with samples diagnosed as cancerous. All ten spots associated with healthy non-cancerous samples contain less than 3.6% spectra classified as cancer (Supplementary Figure 8). The Biomax, Inc., provides tissue array samples that are standardized and annotated by two pathologists.

3. Results and discussion

Exemplary for this study we display a tissue section from a large resected colon section of a patient with an invasive colon adenocarcinoma in this paper. This tissue section, henceforth also referred to as the *validation section*, (see Figure 2), ca. 16.5 mm by 15.1 mm in size, was obtained from a paraffin-embedded tissue block at Bergmann's Heil Hospital in Bochum, and was provided by one of the co-authors (AT). An infrared spectral dataset, which shows a cross section of the colon, was collected from the

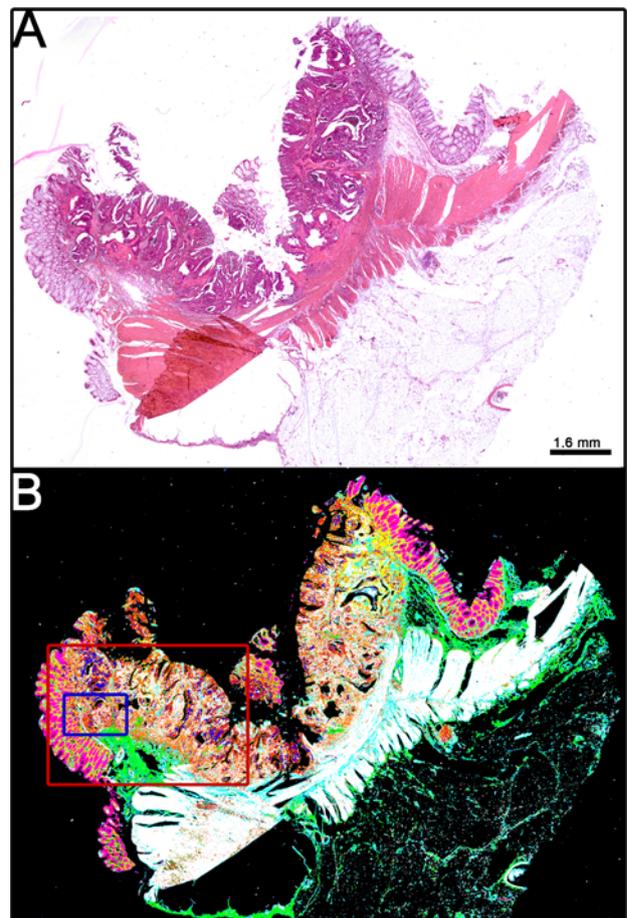


Figure 2 (online color at: www.biophotonics-journal.org) (A) Photomicrograph of an H & E stained colon tissue section, measuring ca. 16.5 mm \times 15.1 mm 2 . (B) Infrared pseudo-color image of the tissue section shown in A, constructed from the trained RF classifier. The area marked by the red rectangle is the region shown in Figure 3, and the one marked in blue in Figure 4. In Panel 2B (and all RF-based images in the next Figures), green and yellow hues denote connective tissue, white: musculature, cyan: connective tissue with supporting cells, pink: lumen of crypts, olive: blood, blue: pathological connective tissue, orange: inflammatory tissue, red: carcinoma. At this magnification, areas of cancer can barely be seen.

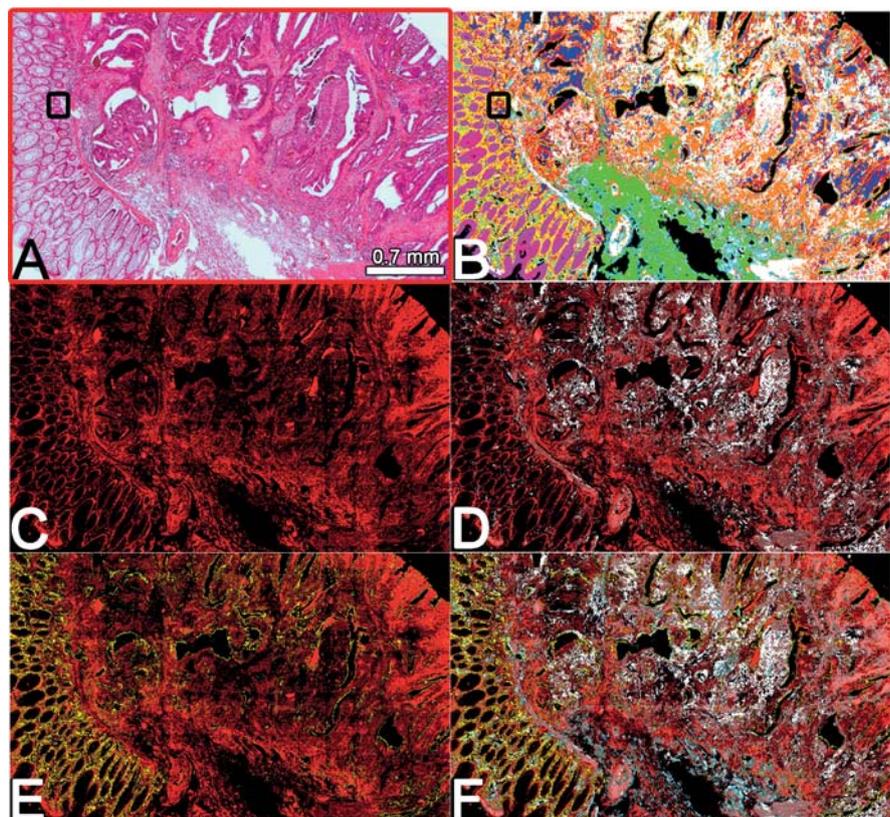
whole tissue section, consuming two days of measurement time on the Cary. Figures 3–5 depict regions of this tissue section for a more detailed comparison of SHP, H&E and IHC images. In order to train spectral classifiers, our study involved six further tissue samples from six different patients, henceforth referred to as training sections, along with three samples from Biomax tissue arrays. Note, that 96 different patients were measured and used for further validation. The training sections were measured using the Hyperion, while spectral images of the validation section were recorded on the Cary microscope. This allows a cross-platform validation of spectral histopathology, which is important for a broad application. Note that the difference in transfer functions between the different platforms is resolved through the resonance Mie correction and the averaging to 100 wavenumbers performed during spectral preprocessing. Sample spectra before and after resonance Mie correction are displayed in Supplementary Figure 5.

3.1 Work flow/Training stage

As shown in Figure 1, the data analysis essentially works in two phases, a training stage and a validation stage. The primary purpose of the training

phase is to obtain representative spectra for different tissue classes or disease states of different patients in order to train a supervised classifier. To this end, spectral images are segmented into differently coloured sub-regions using hierarchical cluster analysis (HCA). Following the procedures reported by Lasch (2004), we applied Ward's method in combination with correlation distance for unsupervised clustering (using 2nd derivative spectra obtained by applying a 9 point Savitzky-Golay algorithm) [32] on six tissue samples from six different patients and three Biomax tissue array samples also from different patients. Cutting the resulting dendrogram at a useful level yielded pseudo-color images that were overlaid with the H&E stained image of the same tissue region. The HCA index color images were annotated by an experienced pathologist (AT) through identifying a tissue component within each index colour. This yielded a set of training spectra for 14 classes of tissue components. Spectra were reduced to 100 data points between 1790–950 cm^{-1} by choosing 100 bands on an equidistant wavenumber scale; the absorption of each of these reduced bands is taken as the average of the (roughly four) bands covered on the original wavenumber scale. In each class, outlier spectra were eliminated by plotting all spectra belonging to one class and eliminating all those spectra which exhibit obvious deviations from the majority of the remaining spectra. This curation step yields our final

Figure 3 (online color at: www.biophotonics-journal.org) (A) H&E stained tissue section shown by red rectangle in Figure 2A. (B) RF-based pseudo-color image (green and yellow hues denote connective tissue, white: musculature, cyan: connective tissue with supporting cells, pink: lumen of crypts, olive: blood, blue: pathological connective tissue, orange: inflammatory tissue, red: carcinoma). (C) Anti-desmin fluorescence stained image for musculature (D) agreement plot of the Anti-desmin fluorescence image with the RF class for musculature (E) overlay of the Anti-desmin fluorescence staining with the RF class for mucosa (F) overlay of the Anti-desmin fluorescence staining with the RF class for musculature, mucosa and supporting cells of connective tissue.



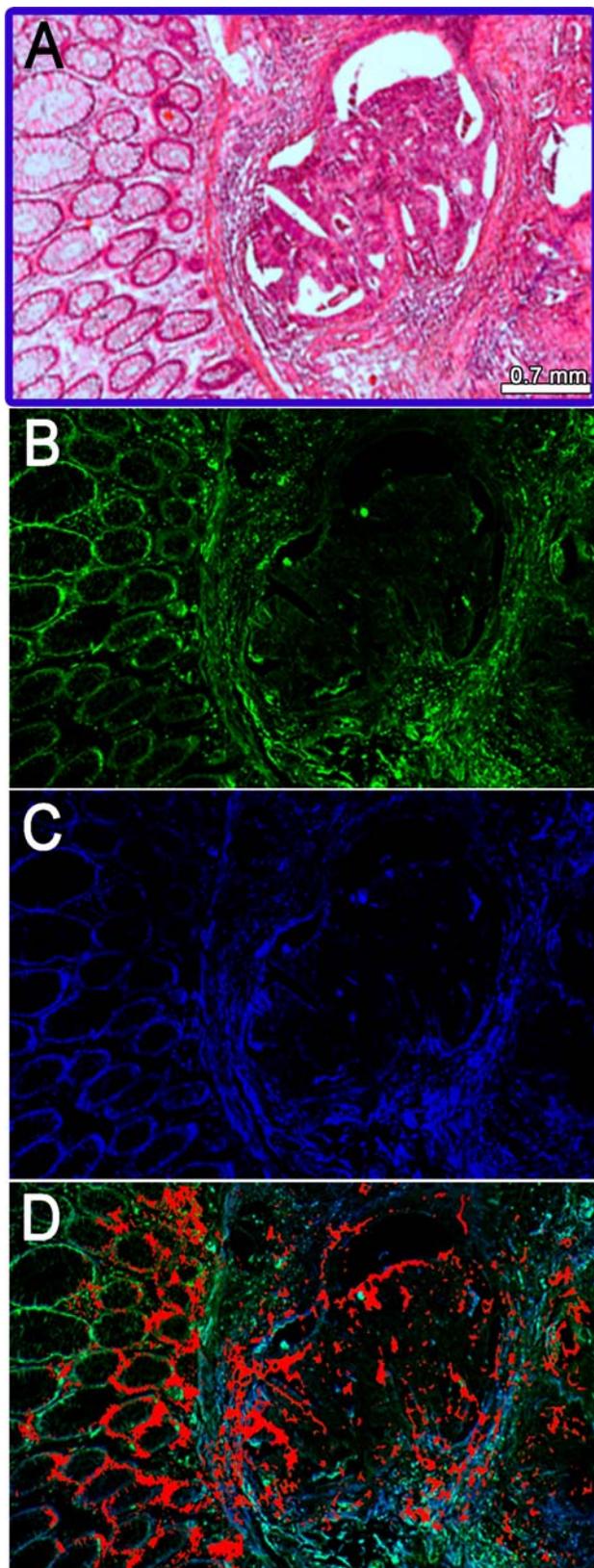


Figure 4 (online color at: www.biophotonics-journal.org) (A) H & E stained image of section shown by blue rectangle in Figure 2A. (B) Anti-p53 fluorescence stained image (C) anti-Ki-67 fluorescence stained image. (D) Overlay of anti-p53 (green), anti-Ki-67 (blue) fluorescence stains, and RF-based cancer class (red).

training data set which is characterized in detail in Supplementary Table 1, comprising 34.964 spectra classified into 14 classes in total.

3.2 Validation stage

Based on the annotations obtained in the training stage, we trained Random Forest classifiers to perform supervised classification. Using the training data set described above, a Random Forest classifier was trained on the pre-processed, but not derivatized absorption spectra using $m = 500$ trees and sampling \sqrt{L} among the $L = 100$ wavenumbers in the spectral range under consideration. The training dataset included only spectra from the Hyperion. Classification results of the RF under different numbers of decision trees as well as different constellations of ANN and SVM classifiers were assessed using ten-fold cross validation. As shown in Supplementary Figure 2, the accuracy of the RF is consistently around 95% accuracy independent of the number of trees. On the other hand, ANNs or SVNs in combination with different versions of the well-established *minimum redundancy maximum relevance* (mRMR) scheme for feature selection [44] achieve comparable accuracy only if an appropriate number of selected features, number of network layers, or kernel function, is chosen. While either the mRMR selected features or the Gini variable importances obtained from the RF (Supplementary Figure 3) can be used to identify relevant spectral bands, we avoid detailed interpretation due to potential band shifts resulting from the RMie correction.

Based on this RF classifier, we obtained the spectral image displayed in Figure 2A, which may be viewed as an automatic annotation performed by the random forest classifier. In order to demonstrate the SHP-based tissue diagnosis, a comparison to the subsequently H & E stained sample and the IR-based pseudo-color image is shown in Figure 2B. It is obvious that the gross tissue morphology is correctly described by the marker-free SHP approach. In the following we will compare in more detail the results of the SHP with the H & E stained and fluorescence labelled sample.

In order to compare in more detail the annotation by automated marker-free SHP with the gold standards of pathology, the marked area in red in

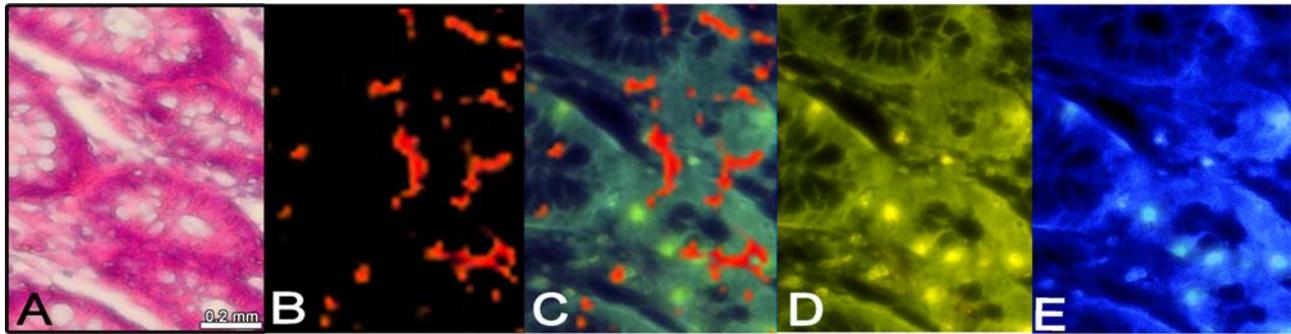


Figure 5 (online color at: www.biophotonics-journal.org) (A) Detail of three crypts marked by black rectangle in Figure 3A. (B) RF-based cancer regions (C) Overlay of the anti-p53 (green), the anti-Ki-67 (blue) fluorescence and the RF-based cancer regions (red). (D) anti-p53 fluorescence stained image (E) anti-Ki-67 fluorescence stained image.

Figure 2B (ca. 4.7 mm × 2.9 mm) is shown enlarged in Figure 3B. It is compared with the corresponding H&E stained section shown in Figure 3A. This region represents a disease-affected area, and contains areas of significant abnormality on the right side, as indicated by the darker-stained regions, as well as early stages of abnormal crypts on the left. The crypts in the left half of Figure 3A show thickening of the cell layer, and somewhat abnormal structure. The same features are identified by SHP in Figure 3B. It shows the corresponding infrared spectral image, constructed from a “random forest” (RF) diagnostic algorithm that was trained from tissue sections of several patients (including a different section of healthy tissue from the same patient). This panel accurately reproduces the structures that are apparent in the H&E-stained image: The crypts, embedded in *lamina propria mucosae* (shown in yellow) appear as glandular (goblet) cells shown in magenta and pink. Here, the magenta areas agree quite well with the layer of nuclei seen at the edges of the crypts, with the pink areas agreeing with the cells’ cytoplasm. The region of *lamina muscularis mucosae* is nicely seen on the left side (shown in white like the musculature); on the right side it is infiltrated by carcinoma. The carcinoma (red) is bounded by pathological connective (blue) and inflammatory tissue (orange). An overlay of the fluorescence-labelled anti-desmin (Figure 3C), (to be discussed in detail in the next section) with some classes of the RF spectral image is shown in Figures 3D–F. This demonstrates that SHP can identify, automatically and marker-free, the same features as seen in an H&E stained sample.

3.3 Comparison with fluorescence labelling

In the next step we compare in more detail the SHP image with the common used fluorescence label image used in IHC. The desmin-rhodamine label speci-

fic for muscle tissue (this was actually done before H&E staining) is shown in Figure 3C. The fluorescence labelled image in Figure 3C is overlaid with the RF class for musculature shown in white in Figure 3B. The results is shown in Figure 3D. In areas of agreement, the RF class fluorescent image is represented transparently, which leads to a pinkish color. The white color of the RF image is maintained, when the fluorescence label did not identify musculature. The regions of *lamina muscularis mucosae* and the other musculature that were identified in Figure 3B (white) are well overlaid with the Anti-desmin staining (Figure 3D + F). This shows nicely, that musculature can be identified with at least the same accuracy as with the IHC marker desmin-rhodamine. However, as seen in Figure 3C desmin-rhodamine labels also musculature in the mucosa. Therefore we overlay in Figure 3E in addition the RF class of mucosa. Figure 3E shows that the SHP can distinguish even better between pure musculature and musculature embedded in the mucosa, shown in yellow. In Figure 3F, connective tissue is shown which seems to be also labeled by desmin-rhodamine as well.

The infrared-based SHP method, on the other hand, is superior in distinguishing these protein classes. It assigned musculature and connective tissues to different classes, because SHP is a method with specific sensitivity to protein secondary structure. Musculature consists of actin and myosin proteins. Actin is a globular protein with significant α -helical contributions, whereas myosin is mostly a α -helical protein. Such structures show spectral patterns very different from the collagen in connective tissue because collagen exhibits distinctly different patterns in the infrared spectra, and thus, can readily be distinguished.

In summary, the comparison of the SHP image with a typical used fluorescence dye clearly demonstrate that tissue classes that show similar response to the Anti-desmin stain, can be more detailed distinguished in SHP. This example demonstrates the advantages of using an inherent signature, rather

than external labels, to distinguish cellular components. In addition, SHP detected collagen signatures in several clusters that were subsequently described by the pathologist as “connective tissue and connective tissue with supporting cells.

3.4 SHP – IHC correlation

An even more detailed comparison between SHP and fluorescence labels is given in Figure 4. Finding the correct translation, rotation and scaling was accomplished by visually overlaying the measured regions in standard image processing software. The area discussed in detail in Figure 4 is marked blue in Figure 2. Figure 4 represents the first correlation between markers typically used in multi-panel IHC and SHP. In particular, we show here that the regions identified by SHP as being cancerous (red areas in the SHP panel, 3B + 4D) co-localize with areas which show positive response to two commonly used IHC cancer markers, p53 and Ki-67.

We use the term “cancer marker” somewhat loosely since Ki-67 is actually a proliferation marker rather than a cancer marker; however, it often is used in multi-panel IHC as a surrogate cancer marker. At present, cancer diagnostics frequently utilized multi-panel IHC to provide the oncologist with treatment options for the disease. In colon cancer, the cocktails may include p53, Ki-67, CK20, MSH and others. Here, we used Ki-67 and p53 monoclonal antibodies. Ki-67 is a protein that is found in the G1, S, G2 and M phase of dividing cells, but not in resting (G0) cells. Thus, it has frequently been used to monitor the percentage of cells undergoing division (the ‘Ki-67 labelling index’), and therewith, is an important marker for mitotic activity and disease outcome. P53 is a cancer suppressing and apoptosis-inducing protein; in many cancers, this protein is found disabled or damaged due to mutations which can be due by a variety of causes, for example, infection by the human papillomavirus (HPV). Staining for abnormal p53 expression is commonly carried out for cancer diagnostics.

The results presented in Figure 4, for the first time, provide a direct visualization of how SHP can detect microscopic cancerous regions, and aid the pathologist in arriving at a more quantitative diagnosis. The left halves of all panels shown in Figure 4 depict areas of the *lamina propria mucosae* with embedded crypts. These crypts show thickened cell layers in the H&E image, but strong p53 response (Panel B) and moderate Ki-67 response (Panel C). However, the SHP image (Panel D) detects significant abnormality in these goblet cells of the crypts. One should remember at this point that the RF classifier was trained to recognize cancer patterns, based

on thousands of spectra diagnosed to origin from cancerous sites. Thus, the emergence of these spectral patterns in the crypts indicates the presence of abnormality in these cells, which is detected independently by the p53 and Ki-67 markers. Furthermore, it is interesting to note that the areas that appear abnormal in the H&E image (see arrows) show strong SHP response, but virtually no p53 and Ki-67 response. This indicates that within the cancerous regions, significant heterogeneity exists. This heterogeneity explains why some cells respond strongly to the p53 marker, while other cells in the same vicinity, although morphologically similar, may not. Furthermore, for most fluorescent labels no information about sensitivity and specificity is available, and the interobserver variability is quite high. In the SHP the de-paraffinization is the only step of sample preparation left so the interobserver variability is minimized. However, the SHP can also be applied to fresh or frozen samples and has the potential to be applied in colposcopy. The use of different instruments (Hyperion versus Cary) has no effect either, as demonstrated further by Supplementary Figure 7. It shows clear cut that SHP is platform independent.

The parallel sensitivity between Ki-67 and SHP is not unexpected, since earlier work demonstrated that infrared spectra of dividing cells have a distinct spectral signature. In particular, it was found that nucleic acid signals were strongest in the S-phase of the cell cycle. This observation was attributed to changes in DNA condensation. Thus, it is not surprising that Ki-67 and SHP are similarly sensitive toward cell proliferation.

The distinction of precancerous and cancerous areas of the section shown in Figures 4 is highly intriguing. Stains for collagen III and musculature did not detect any differences between the regions of the carcinoma and the relatively normal appearing mucosa (see Figure 3). The H&E-stained image, of course, shows distinct areas of the cancer, which are represented in the SHP image as well (the thin, red regions lining the crypts, embedded in inflammatory tissue, at the top left quadrant of Figure 3B). The p53 and Ki-67 immunohistochemistry showed distinct regions of activity within the cancerous region which did co-localize with the H&E and SHP images, see Figure 5. The regions identified as cancerous by SHP exhibit spectra that are different from that of normal glandular tissue by a slight increase of the nucleic acid-related vibrations (ca. 1090 cm^{-1} and 1235 cm^{-1}) and slight changes in the amide I band profile. These spectral changes, in general, are not very distinct when one inspects the spectra visually, but sufficiently pronounced to be detected by diagnostic algorithms. Other forms of cancers (lymphoma, small cell lung carcinomas [34]) have much more pronounced changes in the nucleic acid regions. However, the changes described here

have been observed throughout most SHP-based cancer examples.

This co-localization between IR- and IHC-based images is further emphasized in Figure 5, which shows a small area of crypts that responded strongly to the anti-p53 and the anti-Ki-67 markers. SHP picks up these regions and associates them with cancerous signatures. Normally, the crypts outside the cancerous region have no pathological result, but here it seems that the crypts are already infiltrated from the adjacent cancer, as shown by both IHC and SHP.

Beyond confirming earlier observations reported by Lasch [19]–[23], our contribution significantly expanded the scope of these earlier studies. We showed that SHP is platform independent and used a novel robust and fast bioinformatics approach to analyse the spectral data. In the much larger tissue sections measured in the present study, tissue types were encountered that were not included in Lasch's work. Furthermore, one of the major aims of the present study is to demonstrate that SHP and IHC detect regions of cancer activity. Previous studies by others have demonstrated that regions of high cellular activity, which can be detected in H&E images, correlate with distinct infrared spectral patterns. Extending these previous results, we demonstrate here that regions identified by SHP to contain cancer cells correlate with IHC staining patterns obtained from several markers applied to the same sample. This observation is indicative of the sensitivity of SHP to detect and stage cancerous regions of tissue, exemplified by the abnormalities that we observe in the crypts embedded in the *lamina propria mucosae*.

4. Conclusions

Early results in SHP have demonstrated that this method can distinguish anatomical regions in tissue sections which subsequently could be associated with different tissue types or areas of disease. The distinction of different tissue types can be fairly straightforward: any tissues rich in connective collagen (e.g., fibro-connective tissue, stroma, tendons) show distinct spectral signatures of triple-helical collagen (see below). Similarly, mucin-producing glandular tissue often shows distinct glycoprotein signatures within the cell body, and erythrocytes are nearly completely dominated by hemoglobin spectral features.

In the present study, we demonstrate that the different tissue structures that occur in the colon wall present distinctly different spectro-chemical signatures. In our contribution, the structures identified spectrally, the *lamina propria mucosae*, the *lamina muscularis mucosae*, the crypts and lumen filled with

mucus, connective tissue and cancer, were classified by a random forest algorithm that was trained on tissue sections from different patients, and from a tissue micro-array. However, the results presented here are not to be viewed as a fully comprehensive dataset but rather as a dataset demonstrating the excellent correspondence between SHP, H&E images, and fluorescence images based on several immunohistochemical stains. The prospect of observing individual cancer markers, or their effects on the spectral signatures, presents novel avenues toward diagnostics and prognostics of disease in general, and for cancer in particular. First and foremost, the infrared spectrum of a tissue pixel contains a snapshot of the entire proteome, metabolome and genome of the pixel; this snapshot includes overexpressed genes, oncogenes, protein and non-protein markers of disease (such as elevated glucose levels in diabetes) and biological consequences, such as apoptosis. Since this information is stored digitally in a virtual 'tissue archive' it is possible to retroactively data-mine this databank for cancer markers, disease outcome, and effects of therapeutic intervention.

In order to pave the way for eventual clinical applications of spectral histopathology, readily available and inexpensive substrates need to be utilized. Kevley ("low-e") slides fulfil these criteria. The line shape distortions experienced in infrared spectroscopy of non-homogeneous samples are not due to the measurement mode (transflection vs. transmission), but occur due to interference of particulate samples. The theory of such interferences has been discussed in detail by Kodali and Bhargava [46] and Davis et al. [47].

Second, the spectral information from tissues or body fluids from any (healthy or sick) person may become part of a person's biochemical signature, and can be used in screening and monitoring approaches. Efforts to detect the overall state of health, and possibly the very early onset of disease, from infrared spectral data of body fluids have been reported in the past, and are research areas actively pursued at the authors' laboratory.

Our contributions on the data analysis side suggest that Random Forests, following their increasing popularity in other areas, are an attractive alternative to other supervised classifiers previously used for infrared imaging, mainly due to their simplicity not requiring feature as well as their overall good performance and robustness against overfitting. Presenting the first quantitative evaluation of clustering algorithms applied in infrared imaging, our novel cross-validation scheme for HCA may be a useful tool for systematic comparison of future approaches for infrared image annotation. Quite remarkable, all previous studies were conducted qualitatively, while our approach introduces more objective quantitative tools.

Finally, the fact that this information can be collected in a completely objective and reproducible instrument-based approach, and analyzed by algorithms trained on a combination of all diagnostic methods including pathology, oncology, immunohistochemistry and even gene sequencing, makes infrared spectral methodology a prospect of becoming a truly universal diagnostic and prognostic tool.

Acknowledgements We thank Sascha Krauß for tiling and background correction of fluorescent images and Qiaoyong Zhong for an implementation of the cross validation for hierarchical clustering. This research was supported by the Protein Research Unit Ruhr within Europe (PURE), project 131/1.08–031, Ministry of Innovation, Science and Research (MIWF) of North-Rhine Westfalia, Germany; and the Center for Vibrational Microscopy (CVM), project 005-GW01–139Z, European Regional Development Fund, European Union and North-Rhine Westfalia, Germany.

Author biographies Please see Supporting Information online.

References

- [1] Cancer Facts & Figures 2010, Atlanta: American Cancer Society. Available online: <http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-026238.pdf> (2010).
- [2] A. E. de Jong, M. van Puijenbroek, Y. Hendriks, C. Tops, J. Wijnen, M. G. E. M. Ausems, H. Meijers-Heijboer, A. Wagner, T. A. M. van Os, A. H. J. T. Bröcker-Vriends, H. F. A. Vasen, and H. Morreau, *Clin. Cancer Res.* **10**, 972–980 (2004).
- [3] A. de la Chapelle, *N. Engl. J. Med.* **349**, 209–210 (2003).
- [4] S. Munemitsu, I. Albert, B. Souza, B. Rubinfeld, and P. Polakis, *Proc. Natl. Acad. Sci. USA* **92**, 3046–3050 (1995).
- [5] B. T. Wilhelm, S. Marguerat, S. Watt, F. Schubert, V. Wood, I. Goodhead, C. J. Penkett, J. Rogers, and J. Bähler, *Nature* **453**, 1239–1243 (2008).
- [6] A. J. Mighell, W. J. Hume, and P. A. Robinson, *Oral Dis.* **4**, 217–223 (1998).
- [7] J. A. Ramos-Vara, *Vet. Pathol.* **42**, 405–426 (2005).
- [8] M. Diem, M. Miljkoviæ, B. Bird, T. Chernenko, J. Schubert, E. Marcisins, A. Mazur, E. Kingston, E. Zuser, K. Papamarkakis, and N. Laver, *Spectrosc. Int. J.* **27**, 463–496 (2012).
- [9] L. Chiriboga, P. Xie, H. Yee, V. Vigorita, D. Zarou, D. Zakim, and M. Diem, *Biospectroscopy* **4**, 47–53 (1998).
- [10] B. R. Wood, L. Chiriboga, H. Yee, M. A. Quinn, D. McNaughton, and M. Diem, *Gynecol. Oncol.* **93**(4), 59–68 (2004).
- [11] R. Bhargava, D. C. Fernandez, M. D. Schaeberle, and I. W. Levin, *PittCon. 2002*, New Orleans, Louisiana, USA, 17–22 March 2002.
- [12] W. Steller, J. Eienkel, L.-C. Horn, U.-D. Braumann, H. Binder, R. Salzer, and C. Krafft, *Anal. Bioanal. Chem.* **384**, 145–154 (2006).
- [13] C. Krafft and R. Salzer, *Neuro-oncological Applications of Infrared and Raman Spectroscopy*, in: *Vibrational Spectroscopy for Medical Diagnosis*, M. Diem, P. R. Griffiths, and J. M. Chalmers (eds.) (John Wiley & Sons, Chichester, UK, 2008).
- [14] N. Amharref, A. Beljebbar, S. Dukic, L. Venteo, L. Schneider, M. Pluot, R. Vistelle, and M. Manfait, *Biochem. Biophys. Acta* **1758**(7), 892–899 (2006).
- [15] M. J. Romeo, and M. Diem, *Vib. Spectrosc.* **38**, 115–119 (2005).
- [16] B. Bird, M. Miljkoviæ, S. Remiszewski, A. Akalin, M. Kon, and M. Diem, *Lab. Invest.*, **92**, 1358–1373 (2012).
- [17] E. Gazi, M. Baker, J. Dwyer, N. P. Lockyer, P. Gardner, J. H. Shanks, R. S. Reeve, C. A. Hart, N. W. Clarke, and M. D. Brown, *Eur. Urol.* **50**(4), 750–761 (2006).
- [18] C. M. Krishna, G. Kegelaer, I. ADT, S. Rubin, V. B. Kartha, M. Manfait, and G. D. Sockalingum, *Biopolymers* **85**(5), 462–470 (2006).
- [19] P. Lasch, *Computergestuetzte Bildrekonstruktion auf Basis FTIR-mikrospektrometrischer Daten humaner Tumoren (Humanmedizin, Freie Universität Berlin, Germany, 1999)*, p. 124.
- [20] P. Lasch, W. Haensch, E. N. Lewis, L. H. Kidder, D. Naumann, *Appl. Spectrosc.* **56**, 1–9 (2002).
- [21] P. Lasch, W. Haensch, D. Naumann, and M. Diem, *Biochim. Biophys. Acta* **1688**(2), 176–86 (2004).
- [22] P. Lasch, and D. Naumann, *Cell. Mol. Biol.* **44**(1), 189–202 (1998).
- [23] P. Lasch, M. Diem, W. Hänsch, and D. Naumann, *J. Chemom.* **20**(5), 209–220 (2007).
- [24] D. C. Fernandez, R. Bhargava, S. M. Hewitt, and I. W. Levin, *Nat. Biotechnol.* **23**, 469–474 (2005).
- [25] F. N. Pounder and R. Bhargava, in: *Toward Automated breast Histopathology Using Mid-IR Spectroscopic Imaging*, in *Vibrational Spectroscopic Imaging for Biomedical Applications*, G. Srinivasa (ed.) (McGraw Hill, New York, USA, 2010).
- [26] M. Diem, K. Papamarkakis, J. Schubert, B. Bird, M. J. Romeo, and M. Miljkoviæ, *Appl. Spectrosc.* **63**(11), 307A–318A (2009).
- [27] L. Hartsuiker, N. J. L. Zeijen, L. W. M. M. Terstappen, and C. Otto, *Analyst* **135**, 3220–3226 (2010).
- [28] N. S. Ozek, S. Tuna, A. E. Erson-Bensan, and F. Severcan, *Analyst* **135**, 3094–3102 (2010).
- [29] *antikoerper-online.de*. Available online: <http://www.antikoerper-online.de/resources/17/867/Labeling+Kits/>
- [30] P. Bassan, H. J. Byrne, F. Bonnier, J. Lee, P. Dumas, and P. Gardner, *Analyst* **134**, 1586–1593 (2009).
- [31] P. Bassan, A. Kohler, H. Martens, J. Lee, H. J. Byrne, P. Dumas, Ehsan Gazi, M. Brown, N. Clarke, and P. Gardner, *Analyst* **135**, 268–277 (2010).
- [32] A. Savitzky and M. J. E. Golay, *Anal. Chem.* **36**(8), 1627–1639 (1964).
- [33] H. Xiao, Y. Li, J. Du, and A. Mosig, *Bioinformatics* **27**, 564–571 (2011).

- [34] X. Wu, M. G. Piper-Hunter, M. B. S. Crawford, G. J. Nuovo, C. B. Marsh, G. A. Otterson, and S. P. Nana-Sinkam, *J. Thorac. Oncol.* **4**(8), 1028–1034 (2009).
- [35] S. F. Chew, B. R. Wood, C. Kanaan, J. Browning, D. MacGregor, I. D. Davis, J. Cebon, B. D. Tait, D. McNaughton, *Tissue Antigens* **69**, 252–258 (2007).
- [36] N. Amharref, A. Beljebbar, S. Dukic, L. Venteo, L. Schneider, M. Pluot, M. Manfait, *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1768**(10), 2605–2615 (2007).
- [37] L. Breiman, *Random Forests. Machine Learning* **45**(1), 5–32, (2001).
- [38] R. Diaz-Uriarte and S. A. De Andres, *BMC Bioinformatics* **7**(1), 3 (2006).
- [39] P. Bassan, A. Sachdeva, A. Kohler, C. Hughes, A. Henderson, J. Boyle, J. H. Shanks, M. Brown, N. W. Clarke, and P. Gardner, *Analyst* **137**(6), 1370–1377, (2012).
- [40] A. Statnikov, L. Wang, and C. F. Aliferis, *BMC Bioinformatics* **9**(1), 319 (2008).
- [41] J. C. Vickerman, *Analyst* **136**(11), 2199–2217, (2011).
- [42] D. R. Ifa, C. Wu, Z. Ouyang, and R. G. Cooks, *Analyst* **135**(4), 669–681 (2010).
- [43] L. J. Jensen and A. Bateman, *Bioinformatics* **27**(24), 3331–3332 (2011).
- [44] H. Peng, F. Long, and C. Ding, *IEEE Transact. Pattern Analysis Machine Intellig.* **27**(8), 1226–1238 (2005).
- [45] Q. Wang, W. Sanad, L. M. Miller, A. Voigt, K. Klingel, R. Kandolf, K. Stangl, and G. Baumann, *Vib. Spectrosc.* **38**(1), 217–222 (2005).
- [46] A. K. Kodali, R. Bhargava, *Theor. Chem. Accounts* 1–10 (2011).
- [47] B. J. Davis, P. S. Carney, and R. Bhargava, *Anal. Chem.* **82**(9), 3487–3499 (2010).
- [48] <http://code.google.com/p/randomforest-matlab/>
- [49] http://www.dcs-diagnostics.de/data/IHC_Entparaffinierung_web.pdf