BBA - Molecular Basis of Disease 1868 (2022) 166279

Contents lists available at ScienceDirect

BBA - Molecular Basis of Disease

journal homepage: www.elsevier.com/locate/bbadis

Identification of remodeled collagen fibers in tumor stroma by FTIR Micro-spectroscopy: A new approach to recognize the colon carcinoma

Serena De Santis [a,*], Francesco Porcelli [a], Giovanni Sotgiu [a], Anna Crescenzi [b], Anita Ceccucci [a], Martina Verri [b], Marco Caricato [c], Chiara Taffon [b], Monica Orsini [a]

[a] Department of Engineering, Roma Tre University, via Vito Volterra 62, Roma, Italy
[b] Pathology Unit, University Hospital Campus Bio-Medico, Rome, Italy
[c] Colorectal surgery Unit, University Campus Bio-Medico of Rome, Italy

ARTICLE INFO

Keywords:
Collagen
Stroma
colon cancer
micro-FTIR spectroscopy

ABSTRACT

The tumor stroma plays a pivotal role in colon cancer genesis and progression. It was observed that collagen fibers in the extracellular matrix (ECM) of cancer stroma, undergo a strong remodeling. These fibrous proteins result more aligned and compact than in physiological conditions, creating a microenvironment that favors cancer development.

In this work, micro-FTIR spectroscopy was applied to investigate the chemical modifications in the tumor stroma. Using Fuzzy C-means clustering, mean spectra from diseased and normal stroma were compared and collagen was found to be responsible for the main differences between them. Specifically, the modified absorptions at 1203, 1238, 1284 cm^{-1} and 1338 cm^{-1} wavenumbers, were related to the amide III band and CH_2 bending of side chains. These signals are sensitive to the interactions between the α-chains in the triple helices of collagen structure. This provided robust chemical evidence that in cancer ECM, collagen fibers are more parallelized, stiff and ordered than in normal tissue. Principal Component Analysis (PCA) applied to the spectra from malignant and normal stroma confirmed these findings. Using LDA (Linear Discriminant Analysis) classification, the absorptions 1203, 1238, 1284 and 1338 cm^{-1} were examined as spectral biomarkers, obtaining quite promising results. The use of a PCA-LDA prediction model on samples with moderate tumor degree further showed that the stroma chemical modifications are more indicative of malignancy compared to the epithelium. These preliminary findings have shown that micro-FTIR spectroscopy, focused on collagen signals, could become a promising tool for colon cancer diagnosis.

1. Introduction

Solid cancers are structures consisting of cancer cells immersed in a complex microenvironment. In the past, cancer research considered the tumor development as an autonomous process of cancer cells. In recent years instead, it has become evident that the microenvironment, called tumor stroma, is actively involved in the growth and progression of tumor [1]. Many studies highlighted the pivotal role of the stroma which is composed of different kinds of cells such as fibroblasts, adipocytes, endothelial cells, immune cells, and the extracellular matrix (ECM) [2,3].

The ECM can be mainly classified into structural proteins, including collagens and elastin, glycoproteins, proteoglycans, and polysaccharides. This complex matrix structure of noncellular components with its biochemical and physical properties supports the architecture of the tissue. In addition, it can take part in some crucial processes related to the cell behavior and development [4]. Based on this strong interaction between the ECM and the growing cancer cells, there is increasing interest around the characteristics and role of the ECM [5]. It is widely reported that the ECM macromolecules change and reorganize in most solid tumors including colorectal cancer [6,7]. Within these processes, collagen fibers, which are the major components of the ECM, are strongly altered [8,9].

Collagen consists of three polypeptide chains, called α-chains, which form a right-handed triple-helical structure. Each one of the α-chains contains the repeating amino acid motif Gly-X-Y in which the X and Y can be any amino acid, although in these positions proline and hydroxyproline respectively are the most common. The triple helical

* Corresponding author.
E-mail address: serena.desantis@uniroma3.it (S. De Santis).

https://doi.org/10.1016/j.bbadis.2021.166279
Received 10 June 2021; Received in revised form 20 September 2021; Accepted 22 September 2021
Available online 30 September 2021

structures can be arranged forming fibrillar collagens of different types, and in the interstitial ECM the main fibrous protein is the type I fibrillar collagen [10].

Under physiological conditions, stroma cells such as fibroblasts produce collagen chains. After their synthesis, these macromolecules are subjected to significant post-translational modifications such as hydroxylation, glycosylation of lysine and hydroxylysine residues. These procollagens in turn form the triplex helices which finally lead to the formation of collagen fibrils after the secretion into the extracellular space and further transformations. At this level, enzymes, such as lysyl oxidase (LOX), catalyze the formation of covalent cross-links among the fibrils to obtain the collagen fibers.

Colorectal carcinoma is the most common type of gastrointestinal cancer, and its incidence is rising worldwide [11]. Establishing risk factors, biological pathways, and therapeutic strategies is a current goal of cancer research. It has been recently reported that in colon carcinoma tumor cells interact with fibroblasts which deposit an increased amount of collagen. Many studies showed that the aforementioned processes related to the formation of collagen fibrils are altered. Moreover, LOX resulted overexpressed, creating an increase in cross-links [12,8,1]. As a result of these alterations on the procollagen chains and in the assembly of fibrils, collagen fibers undergo a significant architectural remodeling. They assume a more compact and aligned arrangement and the ECM becomes stiffer and more fibrous than in physiologic conditions. This important restructuring of collagens affects cellular differentiation, proliferation and migration and in turn tumor development and treatment response [13].

In this context, it is evident that a careful investigation and detection of biochemical changes undergone by the collagen fibers is an important topic. This could help to understand the physiologic aspects of tumor development and furthermore, through an early identification of collagen alterations, to diagnose diseased conditions.

Different approaches were reported to detect the modifications of collagen fibers in colon tumor ECM. Immunohistochemical analysis has been applied, though generally it requires specific and expensive reagents with an accurate preparation of sample [14]. Some chemical analytical procedures have been identified to extract the collagens' biomolecules, but these methods can be complex and destructive [15]. Indirect techniques which give insight into the elastic behavior of ECM have also been used, as shear wave elastography [16] and the Atomic Force Microscopy (AFM) [12] although the risk of introducing artifacts remains an issue in these cases.

Fourier Transformed Infrared spectroscopy (FTIR) is a non-destructive and reliable technique which gives information at molecular level of analyzed materials and tissues based on the chemical bond vibrations of molecules [17–18,19]. The combination of infrared spectroscopy and microscopy (micro-FTIR) permits a spatial biochemical investigation of tissues. IR maps can be obtained in a label-free manner each pixel holds chemical information from the associated spectrum [20].

In this work micro-FTIR was used to identify and evaluate the chemical differences of ECM in malignant and normal stroma in patients diagnosed of the adenocarcinoma colon tissue. The spectral maps of diseased and normal tissues were recorded, next by Fuzzy-Clustering analysis, which groups spectra with similar biochemical characteristics, digitally colored images related to histological features were built.

The average spectra from the clusters assigned to the malignant and normal stroma were compared to obtain information on the biochemical differences and to evaluate relationships with the collagen fibers.

To further investigate the role of collagen fibers, Principal Component Analysis (PCA) was also applied to all spectra from normal and malignant stroma of analyzed colon tissues.

Using PCA and Linear Discriminant Analysis (LDA) as a prediction model, spectral images were constructed. Interesting results related to the tumor stroma were attained from the diseased samples in which the degree of malignancy was moderate.

In the long term, micro-FTIR technique based on collagen fibers signals' examination in the stroma, could be used as a non-invasive, label-free, rapid and direct clinical tool for the early detection of colon cancer.

2. Materials and methods

2.1. Sample preparation

Samples used in this work were obtained from human colon specimens surgically removed after diagnosis of adenocarcinoma on endoscopic biopsy from six patients. From each patient areas of normal colon tissues and areas of different degree (moderate to severe) of malignancy were selected and identified by a pathologist. Two consecutive slices of 10 μm were obtained from the frozen tissue samples. The first slice was stained with hematoxylin and eosin (HE) for the pathological exam and used as morphological reference. The adjacent slice was directly mounted onto a calcium fluoride (CaF_2) IR window and examined under FTIR microscope. For each sample, representative areas were selected, microscope images were captured, and corresponding IR maps were recorded.

All the materials and solvents used in this work were purchased from Merk Life Science (Milano, Italy) and used as received.

The study was approved by Ethics Committee of University Campus Bio-Medico, prot 33.15 TS ComEt CBM

2.2. Data acquisition

Infrared spectra were collected by a Nicolet iN10 infrared microscope (Thermo Fisher) equipped with a Mercury-Cadmium-Telluride (MCT-A) nitrogen-cooled detector, in transmission mode. Maps were recorded in the 4000–650 cm^{-1} range, using a 15×15 μm^2 aperture and a step size of 15 μm, both on the x and y axis, with a spectral resolution of 4 cm^{-1}. Sixty-four (64) interferograms were averaged per pixel spectrum and apodized using a Blackman-Harris correction. A background spectrum was collected on a tissue-free area of CaF_2 before any map acquisition. A minimum of 25 to a maximum of 56 spectra for x and/or y direction was acquired.

2.3. Data analysis and processing

2.3.1. Data pre-processing

Spectral pre-processing and multivariate data analysis were performed with Matlab R2019b (The MathWorks, Inc.) [21] in the 1000–1800 cm^{-1} range. Raw FTIR spectra exhibiting a too low signal-to-noise ratio due to the presence of empty or damaged areas in the analyzed tissue were discarded applying an intensity filter on amide I and appear as black point or areas in images. Spectra were then subjected to EMSC algorithm [22,23]. Baseline correction using the rubberband method followed by min-max (Amide *I*) normalization was then performed [24]. Spectra comparison was carried out using Savitzky–Golay second derivative (eleven points smoothing).

2.3.2. Clustering analysis

Unsupervised, Fuzzy c-means (FCM) clustering analysis was applied on pre-processed spectra in order to collect data based on their histological regions grade of membership. In FCM each pixel of the chemical map is encoded based on their cluster membership probability, thus it was possible to assign each spectrum to multiple clusters with a cluster participation from 0 to 1. Compared to hard clustering K-means, FCM allowed better representation of tissue regions, in particular at the boundaries of two different histological components with high quality chemical image segmentation. In this work, a number of clusters varying from 2 to 3 with a fuzzy's index of 2 [25,26] was used, in order to distinguish only epithelium from stroma. A third cluster was applied solely when tissue regions damaged by the preparation procedure where

present. Spectra from stroma and epithelium were extracted by FCM chemical maps considering a cluster membership probability >0.8. A total of 3649 spectra for normal and malignant epithelium and 7743 for normal and malignant stroma were obtained.

2.3.3. Principal component analysis (PCA)

Principal component analysis (PCA) is an algorithm used to extract relevant information from confusing data sets highlighting similarities and differences. The original data set of variables is reduced to smaller number of orthogonal variables called the principal components (PCs). In the analysis of FTIR spectra by PCA the scores report the contribution of each spectrum related to the selected principal component and the corresponding loadings show which wavenumber is responsible for such contribution. In this work the PCA was carried out on all data from malignant and normal stroma extracted from the samples by FCM. Two different spectral ranges 1000–1800 and 1160–1400 cm^{-1} were analyzed.

2.3.4. LDA classification

The loadings plot related to PCA applied to all data from malignant and normal stroma in the spectral range 1160–1400 cm^{-1}were used for linear discriminant analysis (LDA) classification. These absorbances were subjected to iterative one-way ANOVA test. 100,000 iterations were carried out and at each one ten spectra of malignant and normal stroma were randomly selected. The one-way ANOVA test was then performed on the selected variables and *p*-value obtained was recorded at each iteration. For all selected wavenumbers the test gave a p-value <0.05 at least 99,900 of times. Data were divided in training (70%), validation (15%) and test set (15%) using Kennard-Stone algorithm [27,28]. After evaluation of model robustness on internal validation and test set, the prediction model was applied on spectra recorded from samples of four different patients employed as external data.

2.3.5. Supervised PCA-LDA prediction model

A subset of spectra recorded on normal and malignant tissue samples from two different patients was used to build PCA-LDA model. Four different classes were considered based on specific histopathological features: normal stroma, normal epithelium, malignant stroma and malignant epithelium. Dataset was split into training, validation and test set with the same procedure described above. PCA was applied on training set for dimensional reduction corresponding to 8 PCs (> 98% of cumulative explained variance) obtained with venetian blind cross validation (five groups) in the 1000–1800 cm^{-1} range. The robustness of the model built with the data corresponding to training was evaluated on internal and validation set. The model was applied on spectra recorded on samples from four different patients as external data set and prediction performance was evaluated as predicted chemical maps where each colour (blue-normal stroma, cyan -normal epithelium, yellow-malignant stroma and red-malignant epithelium) correspond to specific predicted histopathological class.

3. Results and discussion

3.1. Clustering analysis

Fuzzy C-means clustering (FCM) was applied to determine the spectral differences derived from the main histological characteristics of normal and cancerous colon tissues. In FCM a spectrum can belong to more than one cluster and its percentage of participation is based on its degree of similarity. Therefore, the number of clustering images produced is equal to the number of the clusters established to represent the analyzed sample. This approach — called “soft” — finely refers the chemical information contained in each spectrum, and an effectively histological structure representation can be achieved. This method was applied to the spectral data in the range of 1000–1800 cm^{-1} extracted from samples of normal and malignant colon tissue, and the proper digital stained images for each sample are constructed (Fig. 1). In this study for non-tumoral tissue as well as the tumoral colon tissue, only two clusters were chosen which allowed to effectively distinguish between the two main histological structures of the colon tissue: stroma (Fig. 1, cluster 1) and epithelium (Fig. 1, cluster 2). In cluster 1 images corresponding to samples **1n** (normal) and **1m** (malignant), red pixels are associated with spectra possessing stromal chemical composition. On the contrary, in the cluster 2 images the red pixels are assigned to normal and malignant epithelium composition. The effectiveness of this clustering method can be verified by comparing the optical microscope images **1n** and **1m** with the corresponding stained clustering images. This approach was used for all analyzed samples, as shown in Figs. S1 and S2 in Supporting Information.

3.2. Spectral analysis

In this work, we focused on the chemical changes which take place in the stroma malignant tissue to evaluate modifications of the biomolecules of the extracellular matrix. IR signals in the spectra associated with the malignant stromal tissue were then compared with their counterparts in normal stroma. Consequently, relying on the Fuzzy C-means clustering, the spectra with a percentage of participation equal to or greater than 0.8 of all clusters 1 (Fig. 1, S1 and S2 in Supporting Information) corresponding to normal and malignant stroma of analyzed samples were accurately extracted. The average spectra (red malignant and blue normal), considering the standard deviation of these two spectra groups, were obtained and compared (Fig. 2A). It was evident that the most discriminant wavenumbers between the two average spectra were present in the 1100 cm^{-1} to 1500 cm^{-1} region. In this spectral range significant vibrational modes of important biomolecules involved in the chemical modifications of cancer tissue can be found, as already observed in other studies [29–30,31]. To better identify the values of these wavenumbers, second derivative of both average spectra was considered, as reported in Fig. 2C. By comparison of these two curves, the signals 1454, 1396, 1338, 1284, 1238 and 1203 cm^{-1} can be identified as discriminating between the malignant and the normal stroma. This set of absorptions can be assigned to the characteristic vibration modes of collagen fibers [32]. In this spectral region, the stretching vibrations of the phosphate groups of the DNA and RNA also appears, and many authors [33,28] have often mainly considered the variations of these signals. However, it is reasonable to consider the modifications of collagen absorptions are overriding, since in the stromal tissue the content of collagen is prevalent [34]. Specifically, the signals that changed in malignant stromal tissue are associated with two typical spectral intervals of collagens bands: 1300–1180 cm^{-1} and 1480–1350 cm^{-1}. In the first range, typically called the amide III region, the bands were associated with C—N stretching, N—H bending, C—C stretching and C—H bending vibrations, while in the second zone the signals were attributed to the C—H_2 and C—H_3 bending absorptions. The amide III zone gives information about the proteins secondary structure and, as shown in Fig. 2C, this range is characterized by three main signals at 1284, 1238 and 1203 cm^{-1} [35]. As can be seen, the positions of all these absorptions have not changed significantly passing from normal to malignant stroma, while a variation of intensities can be observed. Specifically, the peak centered at 1284 cm^{-1} results less defined in tumor stroma than in normal condition. Furthermore, the vibration at 1203 cm^{-1} almost disappears in the diseased tissue. As reported in literature, the signal at 1238 cm^{-1} is related to the N—H groups, which are involved in the hydrogen bonds with the C═O of the adjacent α-chain. In a similar way, the component at 1284 cm^{-1} can be assigned to the C_{α}-H of amino acid in Y position or to Gly residue which form hydrogen bonds with the C═O in the X position of the following α-chain. The signal at 1203 cm^{-1} is considered to be associated to the bending vibrations of amino acids side chains which contribute to the stability of the triple helices of proteins [36]. These signals are closely related to the network of interactions among the α-chains in the collagen

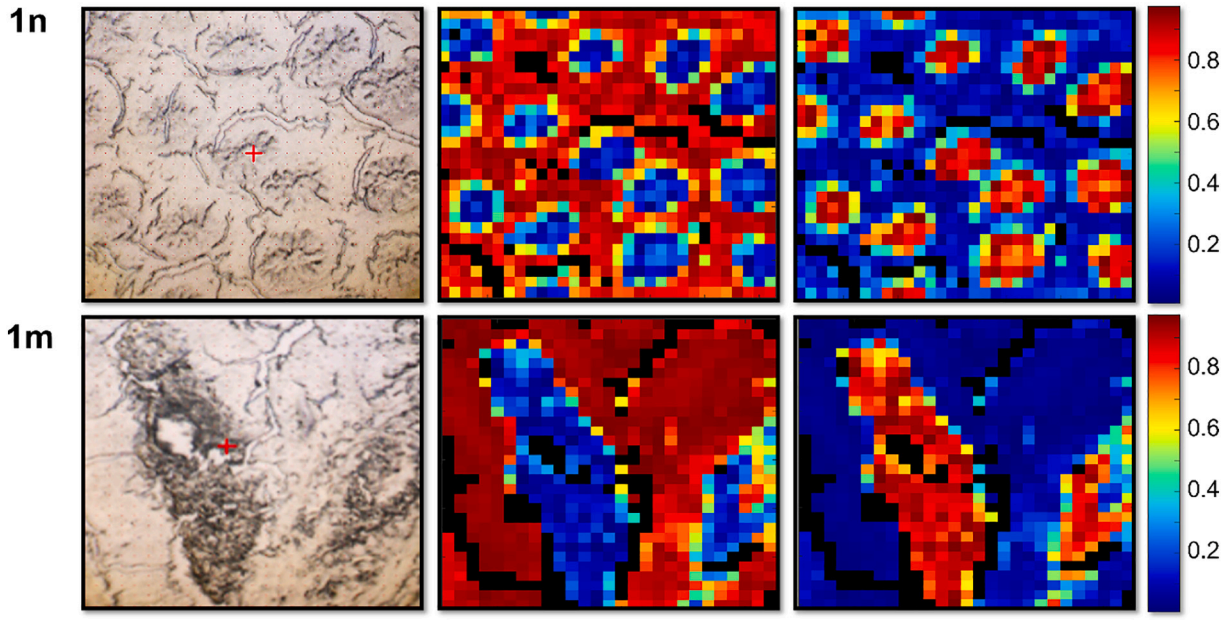

Fig. 1. FTIR microscope images of normal (1n) and malignant (1 m) tissue and the corresponding colored Fuzzy C-means images using two clusters to distinguish between stroma (cluster 1) and epithelium (cluster 2).

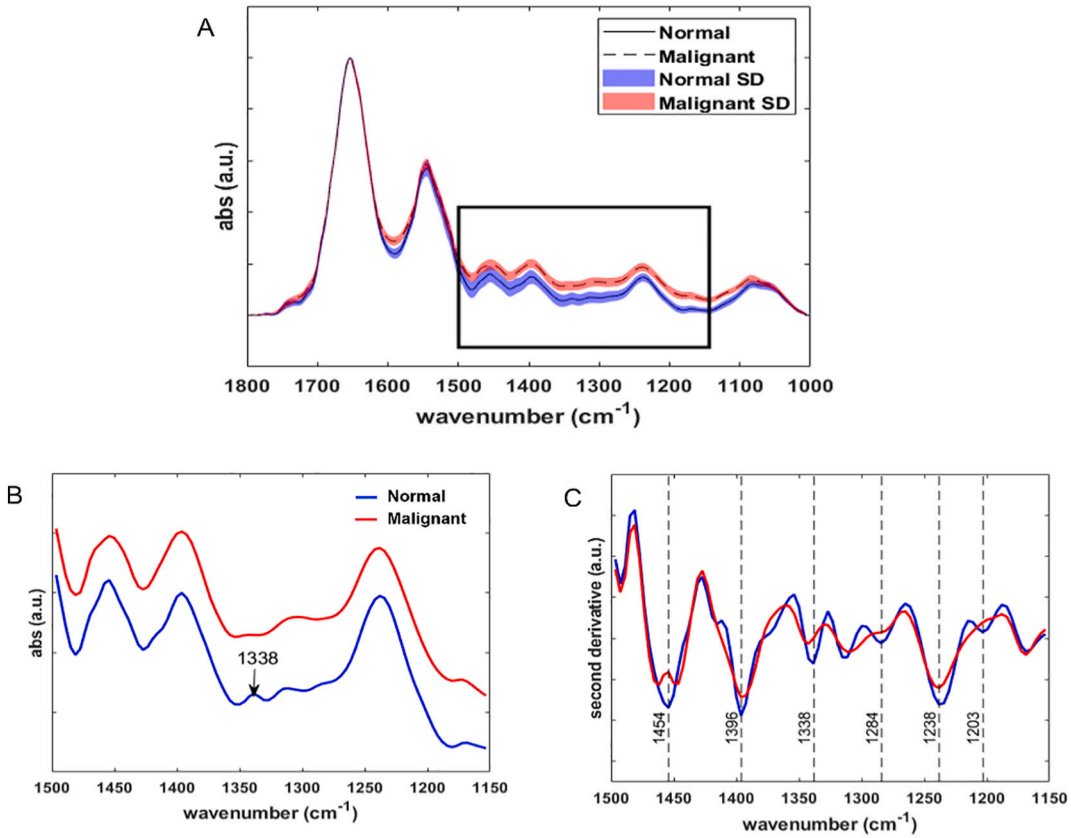

Fig. 2. (A) Average spectra of normal (blue line) and malignant (red line) stroma extracted by FCM chemical maps with standard deviation. (B) A zoom in the interval 1150–1500 cm^{-1} of average spectra of normal (blue line) and malignant (red line) stroma. (C) Second derivative of average spectra of normal (blue line) and malignant (red line) stroma spectra in range 1150–1500 cm^{-1}.

secondary structure. Thus, it can be asserted that their variations in tumor stroma highlights a modification of the triple helix assembly. Accordingly, the observed changes from normal to malignant stroma in second derivative spectra (Fig. 2C) for these three absorptions are similar to the results found in literature [35] considering the thermal denaturation of type I collagen. As reported, in this process the collagen undergoes a structural reorganization which tends to approach the antiparallel β-sheet structure. Based on the evaluation of these specific

signals it can be proved that in cancer stroma the α-chains in the triple helices result more aligned and form a more compact structure than in normal connective tissue.

A further evidence that confirms the alteration of interactions among the adjacent strands is the variation of the peak centered at 1338 cm^{-1} (Fig. 2B and C). This signal was assigned to the C—H_2 wagging vibration of proline [34]. The decrease of absorbance of this signal corresponds to a different arrangement of the α-chains which compose the triple helices in the collagen fibers [37,38].

The obtained results, fully in accordance with literature data obtained by indirect biochemical methods, underlined the formation of altered collagen due to the abnormal post-translational modifications in cancerous tissue. In this way the assembled collagens fibrils are mostly parallelized, ordered and compact, giving rise to an ECM stiffer and denser than in healthy tissue.

IR spectroscopy can identify directly and unequivocally these chemical alterations of collagens α-chains and it can represent a useful and reliable analytical technique to detect the malignancy-associated features in colon adenocarcinoma.

3.3. Principal component analysis (PCA) and LDA classification

To verify if it would be possible to discriminate between normal and cancerous tissue in colon adenocarcinoma only based on stromal tissue investigation, principal component analysis (PCA) was carried out. The spectra used in this analysis were all those associated with normal and malignant stromal region identified by Fuzzy C-means clustering in all analyzed samples (Figs. 1, S1 and S2 in SI). The spectral interval between 1000 and 1800 cm^{-1}, which includes the Amide *I* protein band, was initially selected. The obtained PCA score plot PC1 *vs* PC2 showed a good separation of the normal and malignant colon tissue with a high explained variance for PC1 component (Fig. S3). However, as expected, when the selected spectral range was reduced to 1150–1400 cm^{-1} the explained variance for PC1 component increased to 96% (Fig. 3A). Indeed, as observed by the comparison of stroma average spectra of the normal and malignant tissue, this region provides most of the information on the biochemical changes.

Considering the loadings associated to this range, the most discriminant wavenumbers are once again the signals at 1203, 1238, 1284 and 1338 cm^{-1} related to collagens (Fig. 3B). These results further confirmed that the main chemical modifications that took place in the malignant stroma are effectively ascribable to collagen fibers.

Based on their predominant role in discriminating healthy from diseased condition, the spectral signatures related to the stroma tissue identified by PCA (1203, 1238, 1284 and 1338 cm^{-1}), were subjected to linear discriminant analysis (LDA). This analysis allowed to evaluate if they can be considered as stand-alone diagnostic tool in the identification of colon adenocarcinoma. Fig. 3C shows the scatter plot of the linear discriminant scores of normal and malignant stroma spectra from external raw data considering four different patients. The prediction performance on the external set is also shown in more details in form of a confusion matrix in Fig. 4 (the prediction accuracy evaluated on internal validation and test set was 100%, see Fig. S4 of SI). For all samples, a good agreement between true and predicted classes was obtained.

Despite these preliminary results required to be reinforced by testing a larger number of patients, they confirm micro-FTIR spectroscopy as a fast, robust technique. Moreover, early diagnosis of colon cancer can be reliably achieved relying only on the evaluation of collagen signals changes in stroma tissue.

3.4. PCA-LDA prediction model

Validity of IR procedures as a diagnostic technique is often evaluated employing prediction algorithms [39,40]. In this work the variable reduction with PCA followed by LDA was chosen as the prediction model. This algorithm was built using spectra of four colon tissues (two normal and two malignant) from two patients and considering the 1800–1000 cm^{-1} range. Four main histopathological classes: normal epithelium/stroma and malignant stroma/epithelium. The spectra used for training, internal validation and testing of the model were selected using Fuzzy C-means chemical maps of samples depicted in Fig. S1 in SI. The specificity and overall accuracy of the prediction model in the internal validation and test set were checked by the confusion matrix reported in Fig. S5 in SI. This evaluation showed a perfect correlation between the predicted and the corresponding histopathological true spectral classes. Considering this good result, external validation was carried out on the spectral data from remaining samples (Fig. 1 and Fig. S2 in SI). Classification performance of PCA-LDA based model on external samples are presented in the form of predicted chemicals maps and some significative results are shown in Fig. 5. For sample **6n** corresponding to the normal tissue, the predicted map (panel A, Fig. 5) was in agreement with the morphology of the microscope image. A correct classification of histopathological features was obtained, with only 0.62% of the total spectra incorrectly classified as malignant stroma. Samples **6m** and **7m** were both extracted from a malignant colon tissue, but they present a different tumor degree as confirmed by a pathologist. The undifferentiated sample **6m** was associated to a high degree, while the sample **7m,** which still showed well-differentiated glands from its adjacent stroma, was related to a moderate tumor degree by histopathological analysis. When the prediction model was applied to sample **6m**, the predicted map (panel B, Fig. 5) clearly represented the morphology of the microscope image, and an exact classification of histopathological features was achieved. Interestingly, different histopathological classes were assigned by the model to epithelium and stroma for sample **7m** (panel C, Fig. 5). Spectra of stroma region were assigned to the malignant class, while spectra corresponding to epithelium were classified as normal. This result highlighted that the chemical changes, occurring inside the stroma during the development of tumor, are detected earlier than those in epithelium structure. Predicted maps for some other samples are reported in Figs. S6–S7 of SI.

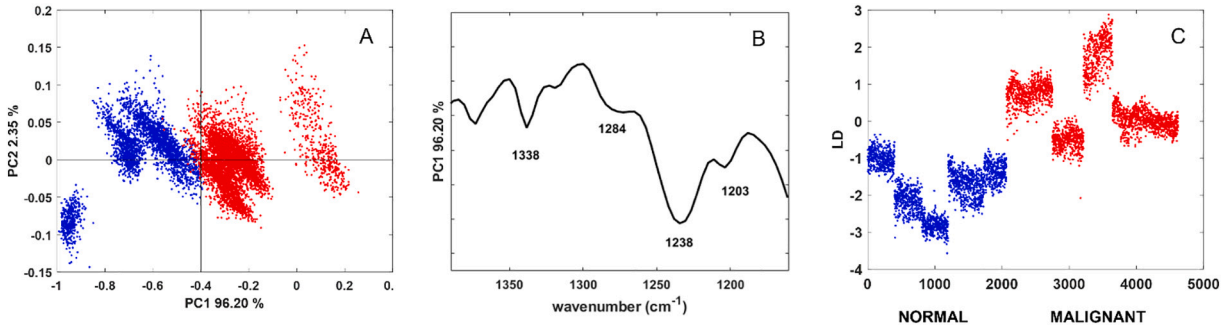

Fig. 3. (A) Score plot PC1 *vs* PC2 for normal (blue dot) and malignant (red dot) stroma in range 1150–1400 cm^{-1}. (B) Loadings associated with the score plot B. (C) LDA score plot of external samples, (blue dots represent normal Stroma, red dots represent malignant Stroma.

Fig. 4. Confusion matrix obtained by malignant (M) and normal (N) tissue from four different patients. Diagonal (green) and off-diagonal (red) cells correspond to correctly and incorrectly classified observations respectively. Number of spectra and percentage are reported at each cell. The grey column on the right gives the percentage of observations correctly (green font) and incorrectly (red font) classified for each predicted class. The bottom row gives the percentage of observations correctly and incorrectly classified for each true class. Overall accuracy is displayed at bottom right cell of each confusion matrix.

Despite the relatively restricted number of colonic tissues considered for model building, it can be observed that spectra corresponding to the stroma regions of malignant tissues were correctly classified as members of malignant stroma class by PCA-LDA for all analyzed samples. Moreover, considering the outcome from the diseased samples in which well-differentiated glands were still present, the classification results support the assumption that the stroma is more sensitive to biochemical modification inducted by tumoral growth than associated epithelium. Consequently, limiting biomarkers research only to histological features corresponding to stroma in colonic tissue could result in enhanced sensitivity.

4. Conclusions

In colon adenocarcinoma micro-FTIR spectroscopy turned out to be an appropriate technique to detect the chemical changes taking place in the extracellular matrix, the main constituent of stroma. To distinguish between stroma and epithelium in malignant as well as in normal tissue, Fuzzy C-means clustering was used, and spectra associated to the stroma histopathological structure were extracted. The main differences between the average spectra of malignant and normal stroma are related to signals assigned to collagens. The most significant variations were noted for the absorptions at 1203, 1238, 1284 cm^{-1} of amide III band and 1338 cm^{-1} assigned to the C—H_2 wagging vibration of proline side chains. These signals are sensitive to the interactions between the α-chains in the triple helices of collagens. The changes of these signals in cancer stroma are compatible with the formation of collagen fibers that are more aligned, compact and ordered than in normal tissue. These results are in accordance with studies reported in literature ascertaining a highly stiff and fibrous ECM in the cancer condition. PCA analysis applied to the spectra from malignant and normal stroma of all analyzed patients confirmed that the highest explained variance for PC1 correspond to spectral interval 1150–1400 cm^{-1}. PCA loadings further confirmed that the most discriminant wavenumbers are associated with collagen absorptions. The application of a PCA-LDA prediction model further established that the collagen chemical modifications in the ECM are a sensitive sign of malignancy using IR spectroscopy. Indeed, in the predicted maps of the diseased samples with moderate tumor degree, the

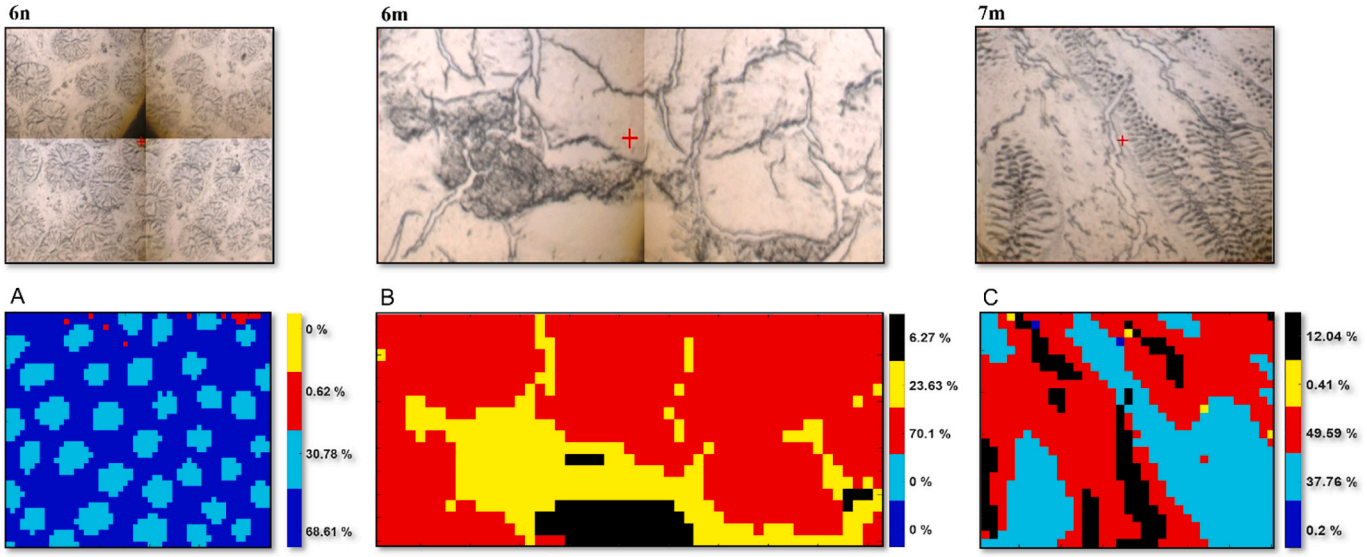

Fig. 5. PCA-LDA predicted maps for normal (6n) and malignant (6 m and 7 m) colon tissue. The label and colour at counter plot are referred to blue-NS (Normal Stroma), cyan-NE (Normal Epithelium), red-MS (Malignant Stroma) and yellow-ME (Malignant epithelium).

stroma was correctly classified as malignant while the epithelium was still assigned to normal condition. Considering the key role of the collagen modifications in the tumor progression, the absorptions 1203, 1238, 1284 and 1338 cm^{-1} were tested as spectral biomarkers. Although this study considers a limited number of patients, the preliminary findings have been promising and strongly indicate that the micro-FTIR spectroscopy could be used alongside the conventional histopathology for an early and reliable colon cancer diagnosis. Furthermore, the modifications of collagen highlighted by micro-FTIR can contribute to the knowledge of the role of the stroma in tumor progression and metastasis.

CRediT authorship contribution statement

Serena De Santis: Conceptualization, Investigation, Validation, Writing – original draft, Writing – review & editing. **Francesco Porcelli:** Methodology, Software, Formal analysis, Validation. **Giovanni Sotgiu:** Investigation, Writing – review & editing, Supervision. **Anna Crescenzi:** Conceptualization, Investigation. **Anita Ceccucci:** Formal analysis. **Martina Verri:** Investigation. **Marco Caricato:** Investigation. **Chiara Taffon:** Investigation. **Monica Orsini:** Conceptualization, Resources, Writing – original draft, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbadis.2021.166279.

References

[1] M. Giussani, T. Triulzi, G. Sozzi, E. Tagliabue, Tumor extracellular matrix remodeling: new perspectives as a circulating tool in the diagnosis and prognosis of solid tumors, Cells 8 (2019) 1–17, https://doi.org/10.3390/cells8020081.
[2] J. Winkler, A. Abisoye-Ogunniyan, K.J. Metcalf, Z. Werb, Concepts of extracellular matrix remodelling in tumour progression and metastasis, Nature Comm. 11 (2020) 1–19, https://doi.org/10.1038/s41467-020-18794-x.
[3] M.A. Karsdal, M.J. Nielsen, J.M. Sand, K. Henriksen, F. Genovese, A.-C. Bay-Jensen, V. Smith, J.I. Adamkewicz, C. Christiansen, D.J. Leeming, Extracellular matrix remodeling: the common denominator in connective tissue diseases, Assay Drug Dev Techn 11 (2013) 70–92, https://doi.org/10.1089/adt.2012.474.
[4] J.F. Hastings, J.N. Skhinas, D. Fey, D.R. Croucher, T.R. Cox, The extracellular matrix as a key regulator of intracellular signalling networks, Br. J. Pharmacol. 176 (2019) 82–92, https://doi.org/10.1111/bph.14195.
[5] V. Poltavets, M. Kochetkova, S.M. Pitson, M.S. Samuel, The role of the extracellular matrix and its molecular and cellular regulators in cancer cell plasticity, Front. Oncol. 8 (2018) 1–19, https://doi.org/10.3389/fonc.2018.00431.
[6] S. Crotti, M. Piccoli, F. Rizzolio, A. Giordano, D. Nitti, M. Agostini, Colorectal cancer: how surrounding microenvironment affects cancer cell behavior? J. Cell. Physiol. 232 (2017) 967–975, https://doi.org/10.1002/jcp.25658.
[7] S.Z. Despotović, Đ.N. Milićević, A.J. Krmpot, A.M. Pavlović, V.D. Živanović, Z. Krivokapić, V.B. Pavlović, S. Lević, G. Nikolić, M.D. Rabasović, Altered organization of collagen fibers in the uninvolved human colon mucosa 10 cm and 20 cm away from the malignant tumor, Sci. Rep. 10 (2020) 1–11, https://doi.org/10.1038/s41598-020-63368-y.
[8] M. Egeblad, M.G. Rasch, V.M. Weaver, Dynamic interplay between the collagen scaffold and tumor evolution, Curr. Opin. Cell Biol. 22 (2010) 697–706, https://doi.org/10.1016/j.ceb.2010.08.015.
[9] S. Xu, H. Xu, W. Wang, S. Li, H. Li, T. Li, W. Zhang, X. Yu, L. Liu, The role of collagen in cancer: from bench to bedside, J. Transl. Med. 17 (2019) 1–22, https://doi.org/10.1186/s12967-019-2058-1.
[10] D. Hulmes, Collagen diversity, synthesis and assembly, in: P. Fratzl (Ed.), Collagen, Springer, Boston, MA, 2008, pp. 15–47, https://doi.org/10.1007/978-0-387-73906-9_2.
[11] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 68 (2018) 394–424, https://doi.org/10.3322/caac.21492.
[12] E. Brauchle, J. Kasper, R. Daum, N. Schierbaum, C. Falch, A. Kirschniak, T. E. Schäffer, K. Schenke-Layland, Biomechanical and biomolecular characterization of extracellular matrix structures in human colon carcinomas, Matrix Biol. 68–69 (2018) 180–193, https://doi.org/10.1016/j.matbio.2018.03.016.
[13] T.T. Vellinga, S. den Uil, I.H.B. Rinkes, D. Marvin, B. Ponsioen, A. Alvarez-Varela, S. Fatrai, C. Scheele, D.A. Zwijnenburg, H. Snippert, L. Vermeulen, J.P. Medema, H. B. Stockmann, J. Koster, R.J.A. Fijneman, J. de Rooij, O. Kranenburg, Collagen-rich stroma in aggressive colon tumors induces mesenchymal gene expression and tumor cell invasion, Oncogene 35 (2016) 5263–5271, https://doi.org/10.1038/onc.2016.60.
[14] J.M. Moilanen, N. Kokkonen, S. Löffek, J.P. Väyrynen, E. Syväniemi, T. Hurskainen, M. Mäkinen, K. Klintrup, J. Mäkelä, R. Sormunen, L. Bruckner-Tuderman, H. Autio-Harmainen, K. Tasanen, Collagen XVII expression correlates with the invasion and metastasis of colorectal cancer, Hum. Pathol. 46 (2015) 434–442, https://doi.org/10.1016/j.humpath.2014.11.020.
[15] M.K. Bode, T.J. Karttunen, J. Mäkelä, L. Risteli, J. Risteli, Type I and III collagens in human colon cancer and diverticulosis, Scand. J. Gastroenterol. 35 (2000) 747–752, https://doi.org/10.1080/003655200750023435.
[16] O. Rouvière, C. Melodelima, A.H. Dinh, F. Bratan, G. Pagnoux, T. Sanzalone, S. Crouzet, M. Colombel, F. Mège-Lechevallier, R. Souchon, Stiffness of benign and malignant prostate tissue measured by shear-wave elastography: a preliminary study, Eur. Radiol. 27 (2017) 1858–1866, https://doi.org/10.1007/s00330-016-4534-9.
[17] Z. Movasaghi, S. Rehman, I. Rehman, Fourier transform infrared (FTIR) spectroscopy of biological tissues, Appl. Spectrosc. Rev. 43 (2008) 134–179, https://doi.org/10.1080/05704920701829043.

[18] S. De Santis, G. Sotgiu, A. Crescenzi, C. Taffon, A.C. Felici, M. Orsini, On the chemical composition of psammoma bodies microcalcifications in thyroid cancer tissues, J. Pharm. Biomed. Anal. 190 (2020), 113534, https://doi.org/10.1016/j.jpba.2020.113534.
[19] C. Battocchio, S. Concolato, S. De Santis, M. Fahlman, G. Iucci, M. Santi, G. Sotgiu, M. Orsini, Chitosan functionalization of titanium and Ti6Al4V alloy with chloroacetic acid as linker agent, Mater. Sci. Eng. C 99 (2019) 1133–1140, https://doi.org/10.1016/j.msec.2019.02.052.
[20] M.J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H.J. Butler, K.M. Dorling, P.R. Fielden, S.W. Fogarty, N.J. Fullwood, K.A. Heys, C. Hughes, P. Lasch, P.L. Martin-Hirsch, B. Obinaju, G.D. Sockalingun, J. Sulé-Suso, R.J. Strong, M.J. Walsh, B.R. Wood, P. Gardner, F.L. Martin, Using Fourier transform IR spectroscopy to analyze biological materials, Nat. Protoc. 9 (2014) 1771–1791, doi:https://doi.org/10.1038/nprot.2014. 110.
[21] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models, PLS-DA. Anal. Methods 5 (2013) 3790–3798, https://doi.org/10.1039/C3AY40582F.
[22] H. Martens, J.P. Nielsen, S.B. Engelsen, Light scattering and light absorbance separated by extended multiplicative signal correction, application to near-infrared transmission analysis of powder mixtures, Anal. Chem. 3 (2003) 394–404, https://doi.org/10.1021/ac020194w.
[23] N.K. Afseth, A. Kohler, Extended multiplicative signal correction in vibrational spectroscopy, a tutorial, Chemom. Intell. Lab. Syst. 117 (2012) 92–99, https://doi.org/10.1016/j.chemolab.2012.03.004.
[24] A. Zancla, S. De Santis, A. Scorza, S.A. Sciuto, G. Sotgiu, M. Orsini, A preliminary study on the importance of normalization methods in infrared micro-spectroscopy for biomedical applications, 24th IMEKO TC4 International Symposium and 22nd International Workshop on ADC and DAC Modelling and Testing (2020) 342–346.
[25] A. Zancla, S. De Santis, G. Sotgiu, C. Taffon, A. Crescenzi, M. Orsini, Micro-FTIR spectroscopy as robust tool for psammoma bodies detection in papillary thyroid carcinoma, Spectrochim. Acta A 229 (2020), 117984, https://doi.org/10.1016/j.saa.2019.117984.
[26] P. Lasch, W. Haensch, D. Naumann, M. Diem, Imaging of colorectal adenocarcinoma using FT-IR microspectroscopy and cluster analysis, Biochim. Biophys. Acta 2004 (1688) 176–186, https://doi.org/10.1016/j.bbadis.2003.12.006.
[27] C.L.M. Morais, M.C.D. Santos, K.M.G. Lima, F.L. Martin, Improving data splitting for classification applications in spectrochemical analyses employing a random-mutation Kennard-Stone algorithm approach, Bioinformatics 35 (2019) 5257–5263, https://doi.org/10.1093/bioinformatics/btz421.
[28] C.L.M. Morais, K.M.G. Lima, M. Singh, Tutorial: multivariate classification for vibrational spectroscopy in biological samples, Nat. Protoc. 15 (2020) 2143–2162.
[29] J. Nallala, O. Piot, M.-D. Diebold, C. Gobinet, O. Bouche, M. Manfait, G. D. Sockalingum, Infrared imaging as a cancer diagnostic tool: introducing a new concept of spectral barcodes for identifying molecular changes in colon tumors, Cytom. A 83A (2013) 294–300, https://doi.org/10.1002/cyto.a.22249.
[30] J. Nallala, M.-D. Diebold, C. Gobinet, O. Bouche´, G. D. Sockalingum, O. Piot, M. Manfait, Infrared spectral histopathology for cancer diagnosis: a novel approach for automated pattern recognition of colon adenocarcinoma, Analyst 139 (2014) 4005–4015. doi:https://doi.org/10.1039/c3an01022h.
[31] C.L. Song, S.G. Kazarian, Effect of controlled humidity and tissue hydration on colon cancer diagnostic via FTIR spectroscopic imaging, Anal. Chem. 92 (2020) 9691–9698, https://doi.org/10.1021/acs.analchem.0c01002.
[32] K. Belbachir, R. Noreen, G. Gouspillou, C. Petibois, Collagen types analysis and differentiation by FTIR spectroscopy, Anal. Bioanal. Chem. 395 (2009) 829–837, https://doi.org/10.1007/s00216-009-3019-y.
[33] A. Zwielly, S. Mordechai, I. Sinielnikov, A. Salman, E. Bogomolny, S. Argov, Advanced statistical techniques applied to comprehensive FTIR spectra on human colonic tissues, Med. Phys. 37 (2010) 1047–1055, https://doi.org/10.1118/1.3298013.
[34] M. Jackson, L.P. Choo, P.H. Watson, W.C. Halliday, H.H. Mantsch, Beware of connective tissue proteins: assignment and implications of collagen absorptions in infrared spectra of human tissues, Biochim. Biophys. Acta 1270 (1995) 1–6, https://doi.org/10.1016/0925-4439(94)00056-v.
[35] C. Stani, L. Vaccari, E. Mitri, G. Birarda, FTIR investigation of the secondary structure of type I collagen: new insight into the amide III band, Spectrochim. Acta A Mol. Biomol. Spectrosc. 229 (2020), 118006, https://doi.org/10.1016/j.saa.2019.118006.
[36] J. Bella, Collagen structure: new tricks from a very old dog, Biochem. J. 473 (2016) 1001–1025, https://doi.org/10.1042/BJ20151169.
[37] R. Cheheltani, J.M. Rosano, B. Wang, A.K. Sabri, N. Pleshko, M.F. Kiania, Fourier transform infrared spectroscopic imaging of cardiac tissue to detect collagen deposition after myocardial infarction, J. Biomed. Opt. 17 (2012), https://doi.org/10.1117/1.JBO.17.5.056014, 056014.
[38] P.A. West, P.A. Torzilli C. Chen, P. Lin, N.P. Camacho, Fourier transform infrared imaging spectroscopy analysis of collagenase-induced cartilage degradation, J. Biomed. Opt. 10 (2005), 014015, https://doi.org/10.1117/1.1854131.
[39] J. Nallala, C. Jeynes, S. Saunders, N. Smart, G. Lloyd, L. Riley, D. Salmon, N. Stone, Characterization of colorectal mucus using infrared spectroscopy: a potential target for bowel cancer screening and diagnosis laboratory investigation 100 (2020) 1102–1110, https://doi.org/10.1038/s41374-020-0418-3.
[40] A. Travo, O. Piot, R. Wolthuis, C. Gobinet, M. Manfait, J. Bara, M.-E. Forgue-Lafitte, P. Jeannesson, IR spectral imaging of secreted mucus: a promising new tool for the histopathological recognition of human colonic adenocarcinomas, Histopathology 56 (2010) 921–931, https://doi.org/10.1111/j.1365-2559.2010.03563.x.