Identification of circulating biomarkers of Crohn's disease and spondyloarthritis using FTIR spectroscopy

Pierre PRADA[1], Benjamin BRUNEL[1,2], David MOULIN[3], Lise ROUILLON[1], Patrick NETTER[3], Damien LOEUILLE[3], Florian SLIMANO[1], Olivier BOUCHE[1], Laurent PEYRIN-BIROULET[4], Jean-Yves JOUZEAU[3], Olivier PIOT[1,5]*

---

[1] Université de Reims Champagne-Ardenne, EA7506-BioSpectroscopie Translationnelle (BioSpecT), Reims, France

[2] Université de Bourgogne Franche-Comté, FEMTO-ST Institute, CNRS UMR-6174, Besançon, France

[3] Ingénierie Moléculaire et Ingénierie Articulaire (IMoPA), UMR-7365 CNRS, Faculté de Médecine, Université de Lorraine et Hôpital Universitaire de Nancy, France

[4] Département de Gastroentérologie, Hôpital Universitaire de Nancy-Brabois, Vandœuvre-lès-Nancy, France

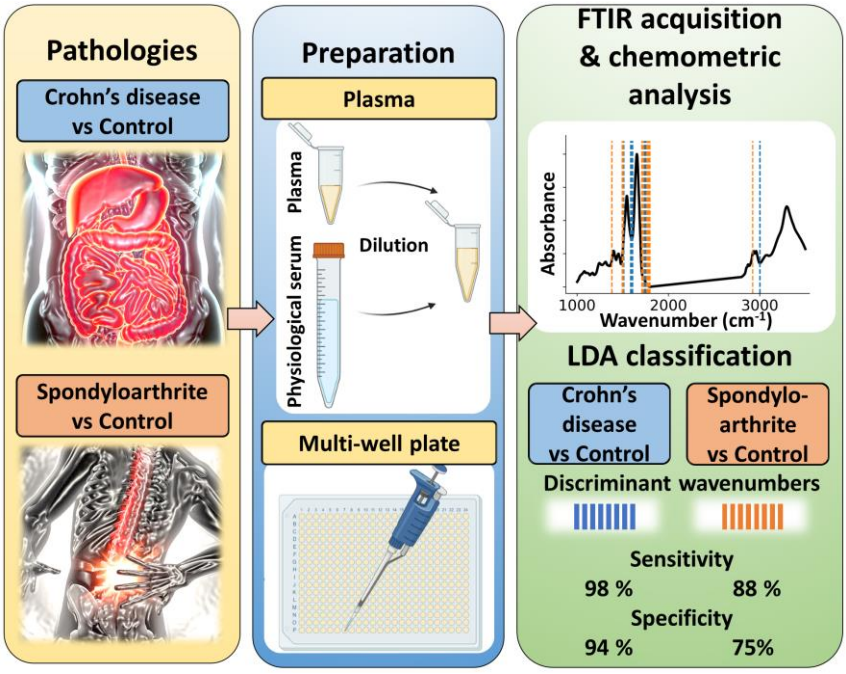
[5] Université de Reims Champagne-Ardenne, Plateforme d'Imagerie Cellulaire ou Tissulaire (PICT), Reims, France

*Correspondence
Olivier PIOT
Université de Reims Champagne-Ardenne, EA7506-BioSpectroscopie Translationnelle (BioSpecT), UFR de Pharmacie, 51 rue Cognacq-Jay, 51097 Reims, France

Email: olivier.piot@univ-reims.fr

Abstract

Crohn's disease (CD) and spondyloarthritis (SpA) are two inflammatory diseases sharing many common features (genetic polymorphism, armamentarium). Both diseases lack diagnostic markers of certainty. While the diagnosis of CD is made by a combination of clinical, and biological criteria, the diagnosis of SpA may take several years to be confirmed. Based on the hypothesis that CD and SpA alter the biochemical profile of plasma, the objective of this study was to evaluate the analytical capability of Fourier transform infrared spectroscopy (FTIR) in identifying spectral biomarkers. Plasma from 104 patients was analyzed. After data processing of the spectra by EMSC and LDA, we demonstrated that it was possible to distinguish CD and SpA from controls with an accuracy of 97% and 85% respectively. Spectral differences were mainly associated with proteins and lipids. This study showed that FTIR analysis is efficient to identify plasma biosignatures specific to CD or SpA.



Keywords: Crohn's disease, Spondyloarthritis, Plasma, Spectroscopy, Fourier Transform Infrared, Chemometrics, Biomarkers, Diagnostic.

1 INTRODUCTION

Crohn's disease (CD) is a chronic inflammatory disease of the gastrointestinal tract, part of the inflammatory bowel diseases (IBD). It results from a complex interaction between environmental factors, genetic susceptibility and altered gut microbiota, leading to a dysregulated immune response.[1,2] The most common symptoms are abdominal pain, chronic

diarrhea, weight-loss and fatigue. There is no sex distribution in adult Crohn's disease. Crohn's disease frequency has been steadily increasing in most parts of the world. Its incidence and prevalence are higher in developed countries, and in urban areas. It is hypothesized that the environment and early life exposures play a role in the risk of developing Crohn's disease. The therapeutic options for CD are emerging, and therefore the need to develop early diagnostic and prognostic biomarkers become increasingly important for personalized medicine strategies.[2]

Spondyloarthritis (SpA) groups together several chronic inflammatory rheumatic diseases, with common clinical symptoms and genetic predisposition. There are two main subtypes according to the location of the diseased joints: peripheral spondyloarthritis (peripheral SpA) and axial spondyloarthritis (axial SpA). Peripheral SpA includes psoriatic arthritis[3,4], reactive arthritis and arthritis associated with inflammatory bowel diseases (IBD). The severe form of axial SpA is ankylosing spondylitis (AS) which affects 0.1-0.5% of the population and is characterized by an attack on the axial skeleton (spine and sacroiliac joints) that can lead to ankylosis. Pathological indicators are inflammatory back pain, radiographic sacroiliitis, excessive spinal bone formation, and a high prevalence of HLA-B27. Peripheral and axial SpA can also occur together. SpA is associated with a significant physical and social burden and can impair everyday life. Therapeutic objectives are relieving symptoms, improving function, maintaining the ability to work, decreasing disease complications, and preventing skeletal damage as much as possible. There is no real cure, hence the need for an early diagnosis of certainty.[5,6]

Crohn's disease and spondyloarthritis are intimately related. On one hand, peripheral joint inflammation affects 10-20% of patients with Crohn's disease and 5-10% of patients with ulcerative colitis. In the more general group of IBD, spondyloarthritis joint symptoms occurs in up to 13% of patients.[7] The most common is peripheral arthritis, followed by sacroiliitis and finally ankylosing spondylitis.[7,8] On the other hand, peripheral arthralgias can be correlated with IBD activity and are more rapidly controlled when IBD remission is achieved. However, axial involvement is generally independent of IBD activity. The diagnosis of spondyloarthritis in the course of inflammatory bowel disease is most often based on a range of arguments and a diagnosis of probability rather than of certainty.

In both pathologies, there is a lack of relevant markers of certainty as their identification is not easy, especially for the diagnosis of SpA which can take up to several years.[9] There is thus a need for new biomarkers, allowing a more certain and/or earlier disease diagnosis. The latter could allow treatment initiation at earlier stages and/or to personalize therapy. As SpA and CD are often but not systematically associated, it is relevant to study them together, to search for some common biological mechanisms and understand how they can be different.

In a context of improving techniques with a cross-sectional approach to medicine, the identification of diagnostic markers by mid-infrared absorption spectroscopy appears to be a promising alternative approach. The analysis of the non-destructive interaction between matter and infrared radiation generates a "spectral fingerprint". This fingerprint represents the biomolecular composition of a studied sample such as a tissue or a biofluid.[10,11] The statistical processing of the spectral data generated from pathological or healthy samples, allows extraction of spectral differences representative for biochemical differences related to the pathology in question.

Vibrational spectroscopy is a simple, fast and cost-effective technique that can be applied to the analysis of different biofluids. By identifying specific spectral markers, it is possible to have an efficient diagnosis.[12] Thus, different cancers can be detected by plasma or serum analysis [13], including hepatocellular carcinoma [14], brain [15,16], breast [17], ovarian [18], prostate [19], pleural[20], and oral [21] cancer. In addition, sputum [22] or saliva [23] can be used for lung diagnosis. Other pathologies can be diagnosed, such as amyotrophic lateral sclerosis in tears [24], Alzheimer's disease and relapsing-remitting multiple sclerosis in cerebral spinal fluid [25,26], pleural mesothelioma in pleural fluid[27] or septic arthritis in synovial fluid.[28]

In the specific case of CD and SpA, previous work was done by Y. Wu *et al.* to diagnose CD on urine samples.[29] But the contribution of infrared spectroscopy has never been tested before in the diagnosis of these two pathologies using plasma. In our study, we tried to set up a diagnosis of CD and SpA using plasma. Plasma offers some advantages over urine including its high protein content and lack of diuresis-related changes. Plasma sampling is frequent in the framework of disease diagnosis, patient follow-up or for a health check-up. For example, A. D. Morris *et al.*, demonstrated that plasma provides very good results compared to alternative biofluids for the diagnosis of ANCA.[30]

Based on the hypothesis that SpA or IBD (Crohn's) modifies the biochemical profile of plasma, the objective of this pilot study was to evaluate the analytical capability of Fourier transform infrared spectroscopy (FTIR) to identify discriminating spectral signatures with diagnostic relevance to these two chronic inflammatory diseases.

2 MATERIAL AND METHODS

2.1 Study design

The study FLORACROHN was approved by the French Ethics Committee for the protection of persons consenting to biomedical research, by the French National Agency for the Safety of Medicines and Health Products (2016-A01720-51) and registered in ClinicalTrials.gov (NCT03072836).

2.2 Patient criteria selection

Inclusion criteria

This is a 3-year consecutive cohort, in which patients were included during a medical visit as part of their longitudinal follow-up. The following criteria were used: i) age over 18 years; ii) IBD diagnosed since at least 3 months based on clinical, biological, radiological, endoscopic, and/or histological findings; iii) spondyloarthritis diagnosed on clinical (inflammatory spinal pain, asymmetric, oligoarticular peripheral arthritis predominating in the lower extremities, sacroiliitis, dactylitis, possible inflammatory eye disease), radiological (MRI [early stage] or X-ray [advanced stage] sacroillitis, MRI peripheral arthritis) and biological (HLA B27 positivity, elevated C-reactive protein) criteria according to the European League Against Rheumatism (EULAR) and the Assessment of Spondyloarthritis international Society (ASAS); iv) ability of the patient to give express and informed written consent. Healthy controls were volunteers recruited after a negative colorectal cancer screening.

Exclusion criteria

Recent (<3 months) use of any antibiotic therapy, current extreme diet (e.g., parenteral nutrition or macrobiotic diet), known history of malignancy, current consumption of probiotics, any gastrointestinal tract surgery leaving permanent residua (e.g., gastrectomy, bariatric surgery, or colectomy), recent colonoscopy (<2 months).

2.3 Data collection

As IBD and SpA are chronic by nature, ongoing treatments and disease activity scores (HBI [Harvey-Bradshaw Index] for IBD and BASDAI [Bath Ankylosing Spondylitis Disease Activity Index] for arthritis) as well , as well as objective signs of inflammation (CRP, last endoscopy and/or abdominal MRI), were recorded. Moreover, demographic and biometric data (age, sex, weight, height, body mass index (BMI)), as well as behavioral habits (smoking and alcohol consumption) were collected.

2.4 Plasma collection

Blood samples were collected from patients undergoing routine biochemical monitoring in the rheumatology or gastroenterology departments of Nancy university hospital. Blood samples were collected in heparinized tubes and centrifuged at 2000 g for 15 minutes for plasma collection. Plasma samples were stored frozen at -80°C until analysis.

2.5 FTIR spectral acquisition

For FTIR analysis, all plasma samples were diluted three times with 0.9% NaCl solution (Miniversol, Dutscher, France) and vortexed. For each sample, 15 drops of 5 µL were deposited separately on the wells of a 384-well IR-transparent silicon plate (Bruker Optics, GmbH, Ettlingen, Germany) and dried by vacuum drying for 2h. The dried drops were examined and found to be homogeneous visually.

After drying, the plate was loaded into the HTS-XT high-throughput extension module coupled to an FTIR spectrometer (Tensor 27 FTIR spectrometer, Bruker Optics). All deposits were analyzed and acquisitions were performed by mixing randomly samples from the three groups within a single multi-well plate, in order to avoid possible bias in data collection. FTIR measurements were performed in transmission mode at a spectral resolution of 4 cm^{-1} using 32 scans, over the 400-4000 cm^{-1} range. The background absorbance of the blank plate was measured under the same conditions before each batch analysis and automatically subtracted from each spectrum (OPUS v6.5 software, Bruker Optics GmbH, Ettlingen, Germany).[31,32]

2.6 Spectral pre-processing

Pre-processing of the spectra was performed using in-house algorithms written in Python programming language. First, wavenumbers below 1000 cm^{-1}, between 1800 cm^{-1} and 2800 cm^{-1}, and above 3500 cm^{-1} were cut-off as they correspond to weak and non-informative signal ranges. Then, normalization and baseline correction were performed using the EMSC

method (Extended Multiplicative Signal Correction), with a polynomial function of the 2nd order and the average spectrum as the target. Finally, a second EMSC was calculated (4th order polynomial function, average spectrum as the target), and a threshold on the error of the model was used for outliers' elimination, as a quality control process. Overall, less than 6% of all spectra were eliminated and no patients were eliminated from quality control.

2.7 Chemometrics analysis

A linear discriminant analysis (LDA) classifier was trained on the pre-processed spectra, to predict if a spectrum derived from a control patient (CP), a patient with CD or with SpA. LDA was performed using the class "LinearDiscriminantAnalysis" from the Python library Scikit-learn, with the default parameters. The discriminative potential was evaluated in terms of accuracy, sensitivity, specificity, F1 score, positive predictive value (PPV), and negative predictive value (NPV) in a 10-fold cross-validation procedure. Therefore, the patients were randomly separated into 10 groups, and the LDA classifier was trained on 9 of them and tested on the last one. By rotating the training and test groups, a prediction for the patients in each group as a test group was obtained. The confusion matrix in Figure 1 shows the summary of the predictions made on the test groups. For each patient, up to 15 spectra were included in the dataset. The class prediction was made for each spectrum, and the predicted class for a patient was chosen as the majority class among its spectra classes.

The LDA model classifies the spectra into the control class or the diseased class according to a calculated score. This score is the dot product between the intensities of the spectrum to be classified and the coefficients of the model. Thus, by studying these coefficients, one can know which wavenumbers have the most weight in the classification. However, the coefficients must be multiplied by the standard deviations of the intensity distributions of each wavenumber in order to obtain an unbiased result. The result is referred as "normalized" coefficients hereafter.

The normalized coefficients were averaged over the ten LDA models resulting from the 10-fold cross validation. The coefficients vector was then smoothed using Savitzky-Golay algorithm (window of 3) to reduce oscillations coming from correlation between intensities of adjacent wavenumbers. Normalized coefficients with the highest intensity (above 2.5 standard deviation) were extracted for vibrational attribution. Indeed, they correspond to wavenumbers which are the most important in the LDA classification. To ensure those wavenumbers were discriminating, a Mann-Whitney U-test (MWU) was performed for each of them. The test

compared preprocessed spectra intensities of patients with and without pathology. The resulting p-values were adjusted by the Bonferroni correction, considering approximately 40 wavenumbers. Wavenumbers with a p-values bellow 0.01 were considered discriminating.

3 RESULTS

In this study, 104 patients were included: 45 with CD, 43 with SpA and 16 without CD or SpA (Control Patients - CP). Of the 104 patients included, 47 were women and 57 were men. The median age was 61 years for control patients, 56 years for patients with spondyloarthritis and 38 years for patients with Crohn's disease. The demography of patient groups is listed in **Table 1**.

TABLE 1: Clinical and demographic data of CD, SpA and CP patients

Population and clinical data	**CD**	**SpA**	**CP**	**Total**
Number of patients	45	43	16	104
Mean age in years (range)	42 ±12 (26 - 67)	55 ±13 (34 - 78)	62 ±13 (39 - 87)	53
Median age (years)	38	56	61	49
Sex				
Female	27	17	3	47
Male	18	26	13	57
Sex-ratio (F/M)	1,5	0,7	0,2	0,8
Smoking status				
Non-smokers	12	11	9	32
Former-smokers	18	17	6	41
Smokers	15	14	1	30
Alcoholic drinkers	5	8	6	19

As detailed in the material and methods section, spectra were pre-processed and used to construct 2 LDA classifiers. Results are represented as confusion matrices (**Figure 1**), showing whether diseased patients are classified as diseased (true positive) or healthy (false negative), and whether control patients are classified as diseased (false positive) or healthy

(true negative). For each model, the classification performances were deduced from the confusion matrix (**Table 2**).

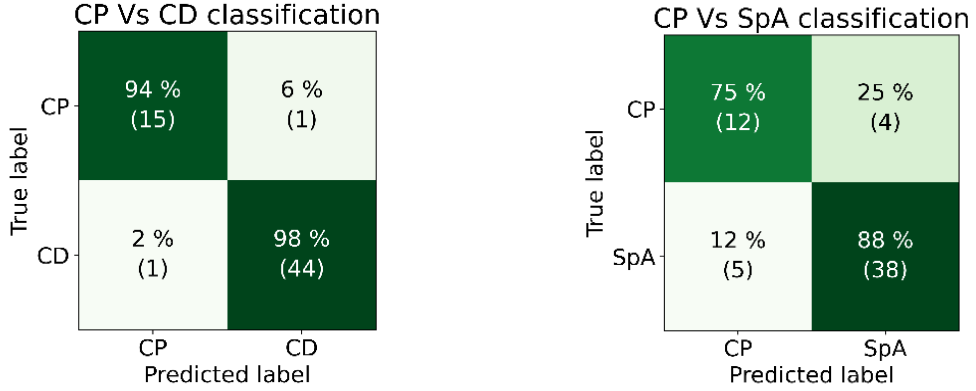


FIGURE 1: Confusion matrices of LDA classification for the diagnosis of Crohn's disease (left) and spondyloarthritis (right). Percentages indicate the fraction of patients classified in the right class (diagonal) or in the wrong class (outside the diagonal).

TABLE 2: Accuracy, sensitivity, specificity, F1 score, PPV and NPV for CD/CP and SpA/CP classifications.

CD /CP	SpA /CP
Accuracy = 97%	Accuracy = 85%
Sensitivity = 98%	Sensitivity = 88%
Specificity = 94%	Specificity = 75%
F1 score = 98%	F1 score = 89%
PPV = 94%	PPV = 71%
NPV = 98%	NPV = 90%

The normalized coefficients of the LDA model are shown in **Figure 2** (bottom rectangles). Wavenumbers that were the most important (greater than 2.5 standard deviation of the coefficients) in the classification were identified with color dashed lines displayed on the median spectra of groups (top rectangles) and the difference spectra (middle rectangles). To further select the discriminant wavenumbers, a MWU test, with a Bonferroni correction, was performed. Only wavenumbers with a p-value below 0.01 were selected. Red lines indicate a positive correlation between absorption intensity and disease probability, while green lines indicate a negative correlation.

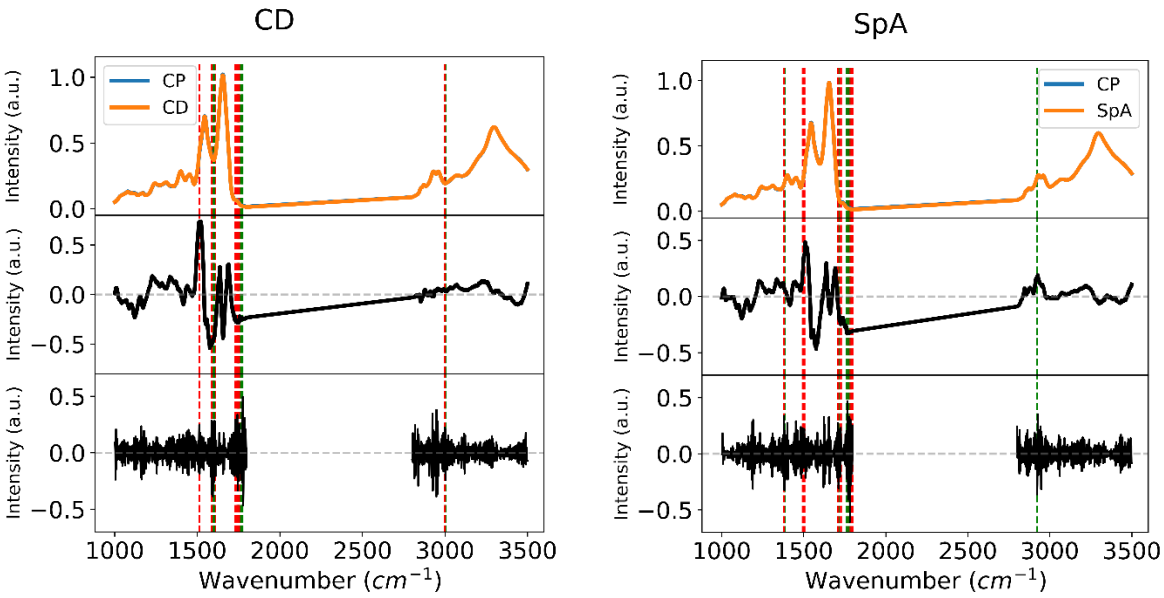


FIGURE 2: Identification of discriminating wavenumbers for CD vs CP (left) and SpA vs CP (right). Median spectra (top) of CD or SpA vs CP present subtle differences that are better seen in the magnified (x60) difference spectra (middle). From the LDA coefficients (bottom), wavenumbers with the greatest impact on the model were extracted and represented by color dashed lines.. Red lines correspond to positive correlations with the disease while green ones correspond to negative correlations.

Selected discriminant wavenumbers formed groups of few contiguous wavenumbers, which is consistent with the spectral resolution being smaller than the infrared peaks width. In some groups, the correlation with the disease was both positive and negative. It means LDA was sensitive to the shape modifications of the infrared peaks (e.g. shift) rather than its intensity. Significantly different absorption levels were observed in various windows of the spectrum, associated with proteins with amide II band (around 1500 cm^{-1}), lipid esters and fatty acids (C=O carbonyl stretching; 1700-1800 cm^{-1}) and lipids and cholesterol (symmetric and asymmetric stretching vibrations of CH_2 and CH_3 groups 2800-3000 cm^{-1}). Discriminant wavenumbers are summarized in **Table 3** for both SpA and CD groups. Main common wavenumbers were found in the regions of amide II band and lipids or fatty acids.

TABLE 3: **Summary of principal discriminant IR wavenumbers for CD vs CP and SpA vs CP (Figure 2).** For each peak, the assignment and the label are reported. Among adjacent discriminating wavenumbers, the wavenumber with the most significant p-value is chosen.

Peak (cm^{-1})	Group		Band assignment[33]	Label	Ref.
	CP-CD	CP-SpA			
1380-1385		x	Methyl (CH_3) symmetrical C-H bending	Phospholipids or ARN	[34]
1500-1510	x	x	Amide II band	Protein	[35]
1560-1590	x		C=N stretching of ring base	Nucleic Base	[34,36]
1700-1800	x	x	C=O stretching band mode of the fatty ester	Lipids	[37]
2900-2920		x	Stretching vibrations of CH_2 and CH_3 of cholesterol, phospholipids		[34,38]
3000	x		CH stretching vibrations of olefins bands or unsaturated fatty acids		[39]

4 DISCUSSION

Crohn's disease and spondyloarthritis are chronic inflammatory diseases whose diagnosis is not always obvious and whose physiopathology is still largely unknown.[40] The detection of these pathologies *via* vibrational spectroscopy analysis of plasma samples appears possible since promising results were obtained in this pilot study by using LDA based-predictive models with accuracy rates of 97% and 85% for CD and SpA respectively. This is the first study to provide such an accurate diagnosis of these two conditions using FTIR.

In vibrational spectroscopy, chemical bonds of molecular components contained in a sample are probed in a label-free manner and are represented by a spectrum. This spectrum allows to discriminate a disease from a healthy control by using several discriminating wavenumbers that can be considered as numerical spectroscopic markers specific of the disease.[11,32,41] The identified wavenumbers can be associated with chemical entities involved in the pathology such as inflammatory mediators or genetic factors.[41] They can be visualized as a digital fingerprint on the spectrum or as a spectral barcode.[42]

In the present study, both common and specific discriminant wavenumbers were found in plasma (Table 3). Common wavenumbers were not unexpected since the diseases share some common features. First, they are both chronic inflammatory diseases which are known to be

associated with circulating protein changes (1500-1510 cm^{-1}) that could be related to increase in C-reactive protein (CRP) or IL-6 plasma levels.[43] Second, gut microbiota dysbiosis is thought to play a pathogenic role in both diseases [44–46]. Common changes in gut bacterial species have been reported [47] which can, in turn, affect plasma metabolites levels. Indeed, a variation of fatty acids profile has been reported in serum in the course of both diseases, as a marker of disease activity[48] and with relevance for diagnostic purpose.[49] Such changes are consistent with the common profile that we observed in lipid bands (1700-1800 cm^{-1}). However, gut dysbiosis is generally more severe in CD than in SpA on the basis of alpha and beta bacterial diversity, meaning that disease-specific modifications of the fatty acids profile can also occur. The lipids-associated wavenumbers (1380-1385, 2900-2920, 3000 cm^{-1}) found in only one of the two inflammatory diseases could be possibly related to changes in short chain fatty acids levels for CD. Finally, a discriminant band associated with nucleic bases (1560-1590 cm^{-1}) and specific to CD vs CP classification, could be related to changes in miRNA profile in plasma, as described by Saccon et al.[50] The disease-specific signature was more marked for CD than for SpA, as illustrated by the higher F1 score (**Table 2**).

In our study, linear discriminant analysis (LDA) was used to process infrared data after EMSC pre-treatment. Various alternative methods of classification such as support vector machines (SVM), partial least squares discriminant analysis (PLS-DA) or again convolutional neural networks (CNN) [21,51], are available and could be compared to LDA; but this idea is out of the scope of the present work. LDA is one of the most widely used supervised methods for high dimensional spectral data. It aims at maximizing the variance between classes through a linear discriminant function and has the advantage of identifying the most informative vibrations in a robust way due to the ease of its optimization.[52] We used a 10-fold cross-validation procedure at the patient level by using 15 spectra for each sample to ensure a certain robustness of the prediction model.

Several studies have demonstrated the diagnostic potential of infrared analysis from biofluid samples, mainly in oncology as mentioned in the introduction of the manuscript. Serum and plasma samples are the most widely used biofluids and they have proven to contain informative molecular components that can be used for diagnostic purposes[16,53]. Recently, Morasso *et al.* showed that Raman microspectroscopy could be used for the diagnosis of

CD.[54] Similarly to our work, they analyzed dried plasma from healthy patients and patients with CD, and obtained a LDA model with an accuracy of 84% and a F1-score of 87%. Here, the use of FTIR combined to a high-throughput acquisition module resulted in higher accuracy and F1-score. They found spectral differences corresponding to carotenoids, aromatic amino acids, lipids, and proteins. The last two match our results, although the wavenumbers were different, which is to be expected since the vibrations to which Raman scattering and infrared absorption are active obey to different photonic principles.
In our investigations on CD and SpA, complementary analysis on serum samples could be carried out and compared to the data obtained from plasma samples. Such a comparison could permit to investigate if fibrinogen, not present in serum and found in increased amount during chronic inflammation, is contributive to the discriminating power of FITR spectral signatures.[55,56]

Despite the advantages of biofluids as easy to collect samples and carrier of diagnostic molecular entities, these samples are highly subject to experimental variability. This is likely to induce non-informative spectral differences in the vibrational analyses and make difficult the identification of discriminant wavenumbers. Consequently, it is paramount to follow standardized protocols from samples collection to spectral acquisition. Reference studies, some of which were carried out in our laboratory, served as a basis for the implementation of our study protocol[57,58]. Thus, for example, the plasma samples were diluted by saline rather than by distilled water.

Before implementation in routine clinic, the approach needs to be validated on a larger cohort in a multicentric way. For this pilot study, samples originated from patients seen in consultation at the Nancy University Hospital during a gastroenterology or rheumatology assessment. One criticism we could make is that this sample collection resulted in an imperfect matching between groups, especially between CD and CP groups that present the highest accuracy rate of classification (90%). Indeed, there were differences in the mean age and sex ratio between the different groups. To check whether the classification models were influenced by the age or sex, we looked at the predictions obtained by these models for subgroups of patients (CD or SpA) with a median age or sex-ratio equivalent to the control group. Results show that the classification performances from these subgroups were as good as those obtained with the entire cohorts (Supporting Information **Table S1** and **Table S2**). This comparison suggests that age or sex has no or insignificant influence on the spectral

classifications to distinguish between control patients and CD or SpA patients. In the literature, Magalhães *et al.* showed, by considering patients between 50 and 95 year-old, that there is a correlation between age and plasma FTIR spectra *via* protein conformation, however on one hand the age range is quite different from our cohort, and in other hand the wavelengths associated with the protein conformation are not the most discriminating bands identified in our study.[59] In addition, Morasso *et al.* in their analysis of spectral differences between plasma of CD and healthy individuals, investigated the impact of age and sex on the discriminative power of Raman spectroscopy.[54] They concluded that even if, for some discriminant peaks, differences in terms of sex or age can be observed, the difference associated with the Crohn disease status remained statistically significant. These data support that the imperfect age and sex-matching between our groups is not a major bias limiting the discriminative potential of our method, even if data warrant to be confirmed at a larger scale including younger control volunteers (not very compatible with patient inclusion during examination for colorectal cancer screening).

In addition, the number of patients in the CP group (n= 16) is relatively small compared to the groups of sick patients (SPA n= 43 and CD n= 46). However, it is assumed that this does not affect our results, as in other studies in the diagnosis of Alzheimer's disease on blood plasma or on the search for biomarkers of osteoarthritis on cartilage which also had unbalanced groups.[60,61] Moreover, sensitivity, specificity, F1 score, PPV and NPV are relevant when dealing with unbalanced groups (**Table 2**).

Overall, data of this pilot study support the hypothesis that plasma from patients with SpA or CD present infrared spectral characteristics that can distinguish them from healthy patients. The development of a diagnostic tool based on FTIR spectroscopy would overcome the lack of means available for the diagnosis of these pathologies. Such an approach, based on liquid biopsy, could contribute to an early personalized management of patients allowing to limit the progression of the disease towards a severe handicap or to follow the response to the treatment.

The next step to validate this approach for possible clinical applications would be to analyze more precisely the impact of possible confounding factors such as the presence of comorbidities or drug treatments, especially immunosuppressants, which are often prescribed in these inflammatory pathologies. Additionally, future research could focus on finding spectral markers correlated with disease stage based on severity and flare or remission period.

CONFLICT OF INTEREST

The authors declare no financial or commercial conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENTS

This work has been supported by the 'Agence Regionale de Santé (ARS)' for funding the doctoral position of Pierre Prada and by the 'Fondation ARC pour la recherche sur le cancer', by funding the post-doctoral position of Dr. Benjamin Brunel. Also, we thank the European Union and the Grand Est region for financing equipment and experimental costs in the frame of the TARGET project. Europe is engaged in the Grand-Est with FEDER.

The authors thank Ms Johanna Protin for her technical assistance.

REFERENCES

[1] W. Elhenawy, S. Hordienko, S. Gould, A. M. Oberc, C. N. Tsai, T. P. Hubbard, M. K. Waldor, B. K. Coombes, *Nat Commun* **2021**, *12*, 2032.
[2] J. Torres, S. Mehandru, J.-F. Colombel, L. Peyrin-Biroulet, *The Lancet* **2017**, *389*, 1741–1755.
[3] M. Dougados, D. Baeten, *The Lancet* **2011**, *377*, 2127–2137.
[4] M. Rudwaleit, D. van der Heijde, R. Landewé, N. Akkoc, J. Brandt, C. T. Chou, M. Dougados, F. Huang, J. Gu, Y. Kirazli, F. V. den Bosch, I. Olivieri, E. Roussou, S. Scarpato, I. J. Sørensen, R. Valle-Oñate, U. Weber, J. Wei, J. Sieper, *Annals of the Rheumatic Diseases* **2011**, *70*, 25–31.
[5] M. M. Ward, A. Deodhar, L. S. Gensler, M. Dubreuil, D. Yu, M. A. Khan, N. Haroon, D. Borenstein, R. Wang, A. Biehl, M. A. Fang, G. Louie, V. Majithia, B. Ng, R. Bigham, M. Pianin, A. A. Shah, N. Sullivan, M. Turgunbaev, J. Oristaglio, A. Turner, W. P. Maksymowych, L. Caplan, *Arthritis & Rheumatology* **2019**, *71*, 1599–1613.
[6] J. D. Taurog, A. Chhabra, R. A. Colbert, *N Engl J Med* **2016**, *374*, 2563–2574.
[7] M. C. Karreman, J. J. Luime, J. M. W. Hazes, A. E. A. M. Weel, *Journal of Crohn's and Colitis* **2017**, *11*, 631–642.
[8] G. E. Fragoulis, C. Liava, D. Daoussis, E. Akriviadis, A. Garyfallos, T. Dimitroulas, *World J Gastroenterol* **2019**, *25*, 2162–2176.
[9] A. Kumar, D. Lukin, R. Battat, M. Schwartzman, L. A. Mandl, E. Scherl, R. S. Longman, *J Gastroenterol* **2020**, *55*, 667–678.
[10] L. Lechowicz, M. Chrapek, J. Gaweda, M. Urbaniak, I. Konieczna, *Mol Biol Rep* **2016**, *43*, 1321–1326.

[11] M. Paraskevaidi, C. L. M. Morais, K. M. G. Lima, J. S. Snowden, J. A. Saxon, A. M. T. Richardson, M. Jones, D. M. A. Mann, D. Allsop, P. L. Martin-Hirsch, F. L. Martin, *Proc Natl Acad Sci U S A* **2017**, *114*, E7929–E7938.
[12] A. L. Mitchell, K. B. Gajjar, G. Theophilou, F. L. Martin, P. L. Martin-Hirsch, *Journal of Biophotonics* **2014**, *7*, 153–165.
[13] J. Ollesch, S. L. Drees, H. M. Heise, T. Behrens, T. Brüning, K. Gerwert, *Analyst* **2013**, *138*, 4092–4102.
[14] I. Taleb, G. Thiéfin, C. Gobinet, V. Untereiner, B. Bernard-Chabert, A. Heurgué, C. Truntzer, P. Hillon, M. Manfait, P. Ducoroy, G. D. Sockalingum, *Analyst* **2013**, *138*, 4006–4014.
[15] J. R. Hands, G. Clemens, R. Stables, K. Ashton, A. Brodbelt, C. Davis, T. P. Dawson, M. D. Jenkinson, R. W. Lea, C. Walker, M. J. Baker, *J Neurooncol* **2016**, *127*, 463–472.
[16] J. M. Cameron, P. M. Brennan, G. Antoniou, H. J. Butler, L. Christie, J. J. A. Conn, T. Curran, E. Gray, M. G. Hegarty, M. D. Jenkinson, D. Orringer, D. S. Palmer, A. Sala, B. R. Smith, M. J. Baker, *Neurooncol Adv* **2022**, *4*, vdac024.
[17] V. E. Sitnikova, M. A. Kotkova, T. N. Nosenko, T. N. Kotkova, D. M. Martynova, M. V. Uspenskaya, *Talanta* **2020**, *214*, 120857.
[18] K. Gajjar, J. Trevisan, G. Owens, P. J. Keating, N. J. Wood, H. F. Stringfellow, P. L. Martin-Hirsch, F. L. Martin, *Analyst* **2013**, *138*, 3917–3926.
[19] D. K. R. Medipally, T. N. Q. Nguyen, J. Bryant, V. Untereiner, G. D. Sockalingum, D. Cullen, E. Noone, S. Bradshaw, M. Finn, M. Dunne, A. M. Shannon, J. Armstrong, F. M. Lyng, A. D. Meade, *Cancers* **2019**, *11*, 925.
[20] D. Yonar, M. Severcan, R. Gurbanov, A. Sandal, U. Yilmaz, S. Emri, F. Severcan, *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **2022**, *1868*, 166473.
[21] R. Wang, Y. Wang, *International Journal of Molecular Sciences* **2021**, *22*, 1206.
[22] P. D. Lewis, K. E. Lewis, R. Ghosal, S. Bayliss, A. J. Lloyd, J. Wills, R. Godfrey, P. Kloer, L. A. Mur, *BMC Cancer* **2010**, *10*, 640.
[23] X. Li, T. Yang, J. Lin, *JBO* **2012**, *17*, 037003.
[24] D. Ami, A. Duse, P. Mereghetti, F. Cozza, F. Ambrosio, E. Ponzini, R. Grandori, C. Lunetta, S. Tavazzi, F. Pezzoli, A. Natalello, *Anal. Chem.* **2021**, *93*, 16995–17002.
[25] M. Griebe, M. Daffertshofer, M. Stroick, M. Syren, P. Ahmad-Nejad, M. Neumaier, J. Backhaus, M. G. Hennerici, M. Fatar, *Neuroscience Letters* **2007**, *420*, 29–33.
[26] D. Yonar, L. Ocek, B. I. Tiftikcioglu, Y. Zorlu, F. Severcan, *Sci Rep* **2018**, *8*, 1025.
[27] S. Abbas, N. Simsek Ozek, S. Emri, D. Koksal, M. Severcan, F. Severcan, *J Biomed Opt* **2018**, *23*, 1–14.
[28] J.-D. Albert, V. Monbet, A. Jolivet-Gougeon, N. Fatih, M. Le Corvec, M. Seck, F. Charpentier, G. Coiffier, C. Boussard-Pledel, B. Bureau, P. Guggenbuhl, O. Loréal, *Joint Bone Spine* **2016**, *83*, 318–323.
[29] Y. Wu, Z. Wang, M. Xing, B. Li, Z. Liu, P. Du, H. Yang, X. Wang, *JIR* **2022**, *15*, 897–910.
[30] A. D. Morris, C. L. M. Morais, K. M. G. Lima, D. L. D. Freitas, M. E. Brady, A. P. Dhaygude, A. W. Rowbottom, F. L. Martin, *Sci Rep* **2021**, *11*, 9981.
[31] V. Untereiner, R. Garnotel, G. Thiéfin, G. D. Sockalingum, *Anal Bioanal Chem* **2020**, *412*, 805–810.
[32] S. Derruau, C. Gobinet, V. Untereiner, G. D. Sockalingum, A. Nassif, M. Viguier, O. Piot, S. Lorimier, *Journal of Biophotonics* **2021**, *14*, e202000327.
[33] Z. Movasaghi, S. Rehman, Dr. I. ur Rehman, *Applied Spectroscopy Reviews* **2008**, *43*, 134–179.
[34] G. I. Dovbeshko, N. Ya. Gridina, E. B. Kruglova, O. P. Pashchuk, *Talanta* **2000**, *53*, 233–246.
[35] H. Fabian, P. Lasch, D. Naumann, *JBO* **2005**, *10*, 031103.

[36] H. Schulz, M. Baranska, *Vibrational Spectroscopy* **2007**, *43*, 13–25.
[37] G. Shetty, C. Kendall, N. Shepherd, N. Stone, H. Barr, *Br J Cancer* **2006**, *94*, 1460–1464.
[38] M. Huleihel, A. Salman, V. Erukhimovitch, J. Ramesh, Z. Hammody, S. Mordechai, *Journal of Biochemical and Biophysical Methods* **2002**, *50*, 111–121.
[39] J.-G. Wu, Y.-Z. Xu, C.-W. Sun, R. D. Soloway, D.-F. Xu, Q.-G. Wu, K.-H. Sun, S.-F. Weng, G.-X. Xu, *Biopolymers* **2001**, *62*, 185–192.
[40] B. A. Hendrickson, R. Gokhale, J. H. Cho, *Clin Microbiol Rev* **2002**, *15*, 79–94.
[41] W. Ansar, S. Ghosh, *Biology of C Reactive Protein in Health and Disease* **2016**, 67–107.
[42] J. Nallala, O. Piot, M.-D. Diebold, C. Gobinet, O. Bouché, M. Manfait, G. D. Sockalingum, *Cytometry Part A* **2013**, *83A*, 294–300.
[43] D. Benfaremo, M. M. Luchetti, A. Gabrielli, *Journal of Immunology Research* **2019**, *2019*, e8630871.
[44] S. Abdollahi-Roodsaz, S. B. Abramson, J. U. Scher, *Nat Rev Rheumatol* **2016**, *12*, 446–455.
[45] E. Gracey, L. Vereecke, D. McGovern, M. Fröhling, G. Schett, S. Danese, M. De Vos, F. Van den Bosch, D. Elewaut, *Nat Rev Rheumatol* **2020**, *16*, 415–433.
[46] Z. Qaiyum, M. Lim, R. D. Inman, *Semin Immunopathol* **2021**, *43*, 173–192.
[47] F. Salem, N. Kindt, J. R. Marchesi, P. Netter, A. Lopez, T. Kokten, S. Danese, J.-Y. Jouzeau, L. Peyrin-Biroulet, D. Moulin, *United European Gastroenterology Journal* **2019**, *7*, 1008–1032.
[48] E. A. Scoville, M. M. Allaman, D. W. Adams, A. K. Motley, S. C. Peyton, S. L. Ferguson, S. N. Horst, C. S. Williams, D. B. Beaulieu, D. A. Schwartz, K. T. Wilson, L. A. Coburn, *Sci Rep* **2019**, *9*, 2882.
[49] Z. Ito, K. Uchiyama, S. Odahara, S. Takami, K. Saito, H. Kobayashi, S. Koido, T. Kubota, T. Ohkusa, M. Saruta, *DDI* **2018**, *36*, 209–217.
[50] T. D. Saccon, J. M. Dhahbi, A. Schneider, Y. O. Nunez Lopez, A. Qasem, M. B. Cavalcante, L. K. Sing, S. A. Naser, M. M. Masternak, *Biology (Basel)* **2022**, *11*, 508.
[51] R. Gautam, S. Vanga, F. Ariese, S. Umapathy, *EPJ Techn Instrum* **2015**, *2*, 8.
[52] M. J. Baker, S. R. Hussain, L. Lovergne, V. Untereiner, C. Hughes, R. A. Lukaszewski, G. Thiéfin, G. D. Sockalingum, *Chem. Soc. Rev.* **2016**, *45*, 1803–1818.
[53] X. Zhang, G. Thiéfin, C. Gobinet, V. Untereiner, I. Taleb, B. Bernard-Chabert, A. Heurgué, C. Truntzer, P. Ducoroy, P. Hillon, G. D. Sockalingum, *Translational Research* **2013**, *162*, 279–286.
[54] C. Morasso, M. Truffi, R. Vanna, S. Albasini, S. Mazzucchelli, F. Colombo, L. Sorrentino, G. Sampietro, S. Ardizzone, F. Corsi, *Journal of Crohn's and Colitis* **2020**, *14*, 1572–1580.
[55] G. D. O. Lowe, A. Rumley, I. J. Mackie, *Ann Clin Biochem* **2004**, *41*, 430–440.
[56] L. Xue, L. Tao, X. Li, Y. Wang, B. Wang, Y. Zhang, N. Gao, Y. Dong, N. Xu, C. Xiong, T. Zhou, Z. Liu, H. Liu, J. He, K. Li, Y. Geng, M. Li, *Sci Rep* **2021**, *11*, 16903.
[57] L. Lovergne, J. Lovergne, P. Bouzy, V. Untereiner, M. Offroy, R. Garnotel, G. Thiéfin, M. J. Baker, G. D. Sockalingum, *Journal of Biophotonics* **2019**, *12*, e201900177.
[58] L. Lovergne, P. Bouzy, V. Untereiner, R. Garnotel, M. J. Baker, G. Thiéfin, G. D. Sockalingum, *Faraday Discuss.* **2016**, *187*, 521–537.
[59] S. Magalhães, D. Trindade, T. Martins, I. Martins Rosa, I. Delgadillo, B. J. Goodfellow, O. A. B. da Cruz e Silva, A. G. Henriques, A. Nunes, *Clinica Chimica Acta* **2020**, *502*, 25–33.
[60] P. Carmona, M. Molina, E. López-Tobar, A. Toledano, *Anal Bioanal Chem* **2015**, *407*, 7747–7756.
[61] P. Casal-Beiroa, V. Balboa-Barreiro, N. Oreiro, S. Pértega-Díaz, F. J. Blanco, J. Magalhães, *Diagnostics* **2021**, *11*, 546.

SUPPORTING INFORMATION

Additional supporting information is available in the online version of this article at the publisher's website or from the author.