

Classification of Inflammatory Bowel Disease from Formalin-Fixed, Paraffin-Embedded Tissue Biopsies via Imaging Mass Spectrometry

Oliver Klein,* Franz Fogt, Stephan Hollerbach, Grit Nebrich, Tobias Boskamp, and Axel Wellmann

Purpose: Discrimination between ulcerative colitis (UC) and Crohn's disease (CD) by histologic features alone can be challenging and often leads to inaccurate initial diagnoses in inflammatory bowel disease (IBD) patients. This is mostly due to an overlap of clinical and histologic features. However, exact diagnosis is not only important for patient treatment but it also has a socioeconomic impact. It is therefore important to develop and improve diagnostic tools complementing traditional histomorphological approaches. **Experimental Design:** In this retrospective proof-of-concept study, the utilization of MALDI imaging is explored in combination with multi-variate data analysis methods to classify formalin-fixed, paraffin-embedded (FFPE) colon biopsies from UC (87 biopsies, 14 patients), CD (71 biopsies, 14 patients), and normal colonic (21 biopsies, 14 patients) tissues. **Results:** The proposed method results in an overall balanced accuracy of 85.7% on patient and of 80.4% on sample level, thus demonstrating that the assessment of IBD from FFPE tissue specimens via MALDI imaging is feasible. **Conclusions and Clinical Relevance:** The results emphasize the high potential of this method to distinguish IBD subtypes in FFPE tissue sections, which is a prerequisite for further investigations in retrospective multicenter studies, as well as for a future implementation into clinical routine.

1. Introduction

More than five million people worldwide are affected with Crohn's disease (CD) or ulcerative colitis (UC), with 1.4 million in the United States alone, and three million in Europe.^[1] Distinction between these two main forms of chronic inflammatory bowel disease (IBD) can be achieved in most cases based on clinical history, radiologic and endoscopic findings, and histomorphological parameters. Distinguishing both forms of IBD is important, as patients with CD should, for instance, not be exposed to an ileorectal anastomosis due to wound healing problems and significant clinical issue with a small bowel pouch formation. However, up to 20% of IBD cases cannot be classified at the time of diagnosis, which is mainly due to inconclusive endoscopic and histologic data.^[2]

Until a definite diagnosis can be rendered, the respective patients are often diagnosed as indeterminate chronic IBD.^[3]

Dr. O. Klein, G. Nebrich
Charité – Universitätsmedizin Berlin
Corporate member of Freie Universität Berlin Humboldt-Universität zu Berlin, and Berlin Institute of Health
Augustenburger Platz 1, 13353 Berlin, Germany
E-mail: Oliver.klein@charite.de
Dr. O. Klein, G. Nebrich
Berlin-Brandenburg Center for Regenerative Therapies
Charité – Universitätsmedizin Berlin
Augustenburger Platz 1, 13353 Berlin, Germany
Prof. F. Fogt
Penn Presbyterian Medical Center
Hospital of the University of Pennsylvania
51N 39th Street, Philadelphia, PA 19104, USA

Dr. T. Boskamp
Center for Industrial Mathematics
University of Bremen
Bibliothekstr. 5, 28359 Bremen, Germany
Dr. T. Boskamp
SCiLS
Bruker Daltonik GmbH
Fahrenheitstr. 4, 28359 Bremen, Germany
Prof. A. Wellmann
Institute of Pathology
Wittinger Strasse 14, 29223 Celle, Germany
Prof. S. Hollerbach
Department of Gastroenterology
AKH Celle
Siemensplatz 4, 29223 Celle, Germany

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/prca.201900131>

© 2020 The Authors. *Proteomics – Clinical Applications* published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1002/prca.201900131

When confronted with typical biopsy features of UC, such as rectal involvement, mucosal inflammation and continuous disease without significant involvement of the terminal ileum, the diagnosis can be made reliably in most cases. In contrast, cecal sparing, deep “snail-trail” ulcerations and fissures, epithelioid granuloma formation, fibrous stenosis, and significant ileitis with “cobblestone pattern” discloses features that are typically seen in CD. Although CD is usually considered to represent a segmental disease with transmural involvement, there are significant numbers of cases which will show diffuse and/or superficial involvement.^[3,4]

In order to help substantiate presence of either UC or CD, serological tests have been applied as non-invasive diagnostic tools in IBD patients. Such tests include assays for perinuclear antineutrophil cytoplasmic antibodies (pANCA), antibodies against *Saccharomyces cerevisiae* (ASCA), and serum agglutinating antibodies to anaerobic coccoid rods.^[5] However, due to a low sensitivity of these markers, their clinical value appears to be limited.^[6]

Furthermore, linkage analysis and genome wide association studies have uncovered over 100 loci that have significant association with IBD.^[7] Genetic tests from tissue used in discriminating UC from CD include specific genes associated with IBD, such as solute carrier (SLC) 6A14, SLC 26A2, small protein associated with PDZ domain-containing protein 1 (SPAP), regenerating protein IV (RegIV), Vanin-1, matrix metalloproteinase 7 (MMP-7), and growth-related oncogene alpha (Gro- α). While initially a high accuracy for classification models based on these gene expression features was reported,^[8] these results could not be confirmed in follow-up studies.^[9]

Consequently, accurate and reliable diagnostic assays for differential IBD diagnosis are currently not available. Moreover, tissue based proteomic and genomic techniques require large amounts of homogenized tissue material, and do not allow a direct correlation of differentially expressed molecular profiles to tissue histology.

Matrix assisted laser desorption/ionization imaging mass spectrometry (MALDI imaging) enables the label-free and multiplex determination of locally resolved molecular signatures in tissue (e.g., proteins, peptides, lipids, metabolites) and allows their correlation with alterations in tissue histology.^[10,11] In combination with supervised machine learning algorithms, this method shows a high potential for the development of diagnostic histopathological tests.^[12,13] While most research is focused on applications to tumor typing and subtyping,^[14–16] applications to, for example, neurodegenerative diseases,^[17] cardiology,^[18] diabetes,^[19] and regenerative medicine have also been considered.^[20]

The proposed methods primarily rely on the extraction of spectral features from large sets of training data, involving statistical tests for detecting discriminative spectral peaks, as well as computational methods, such as principal component analysis (PCA),^[21] probabilistic latent semantic analysis (PLSA) or non-negative matrix factorization (NMF).^[22,23] The extracted features then form the basis for constructing a subsequent classification scheme based on linear discriminant analysis (LDA), logistic regression, support vector machine (SVM), or random forests (RF). Moreover, the application of convolutional neural networks to MALDI imaging based tissue typing has been proposed.^[24,25]

Clinical Relevance

More than five million people worldwide are affected with Crohn's disease (CD) or ulcerative colitis (UC), the two main forms of chronic inflammatory bowel disease (IBD). In up to approximately 20% of these cases it is impossible to distinguish between both pathologies at the time of initial diagnosis based on clinical and histological findings. However, a correct diagnosis is essential because surgical treatment and long-term prognosis differ for UC and CD.

MALDI imaging mass spectrometry (MALDI imaging) shows promising potential to support diagnostic procedures on tissue samples in clinical pathology. In particular, earlier studies were able to demonstrate that accurate IBD subtyping is possible through MALDI imaging. Since these studies were conducted on fresh frozen colonic tissue, their findings can be verified in multicenter studies only to a very limited extent. In the present work, we examine the use of MALDI imaging in combination with multi-variate data analysis methods to distinguish between CD and UC on formalin-fixed, paraffin-embedded (FFPE) colonic tissue. Our results emphasize the high potential of this method to distinguish IBD subtypes in FFPE tissue sections.

With respect to IBD subtyping, a first study was able to demonstrate that a MALDI imaging based classifier using signatures from intact proteins and an SVM algorithm allows to discriminate CD and UC from fresh frozen colonic tissue at an accuracy of 76.9%.^[26,27] In this work, a histology guided approach was followed, requiring the prior specification of inflammatory mucosa and submucosa regions by a pathologist. Since tissue biopsies from up to six sites around the colon and rectum are typically investigated,^[28] this process is time consuming and induces a high observer dependence. Moreover, for a wide applicability in clinical pathological routine, compatibility with formalin-fixed, paraffin-embedded (FFPE) tissue samples is mandatory, as these constitute most of the tissue collected and stored by pathologists worldwide.

To overcome these limitations, the present study addresses the use of MALDI imaging as a classification tool to discriminate UC, CD, and normal tissue biopsies from FFPE tissue sample sections. Our results suggest that MALDI imaging is able to provide valuable support for the distinct pathological diagnosis of IBD.

2. Results

Colon biopsies (179 samples, 42 patients) from UC (87 biopsies, 14 patients), CD (71 biopsies, 14 patients), and normal tissue (21 biopsies, 14 patients) were analyzed by MALDI imaging. A total of 166 893 spectra were recorded (UC: 90 654, CD: 56 801, N: 19 483). Average spectra of exemplary UC, CD, and normal tissue biopsies are shown in **Figure 1**.

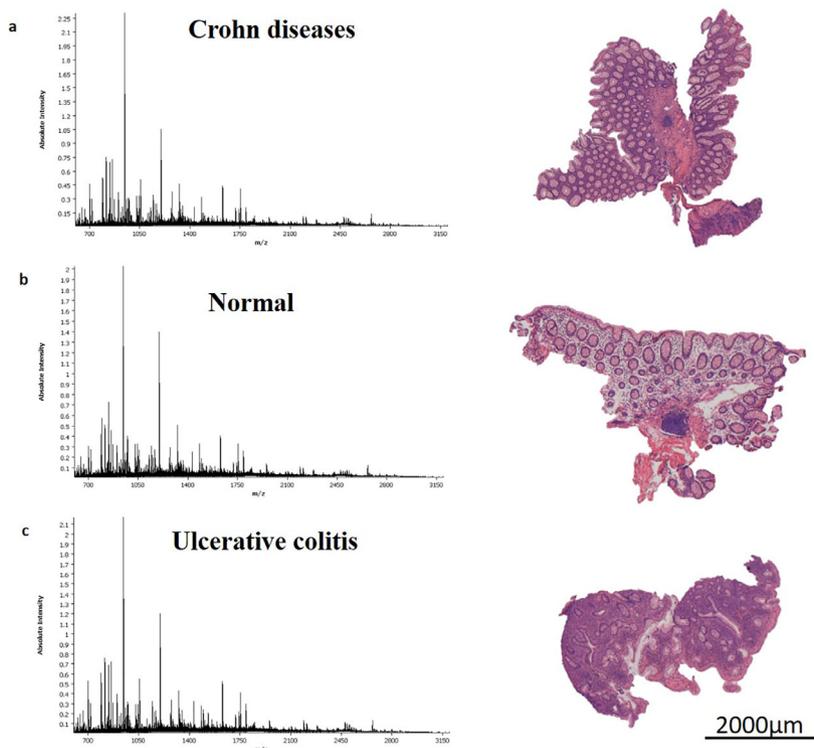


Figure 1. Average spectra of three selected colon biopsies and the corresponding hematoxylin and eosin (HE) staining. a) Crohn's colitis, b) normal colon, and c) ulcerative colitis.

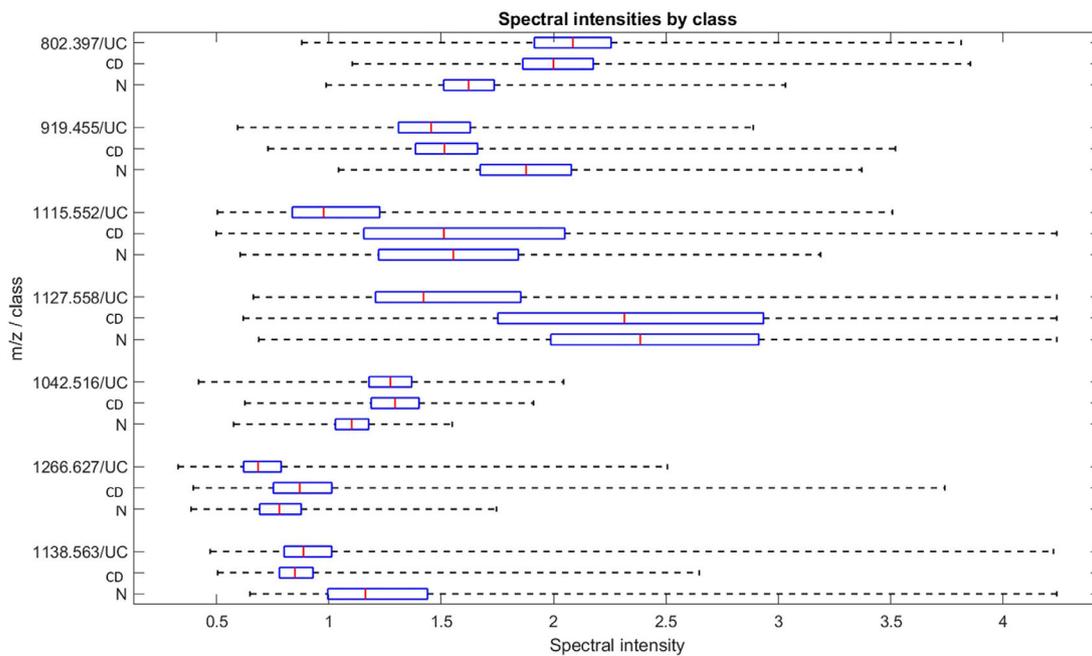


Figure 2. Box plot of spectral intensities for seven discriminative m/z values shown separately for the tissue classes ulcerative colitis (UC), Crohn's disease (CD), and normal (N). Center markers, boxes, and dashed bars represent median values, interquartile and min/max ranges, respectively.

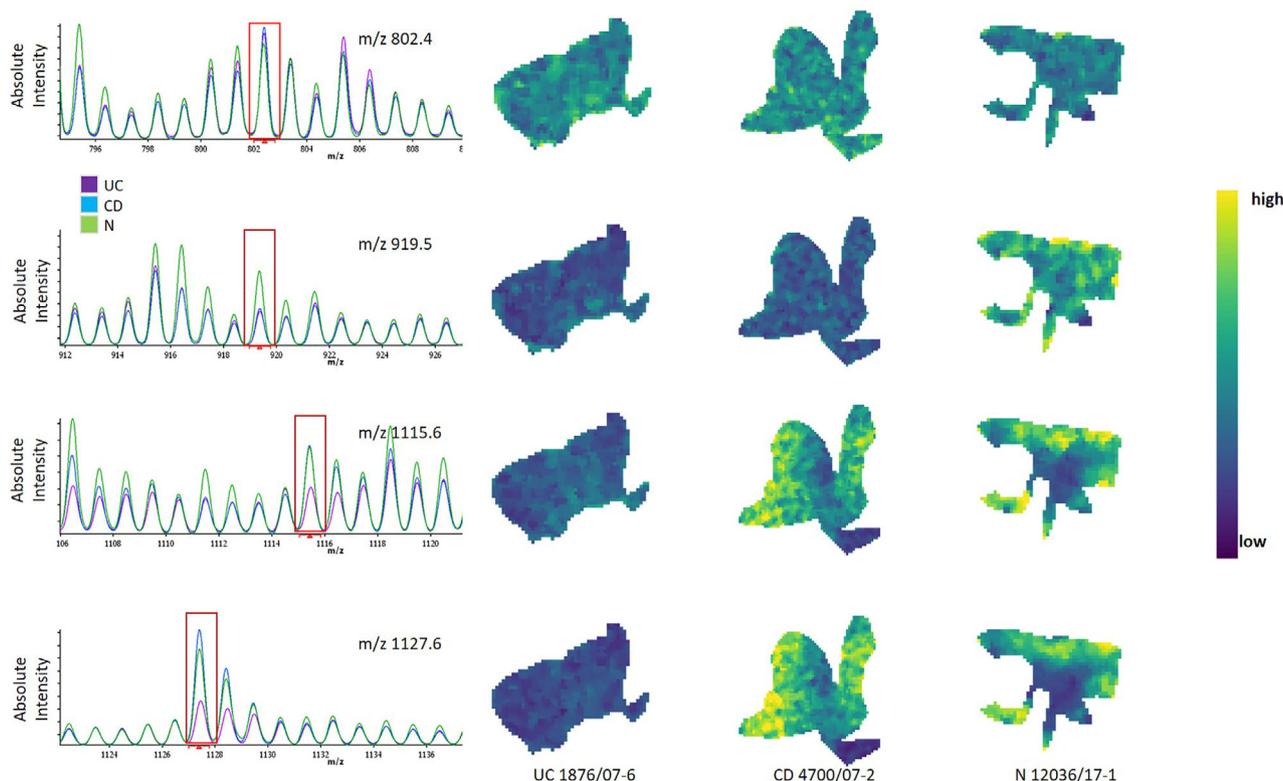


Figure 3. Spectral intensity images of the four most discriminative m/z values shown for three-selected tissue samples corresponding to the classes UC, CD, and N, as well as corresponding mean spectra close-ups.

Table 1. Selected m/z values and their receiver operating characteristic (ROC) values with respect to each of the three tissue classes UC, CD, and N.

m/z	UC	CD	N
802.397	0.648	0.513	0.115
919.455	0.350	0.503	0.854
1115.552	0.192	0.749	0.698
1127.558	0.187	0.744	0.721
1042.516	0.556	0.591	0.167
1266.627	0.252	0.756	0.539
1138.563	0.489	0.359	0.833

2.1. Discriminative Peptide Values

The feature selection procedure described in Section 4, Discriminative Feature Selection, reveals a list of seven m/z values for which the spectral intensity distributions show considerable differences between the three tissue classes (Figure 2, Table 1). An example for intensity images corresponding to the first four of these m/z values and three selected tissue samples (all from the same slide OT1) is shown in Figure 3. As can be seen, the intensities for m/z 919.5, 1115.6, and 1127.6 nicely reflect the overall distributions shown in Figure 2. For m/z 802.4, however, this is not the case in this particular example, which is due to the statistical nature of the feature selection procedure and the strong

variation of the observed features across all samples. The complete set of m/z images for all slides (OT1–OT7) is available as Figure S1, Supporting Information (mz images, generated using SCiLS Lab, with weak denoising and hot spot removal at 97.5% quantile).

2.2. Classification Models for Inflammatory Bowel Disease

The LDA classification models generated in the cross-validation process as described in Section 4, Classification, result in prediction scores for all samples. The distributions of these scores are shown in Figure 4, grouped by slide (OT1–OT7, left to right) and patient (top to bottom, separated by black lines). Each horizontal bar represents prediction scores obtained for a single sample, and incorrect scores, that is, scores indicating the wrong class for this sample, shown in red.

For further evaluating the prediction accuracy on a per-sample and per-patient level, the spot-wise prediction scores obtained from the LDA classification were averaged for each sample and patient, respectively. From the resulting receiver operating characteristic (ROC) curves, the overall multi-class area under curve (AUC) was computed as 90.6% on sample level and 92.8% on patient level. Assuming equal priors for each class, cut-off thresholds were selected to maximize the balanced accuracy, that is, the average of all three sensitivities, resulting in distinct class predictions for each sample and patient (Table 2). As can be seen, discrimination between any of the two conditions of IBD and

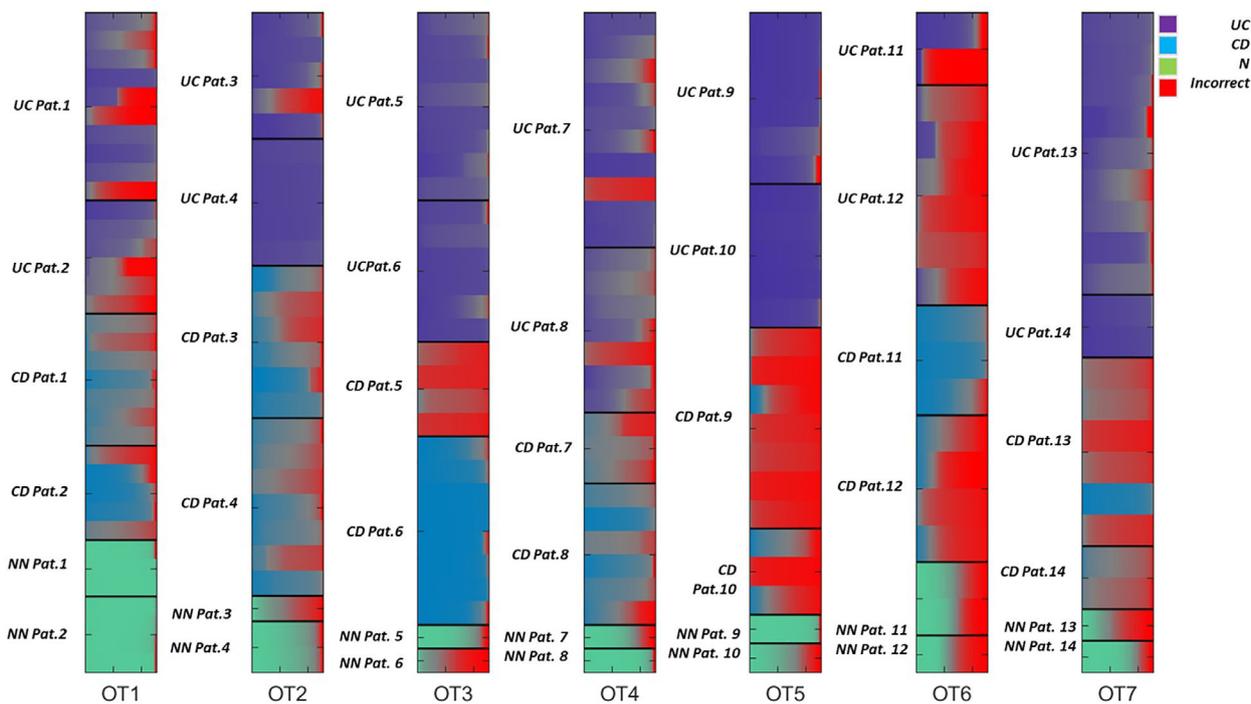


Figure 4. Distribution of classification scores for the seven slides (OT1–OT7, left to right). Horizontal lines represent individual samples, grouped by patient and tissue class. Extent of misclassification shown in red.

Table 2. Classification results per sample and per patient.

Diagnosis	Sample prediction					Patient prediction				
	UC	CD	N	Sensitivity	Specificity	UC	CD	N	Sensitivity	Specificity
UC	64	16	7	73.6%	79.3%	12	2		85.7%	85.7%
CD	19	48	4	67.6%	85.2%	4	10		71.4%	92.9%
N			21	100.0%	93.0%			14	100.0%	100.0%
Balanced accuracy				80.4%					85.7%	

normal tissue is highly accurate on sample level (sens/spec = 93%/100% for colitis versus normal), and even more on patient level (sens/spec = 100%/100%). Discrimination between UC and CD is achieved with a balanced accuracy of 78.6% (UC 85.7%, CD 71.4%) on patient and 70.6% (UC 73.6%, CD 67.6%) on sample level. The overall balanced accuracy for all three classes is 85.7% on patient and 80.4% on sample level.

3. Discussion

Our aim in this pilot study was to investigate the potential of MALDI imaging as a clinical diagnostic tool to differentiate between UC and CD in colonic biopsies. This application scenario requires the compatibility to FFPE tissue specimen, which is why we considered a sample preparation workflow consisting of deparaffination, antigen retrieval, and tryptic digestion. As a result of this complex workflow, the acquired MALDI imaging data contain a significant amount of chemical and technical noise, posing severe challenges to the subsequent data analysis methods.

Our data-preprocessing pipeline included a spatial denoising step to increase the signal-to-noise ratio, as well as downsampling to intervals of 0.4 Da width around the theoretically expected peptide masses. The latter processing step improves the robustness against mass misalignment, which is likely to occur in MALDI TOF data. Moreover, it heavily reduces the data dimensionality, which is beneficial for the subsequent feature selection and classification algorithms. Thus, we were able to develop a completely automated supervised machine learning workflow that does not require prior peak picking or visual feature selection. In order to achieve this level of automation and to simplify the cross-validation process described here, we had to perform our data analysis using a programmable data analysis framework, such as MATLAB. For subsequent validation of the discriminative spectral features derived in our analysis, it will be possible to perform the classification directly in SCiLS Lab, as the classification algorithm used in our work (LDA) is no different from the standard algorithms found in SCiLS Lab or other commercial products.

Based on a set of seven spectral features (m/z values), a classification model was trained that allowed to discriminate between the three classes UC, CD, and normal at an AUC of 92.8% on patient level, evaluated using cross validation on slide level (“leave one slide out”). With respect to discriminating IBD (UC or CD) from normal, even a perfect classification result was achieved on patient level (100% sensitivity and specificity). When considering the IBD cases only, discrimination between UC and CD patients was possible at a balanced accuracy of 78.6%.

In our study design, we decided to include multiple samples per patient (varying between 1 and 10, median = 4), as this is consistent with the typical diagnostic routine workflow, where often up to six biopsy samples or more are investigated. On the other hand, we did not pursue the histology guided MALDI imaging approach, which would require prior identification of “hot spots” of inflammatory regions by the pathologist. Instead, the classification results are obtained on tissue samples that may include both normal and inflammatory regions for an UC or CD patient. In view of this background, we consider the reported accuracy levels even more remarkable.

In order to arrive at a robust classification model linked to putative protein signatures associated with the two forms of IBD, we took care to avoid obvious sources of confounding. For example, we mixed patients of all three diagnostic categories (UC, CD, normal) on each slide, and designed the cross validation scheme to strictly separate on both the patient and the slide levels. Moreover, for improved robustness, cross validation was utilized not only for estimating the classification accuracy, but also for performing the feature selection.

As a caveat, however, we point out that the final feature set of seven m/z values was in fact obtained taking into account data from the full study cohort. Nevertheless, our study highlights the potential of MALDI imaging to derive proteomic signatures for the discrimination between UC and CD, and thus for supporting differential diagnosis of IBD in clinical pathological routine. The disease manifestation and characteristic are heterogeneously distributed across the different biopsies from the individual patients. Based on the HE staining, morphological structures and disease manifestations (crypt reconstruction, inflammation) were assessed and aligned to the peptide intensity maps from the MALDI imaging analysis. In case of, for example, m/z value 1115.6 Da, an increased intensity distribution could be observed in regions of mostly regular crypts in UC and CD. In contrast, areas with high inflammation and massive crypt reconstruction (severe disease manifestation) revealed a decreased spectral intensity (Figure S2, Supporting Information).

The proportion of regions with actual disease manifestations in the IBD biopsies have an influence on the accuracy of the classification. In our study, the misclassified biopsies were associated with histological findings of mostly regular morphological structures and almost no disease manifestations (see example in Figure S2, Supporting Information). However, by considering all available biopsies, the classification accuracy could be increased from 80.4% on sample level to 85.7% on patient level. Thus, different proportions of disease manifestations in the IBD biopsies had only little influence on the classification on patient level, even without further time-consuming annotation of the biopsies by the diagnostic pathologist. Nevertheless, a consideration of a minimum of six biopsies per patient, as usually recommended

in clinical routine^[29] remains advisable also for analysis by MALDI Imaging.

This retrospective study demonstrates that MALDI imaging based diagnostic assessment of IBD from FFPE tissue is technically and practically feasible. As FFPE tissue samples constitute the largest part of all tissue collected in clinical routine, the results of this study lay the foundations for more extensive validation in multi-center studies. Future work will focus on comparative clinical follow-up studies with larger patient cohorts, further investigation of the seven m/z values and their identification by proteomic techniques, as well as analyzing their potential correlation with histologic features.

4. Experimental Section

Study Cohort and Histopathological Examination: FFPE colon biopsies of 42 patients with UC, CD, and no IBD (normal) were retrospectively recruited and diagnostically evaluated by an experienced, board-certified pathologist (A.W.). The research study was implemented in accordance with the relevant guidelines and regulations. Samples used in this study were available from the tissue archive of the Institute of Pathology, Celle, and were subjected to strict anonymization. Utilization of anonymized archival material in retrospective studies without explicit patient consent was approved by the ethical committee of MH Hannover (393-2008).

Chemicals: Alpha-cyano-4-hydroxycinnamic acid (HCCA) was obtained from Bruker Daltonik (Bremen, Germany), ammonium bicarbonate, acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA) from Fluka (St. Louis, USA) and trypsin from Promega (Madison, WI, USA). 100% TFA was purchased from Merck (spectroscopy purity, Darmstadt, Germany) and ammonium hydrogen phosphate from Sigma (St. Louis, USA).

Tissue Preparation: Tissue preparation and MALDI imaging data acquisition were performed as previously described.^[30] Briefly, paraformaldehyde (PFA)-fixed specimens were dehydrated by washing sequentially with increasing concentrations of ethanol, subsequently cleared in xylene and embedded in paraffin. 6 μm FFPE sections were transferred to a total of seven indium-tin-oxide slides (Bruker Daltonik, Bremen, Germany). Sections were dewaxed and passed through decreasing concentrations of ethanol according to a protocol adapted from Casadonte et al.^[11] Citric acid antigen retrieval (CAAR) adapted from a previously described method^[31] was performed. After drying slides for 10 min, tryptic digest was performed. Using an automated spraying device (ImagePrep, Bruker Daltonik, nine spraying cycles), 200 μL trypsin solution (20 μg , 20 mM ammonium bicarbonate/acetonitrile 9:1) was applied onto the section. This procedure took about 15 min; the following settings were used: 39% spray power, 0% modulation (offset: 20), 9 spray cycles, spray time of 1.25 s, drying time of 50 s. After tissue incubation (3 h at 37 °C, moist chamber), matrix solution (1 mL 7 g L⁻¹ HCCA in 50% ACN and 1% TFA) was deposited using ImagePrep (15% power, \pm 40% modulation, 60 spraying cycles).

MALDI Imaging Analysis: MALDI imaging data acquisition was executed as previously described.^[30] Briefly, analysis was performed in mass detection range of m/z 600–3500, 200 laser shots per spot, sampling rate of 1.25 GS s⁻¹, and raster width of 50 μm on a Rapiflex MALDI time of flight (TOF) using flexControl 3.0 and flexImaging 3.0 (Bruker Daltonik). External calibration was performed using a peptide calibration standard (Bruker Daltonik). Spectra processing was performed in flexAnalysis 3.0 (Bruker Daltonik). After MALDI imaging experiments, matrix was removed with 70% ethanol and tissue sections were hematoxylin and eosin (HE) stained.^[11]

Data Preprocessing: Spectral data of each of the seven slides were loaded into SCiLS Lab (version 2018a, SCiLS, Bremen, Germany) separately, and total ion count (TIC) normalization was applied. For subsequent processing, the data were exported to MATLAB (version 2016b, MathWorks, Natick, MA, USA). The spectra were cropped to the mass range m/z 700–2700, and spatial denoising was performed using local averaging

on a 5-by-5 neighborhood around each spot. Next, spectra were resampled to intervals of 0.4 Da width centered around expected peptide masses according to the idealized “averagine” peptide mass model.^[32] Thus, the dimensionality of the data was reduced to 1999 intensity values per spectrum.

Discriminative Feature Selection: Selection of discriminative peptide masses was performed by receiver operating characteristic (ROC) analysis. Based on the patients’ diagnosis, spectra were assigned to one of the three classes UC, CD, and normal (N). For each class and each m/z value the Mann–Whitney–Wilcoxon statistic was computed, which is equivalent to the empirical probability P that the spectral intensity at the given m/z value in the given class is larger than that of a spectrum in any of the other classes.^[33] For selecting those m/z values that were either over- or under-represented in one of the classes, the ROC ranking criterion $r = |P - 0.5|$ (ROC score) was computed, and the top 10 m/z values were chosen. In order to increase the robustness of this feature extraction step, the above analysis was not performed on the full data, but instead on partial data using a k -fold cross validation scheme. Patients were randomly split into $k = 5$ subsets of approximately same size, each including all three diagnoses. By removing one of these subsets in turn, five different training sets were obtained, and ROC ranking was performed on each of these training sets independently, resulting in five ranked lists of 10 m/z values each. Those m/z values occurring in only one or two lists were discarded, leaving a list of seven m/z values with high ROC ranking in at least two training subsets. A complete table of all top ranked m/z values for each of the five cross validation subsets, including the corresponding ROC values for each of the three classes, as well as the maximum ROC score across all classes, is available (Table S1, Supporting Information—Peaks ROC values).

Classification: Given the discriminative features computed in the previous step and the corresponding class information, LDA^[34] was applied using the implementation available in the MATLAB Statistics and Machine Learning Toolbox (MathWorks, Natick, MA, USA). Since the authors wanted to evaluate the classification method’s robustness toward technical variation between individual measurements, they chose to perform cross validation on slide level instead of patient level. Seven cross validation folds were formed, each one consisting of data from six slides as training data on which an LDA classification model was trained. The respective remaining subset in each fold was used as test data. Combining classification results from all folds, a full set of predictions was obtained and compared to the ground truth class information. A multi-class extension of the area under curve (AUC) metric^[35] was used to evaluate the classification result.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors thank Angelika Krajewski for providing excellent technical assistance in the imaging mass spectrometry sample preparation. This work was supported by grants from the BCRT through funding by the German Federal Ministry of Education and Research (BMBF). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data availability Statement

The data sets generated and/or analyzed in the current study are available on request from the corresponding author.

Keywords

Crohn’s disease, imaging mass spectrometry, inflammatory bowel disease, ulcerative colitis

Received: December 12, 2019

Revised: May 25, 2020

Published online: August 4, 2020

- [1] J. Burisch, T. Jess, M. Martinato, P. L. Lakatos, E. EpiCom, *J. Crohn’s Colitis* **2013**, *7*, 322.
- [2] G. E. Tontini, M. Vecchi, L. Pastorelli, M. F. Neurath, H. Neumann, *World J. Gastroenterol.* **2015**, *21*, 21.
- [3] W. J. Tremaine, *Gastroenterol. Hepatol.* **2011**, *7*, 826.
- [4] T. C. DeRoche, S. Y. Xiao, X. Liu, *Gastroenterol. Rep.* **2014**, *2*, 178.
- [5] E. Viennois, Y. Zhao, D. Merlin, *Inflammatory Bowel Dis.* **2015**, *21*, 2467.
- [6] X. Bossuyt, *Clin. Chem.* **2006**, *52*, 171.
- [7] B. D. Ye, D. P. McGovern, *Expert Rev. Clin. Immunol.* **2016**, *12*, 1091.
- [8] P. von Stein, R. Lofberg, N. V. Kuznetsov, A. W. Gielen, J.-O. Persson, R. Sundberg, K. Hellstrom, A. Eriksson, R. Befrits, A. Ost, O. D. von Stein, *Gastroenterology* **2008**, *134*, 1869.
- [9] J. T. Bjerrum, C. Nyberg, J. Olsen, O. H. Nielsen, *J. Intern. Med.* **2014**, *275*, 484.
- [10] A. Walch, S. Rauser, S. O. Deininger, H. Hofler, *Histochem. Cell Biol.* **2008**, *130*, 421.
- [11] R. Casadonte, R. M. Caprioli, *Nat. Protoc.* **2011**, *6*, 1695.
- [12] K. Schwamborn, R. M. Caprioli, *Nat. Rev. Cancer* **2010**, *10*, 639.
- [13] R. D. Addie, B. Balluff, J. V. Bovee, H. Morreau, L. A. McDonnell, *Anal. Chem.* **2015**, *87*, 6426.
- [14] J. Kriegsmann, M. Kriegsmann, R. Casadonte, *Int. J. Oncol.* **2015**, *46*, 893.
- [15] R. Casadonte, M. Kriegsmann, F. Zweynert, K. Friedrich, G. Baretton, M. Otto, S.-O. Deininger, R. Paape, E. Belau, D. Suckau, D. Aust, C. Pilarsky, J. Kriegsmann, *Proteomics* **2014**, *14*, 956.
- [16] K. A. Veselkov, R. Mirnezami, N. Strittmatter, R. D. Goldin, J. Kinross, A. V. M. Speller, T. Abramov, E. A. Jones, A. Darzi, E. Holmes, J. K. Nicholson, Z. Takats, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1216.
- [17] Y. Ucal, Z. A. Durer, H. Atak, E. Kadioglu, B. Sahin, A. Coskun, A. T. Baykal, A. Ozpinar, *Biochim. Biophys. Acta, Proteins Proteomics* **2017**, *1865*, 795.
- [18] F. Spillmann, S. Van Linthout, G. Schmidt, O. Klein, N. Hamdani, T. Mairinger, F. Krackhardt, B. Maroski, T. Schlabs, S. Soltani, S. Anker, E. V. Potapov, D. Burkhoff, B. Pieske, C. Tschöpe, *Eur. Heart. J.* **2019**, *40*, 2164.
- [19] S. Miyamoto, C.-C. Hsu, G. Hamm, M. Darshi, M. Diamond-Stanic, A.-E. Declèves, L. Slater, S. Pennathur, J. Stauber, P. C. Dorrestein, K. Sharma, *EBioMedicine* **2016**, *7*, 121.
- [20] O. Klein, K. Strohschein, G. Nebrich, M. Fuchs, H. Thiele, P. Givalisco, G. N. Duda, T. Winkler, J. H. Kobarg, D. Trede, S. Geissler, *Sci. Rep.* **2018**, *8*, 12677.
- [21] S. O. Deininger, M. Becker, D. Suckau, *Methods Mol. Biol.* **2010**, *656*, 385.
- [22] T. Boskamp, D. Lachmund, J. Oetjen, Y. C. Hernandez, D. Trede, P. Maass, R. Casadonte, J. Kriegsmann, A. Warth, H. Dienemann, W. Weichert, M. Kriegsmann, *Biochim. Biophys. Acta, Proteins Proteomics* **2017**, *1865*, 916.
- [23] J. Leuschner, M. Schmidt, P. Fernsel, D. Lachmund, T. Boskamp, P. Maass, *Bioinformatics* **2018**, *35*, 1940.

- [24] J. Behrmann, C. Etmann, T. Boskamp, R. Casadonte, J. Kriegsmann, P. Maaß, *Bioinformatics* **2018**, *34*, 1215.
- [25] O. Klein, F. Kanter, H. Kulbe, P. Jank, C. Denkert, G. Nebrich, W. D. Schmitt, Z. Wu, C. A. Kunze, J. Sehouli, S. Darb-Esfahani, I. Braicu, J. Lellmann, H. Thiele, E. T. Taube, *Proteomics Clin. Appl.* **2019**, *13*, e1700181.
- [26] E. H. Seeley, M. K. Washington, R. M. Caprioli, A. E. M'Koma, *Proteomics – Clin. Appl.* **2013**, *7*, 541.
- [27] A. E. M'Koma, E. H. Seeley, M. K. Washington, D. A. Schwartz, R. L. Muldoon, A. J. Herline, P. E. Wise, R. M. Caprioli, *Inflammatory Bowel Dis.* **2011**, *17*, 875.
- [28] E. F. Stange, S. P. L. Travis, S. Vermeire, W. Reinisch, K. Geboes, A. Barakauskiene, R. Feakins, J. F. Fléjou, H. Herfarth, D. W. Hommes, L. Kupcinskas, P. L. Lakatos, G. J. Mantzaris, S. Schreiber, V. Villanacci, B. F. Warren, *J. Crohn's Colitis* **2008**, *2*, 1.
- [29] T. Kucharzik, A. U. Dignass, R. Atreya, B. Bokemeyer, P. Esters, K. Herrlinger, K. Kannengießer, P. Kienle, J. Langhorst, A. Lügering, S. Schreiber, A. Stallmach, J. Stein, A. Sturm, N. Teich, B. Siegmund, *Z Gastroenterol.* **2019**, *57*, 1321.
- [30] O. Klein, K. Strohschein, G. Nebrich, J. Oetjen, D. Trede, H. Thiele, T. Alexandrov, P. Giavalisco, G. N. Duda, P. von Roth, S. Geissler, J. Klose, T. Winkler, *Proteomics* **2014**, *14*, 2249.
- [31] J. O. Gustafsson, M. K. Oehler, S. R. McColl, P. Hoffmann, *J. Proteome Res.* **2010**, *9*, 4315.
- [32] M. W. Senko, S. C. Beu, F. W. McLafferty, *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 229.
- [33] T. Fawcett, *Pattern Recognit. Lett.* **2006**, *27*, 861.
- [34] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, Springer, New York **2009**.
- [35] D. J. Hand, R. J. Till, *Mach. Learn.* **2001**, *45*, 171.