### ABSTRACT

**Background** – Crohn’s disease and ulcerative colitis are the primary inflammatory bowel diseases (IBD), and their pathogenesis is related to genetic and environmental factors. Currently, the diagnosis of IBD results in a multidisciplinary approach with significant disadvantages, such as its invasive nature, time spent, and the fact that 10% of patients remain without diagnostic classification. However, new methodologies of analysis have emerged that allowed the expansion of knowledge about IBD, as the metabolomics, the study of metabolites. The presence and prevalence of such metabolites may prove to be useful as biomarkers in the diagnosis of IBD. **Objective** – Analyze fecal samples for metabolic analysis in the diagnosis of inflammatory bowel diseases (IBD), providing differentiation between Crohn’s disease and ulcerative colitis. **Methods** – This is an observational study with 21 patients diagnosed with IBD (ulcerative colitis 11 and Crohn’s disease 10) and 15 healthy controls, all with the consent and clarification. The fecal extracts of all patients are submitted to a high-resolution Nuclear Magnetic Resonance Hydrogen (1H-NMR) spectroscopy combined with multivariate and univariate pattern recognition techniques. Through the metabolomics of fecal extracts, gives us a characterization of employing a noninvasive approach. **Results** – We identify some metabolites, such as lactate, succinate, alanine, and tyrosine, in the Crohn’s disease fecal samples, and leucine, alanine, and tyrosine in ulcerative colitis fecal samples. All the amino acids presented positive covariance for disease correlation. **Conclusion** – The results showed different metabolic profiles between IBD patients and healthy volunteers based on 1H-NMR analysis of fecal extracts. Moreover, the approach discriminated patients with Crohn’s disease and ulcerative colitis. The metabolomics analysis is promising as a novel diagnostic technique for further IBD recognition and surveillance. New studies are necessary to validate these findings.

### HEADINGS


### INTRODUCTION

Inflammatory bowel disease (IBD) is a chronic gastrointestinal tract disorder due to its inflammation and epithelial lesions. Crohn’s disease (CD) and ulcerative colitis (UC) are significant IBD, and its etiology is still unknown. Despite all efforts to reach the diagnosis and to define a more individualized treatment for patients with CD and UC, following the ECCO-ESGAR Guideline for diagnostic assessment in inflammatory bowel disease, a single standard as reference for the diagnosis of CD or UC until now does not exist. Genetic and environmental factors are believed to be involved in its pathogenesis. Also, there is evidence that IBD is related to intestinal microbiota dysbiosis. Currently, in the absence of trustful evidence to support the best modality to assess these diseases, IBD diagnosis is based on a multidisciplinary approach which involves clinical history, endoscopy, radiology, histology, microbiology, and hematology. Colonoscopy is currently used as the gold standard diagnostic method for the evaluation and monitoring of IBD. This process has the following drawbacks: its invasive nature, and 10% to 15% of the patients remain without diagnostic classification. Technologies such as metabolomics and metagenomic have emerged in the last decade, offering alternatives for identification and differentiation of inflammatory bowel diseases. Such technologies are more effective and less invasive, and they open new perspectives in the diagnosis of this class of disease.

The use of metabolomics is an emerging approach for analyzing the physiology and cellular pathophysiology through the identification of metabolic profiles. Metabolomics is in the group of omics (transcriptomics, genomics, and proteomics), related to investigations of biochemical profiles, defined as an analysis of macromolecules (metabolites) produced by the metabolism of an organism, organ or living tissue, considering the genetic expression under the influence of disease, lifestyle, and environmental factors. The metabolomics analysis monitors the metabolic state enabling the understanding of complex biological interactions, and defines biomarkers related to specific conditions; in our study, IBDs.

Preliminary studies reveal that metabolomics can differentiate between individuals with IBD and control groups, as well as between the IBD subtypes, using, for example, metabolic profiles of fecal extracts. Metabolomics is still an under-explored approach to identify biomarkers in fecal samples of patients with IBD, and can be potentially used for disease surveillance. In this sense, several markers have been proposed to assess the disease activity.
Among the tools that have good performance for the study of metabolomics, the High-Resolution Nuclear Magnetic Resonance Hydrogen (1H-NMR) stands out\textsuperscript{11,12}. This methodology is attractive due to the simplicity involved in the sample preparation and for its ability to discover the chemical identities of unknown peaks in the formed spectra. 1H-NMR spectroscopy of fecal samples, tissue, and urine successfully distinguishes patients with CD and UC from healthy individuals\textsuperscript{6,13,14}.

This study investigates the potential of 1H-NMR spectroscopy in the identification of metabolites presented in fecal samples of individuals with IBD. We verify the differences between the metabolic expression profile of healthy patients from those with CD and UC. The focus is on the definition and differentiation of potential biomarkers in the diagnosis of inflammatory bowel diseases through the metabolomics analysis of fecal samples.

METHODS

This was an observational and transversal study. The work was approved by the Ethics Committee of the Federal University of Alagoas (CAAE: 55813816.2.0000.5013). We use an 1H-NMR spectrometer to analyze the stool samples from 36 individuals: CD (n=10, 6 males, 4 females, and age between 25 and 60 years), UC (n=11, 8 females, 3 males, age between 25 and 70 years) and healthy controls, HC (n=15, 9 females, 6 males, age between 30 and 79 years, without a history of any pathology in the gastrointestinal tract).

Fecal specimen collection and preparation (FIGURE 1): First, the fecal samples are collected and, immediately, begins the process obtaining fecal water. The sample is added to the mortar, and the mixture is crushed with distilled water and liquid nitrogen. Next, the samples in Eppendorf are centrifuged and the supernatant is collected and stored for 1H-NMR analysis.

Obtaining the 1H-NMR spectra: All 1H-NMR experiments were performed at 298.15 K (Kelvin) using a Bruker 400 MHz spectrometer (Bruker Biospin). The spectra were acquired with 64K dots, 128 transients (scans) with the spectral width of 10.5 parts per million (ppm) and suppression of the water signal, using the pulses sequence noesypr1d, with 2 seconds (s) of relaxation time, 100 milliseconds (ms) of mixing time. The results were processed in the TOPSPIN (V 3.5, Bruker Biospin) program. The chemical shifts (δ) were calibrated concerning the signal of TSP at 0.0 ppm, and were manually corrected for phase and baseline distortions.

Data analysis

Despite the suppression of the water signal, we removed the water of region δ 5.00-4.70ppm from all spectra. The 1H-NMR spectra data set were normalized by the median, compensating the differences in sample concentration, using MetaboAnalyst 4.0. We applied Partial Least Squares Discriminant Analysis (PLS-DA) to the 1H-NMR spectral data set to facilitate the interpretation, discriminating against the of fecal biochemical profiles of each group. We identified the metabolites responsible for the difference between UC and CD, and between HC and UC by the Variance Analysis of Cross-Validated Residues CV-ANOVA; we set the $P$-value threshold at 0.05 to verify the validity of the models\textsuperscript{15}. The model is considered valid when the permutation test and the CV-ANOVA test are satisfied at the same time.

RESULTS

Investigation of the 1H-NMR

The 1H-NMR spectra came from all the 36 patients (diagnosed with CD, UC, and the HC). The area referring to the aromatic hydrogen signals, δ 6.0 to 9.0, was insufficient to identify the cues and to distinguish each one of them since the signals were of low intensity. Also, it was not possible to visually perceive a pattern of differences between the spectra of the two data sets (HC versus UC, and UC versus CD). Therefore, the use of features such as chemometrics to identify the signals became necessary.

Identification of the signals in the 1H-NMR spectra

The signals were attributed to compounds belonging to different classes, such as short chain fatty acids (SCFA), amino acids, other organic acids intermediate of Krebs cycle, amines, and nucleotides. TABLE 1 shows the assignment of the signals. The identification of the metabolites was possible by comparing with the data in the Human Metabolome Database (available at hmdb.ca) and with the use of CHENOMXTM (evaluation mode).

The metabolites found are in agreement with the literature. The intermediates of the Krebs cycle and SCFA are possible metabolites in fecal samples\textsuperscript{16}, as well as the metabolites such as butyrate, propionate, and lactate\textsuperscript{16}. FIGURE 2 shows the metabolites identified in TABLE 1. Due to the complexity of the spectra, some signals not yet were found.

Analysis with PLS-DA and CV-ANOVA ($P$-value <0.05)

The ANOVA helped reduce the possibility large number of metabolites investigated, keeping only those that showed the most reliable response under the conditions studied, to discriminate against the groups\textsuperscript{15}.  

![FIGURE 1. Material and methods for the preparation of fecal water and analyzed by 1H-NMR.](image-url)
The results were validated through Cross-Validation, with an accuracy of 100.0%. Both samples, HC and CD, were well classified, showing 100% of classification. The distinction between these profiles has already been observed in previous studies using the same methodology \(^{(9)}\).

Through ANOVA, it was possible to identify which metabolites discriminate against the HC and the CD samples. Lactate (δ \(1.33\) \(P=0.0050338\)), succinate (δ \(2.41\) \(P=0.0020639\)), alanine (\(P=0.001\)) and tyrosine (\(P=0.002\)) presented a positive covariance for Crohn’s disease, with a strong correlation for disease discrimination. An important data was the sign of ethanol (δ \(1.19\) \(P=0.00011198\)) with positive covariance for the HC samples, which presented such a greater quantity to discriminate CD.

### Healthy controls (HC) x ulcerative colitis (UC)

The results of PLS-DA between the HC and CD groups (FIGURE 3) showed that the groups are well defined and demonstrate a significant difference in Metabolism. Component # 2 (the variance between spectra \(^1\)H-NMR in the same group) was 22.6%, and Component # 1 (the variance between the sets generated by HC and CD \(^1\)H-NMR spectra), was 40.02%.

The Cross-Validation test resulted in 69% of precision. That is, from a total of 11 UC samples, 6 (54.5%) were well classified and 5 (45.5%) presented similarity to the HC. Among the 15 samples diagnosed as HC, 12 (80.0%) were classified as HC, and 3 (20.0%) were not grouped in any class.

By ANOVA (\(P<0.05\)) the discriminant metabolites of UC comparing with HC, were leucine (δ 3.36 \(P=0.0050338\)), alanine (δ 1.48 \(P=0.000001\)), and tyrosine (δ 6.91 \(P=0.0020639\)). All three amino acids presented positive covariance in disease correlation. Again, the ethanol (δ 1.19) discriminated against the HC samples.

### TABLE 1. Chemical displacement of identified metabolites.

<table>
<thead>
<tr>
<th>Molecular name</th>
<th>Chemical displacement</th>
<th>Multiplicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetate</td>
<td>1.92</td>
<td>Singlet</td>
</tr>
<tr>
<td>Alanine</td>
<td>1.48</td>
<td>Doublet</td>
</tr>
<tr>
<td>Butyrate</td>
<td>0.90</td>
<td>Triplet</td>
</tr>
<tr>
<td>Butyrate</td>
<td>1.36</td>
<td>Doublet</td>
</tr>
<tr>
<td>Butyrate</td>
<td>2.16</td>
<td>Triplet</td>
</tr>
<tr>
<td>Choline</td>
<td>3.21</td>
<td>Singlet</td>
</tr>
<tr>
<td>Creatinine</td>
<td>3.05</td>
<td>Singlet</td>
</tr>
<tr>
<td>Ethanol</td>
<td>1.19</td>
<td>Triplet</td>
</tr>
<tr>
<td>Ethanol</td>
<td>3.66</td>
<td>Quartet</td>
</tr>
<tr>
<td>Formate</td>
<td>8.46</td>
<td>Singlet</td>
</tr>
<tr>
<td>Glutamate</td>
<td>2.36</td>
<td>Doublet of Triplet</td>
</tr>
<tr>
<td>Glycine</td>
<td>3.56</td>
<td>Singlet</td>
</tr>
<tr>
<td>Hypoxanthine</td>
<td>8.2</td>
<td>Doublet</td>
</tr>
<tr>
<td>Lactate</td>
<td>1.33</td>
<td>Doublet</td>
</tr>
<tr>
<td>Lactate</td>
<td>4.12</td>
<td>Quartet</td>
</tr>
<tr>
<td>Leucine</td>
<td>0.96</td>
<td>Triplet</td>
</tr>
<tr>
<td>Leucine</td>
<td>1.71</td>
<td>Multiplet</td>
</tr>
<tr>
<td>Leucine</td>
<td>3.73</td>
<td>Triplet</td>
</tr>
<tr>
<td>Lysine</td>
<td>3.03</td>
<td>Triplet</td>
</tr>
<tr>
<td>Methanol</td>
<td>3.36</td>
<td>Singlet</td>
</tr>
<tr>
<td>Propionate</td>
<td>1.06</td>
<td>Triplet</td>
</tr>
<tr>
<td>Phenyllalanine</td>
<td>7.39</td>
<td>Multiplet</td>
</tr>
<tr>
<td>Succinate</td>
<td>2.41</td>
<td>Singlet</td>
</tr>
<tr>
<td>TMAO</td>
<td>3.27</td>
<td>Singlet</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>6.91</td>
<td>Doublet</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>7.2</td>
<td>Doublet</td>
</tr>
<tr>
<td>Trimethylamine</td>
<td>2.89</td>
<td>Singlet</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>7.55</td>
<td>Doublet</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>7.74</td>
<td>Doublet</td>
</tr>
</tbody>
</table>

The results were validated through Cross-Validation, with an accuracy of 100.0%. Both samples, HC and CD, were well classified, showing 100% of classification. The distinction between these profiles has already been observed in previous studies using the same methodology\(^{(9)}\). Through ANOVA, it was possible to identify which metabolites discriminate against the HC and the CD samples. Lactate (δ \(1.33\) \(P=0.0050338\)), succinate (δ \(2.41\) \(P=0.0020639\)), alanine (\(P=0.001\)) and tyrosine (\(P=0.002\)) presented a positive covariance for Crohn’s disease, with a strong correlation for disease discrimination. An important data was the sign of ethanol (δ \(1.19\) \(P=0.00011198\)) with positive covariance for the HC samples, which presented such a greater quantity to discriminate CD.

### FIGURE 2. Spectrum \(^1\)H-NMR 400 MHz of the fecal water samples.

### FIGURE 3. Fingerprint of fecal metabolomic profiles analyzed by \(^1\)H-NMR. HC vs CD. PLS-DA plot of water fecal obtained from healthy controls (green dots) and CD patients (red dots). Component 1, 40.2% and component 2, 22.6%.

The PLS-DA applied to data set HC x UC (FIGURE 4) showed homogeneity among the \(^1\)H-NMR spectra of the same group, with the value of component #2 of 9.9%, despite component #1 with 26.1% presenting a significant variance. This finding has already reported in previous studies using the same methodology\(^{(9)}\).

The Cross-Validation test resulted in 69% of precision. That is, from a total of 11 UC samples, 6 (54.5%) were well classified and 5 (45.5%) presented similarity to the HC. Among the 15 samples diagnosed as HC, 12 (80.0%) were classified as HC, and 3 (20.0%) were not grouped in any class.

### FIGURE 4. Score plot of PLS-DA of water fecal obtained from healthy controls (green dots) and CD patients (red dots). Components 1 and 2, 40.2% and 22.6%, respectively.

### FIGURE 5. Spectrum \(^1\)H-NMR 400 MHz of the fecal water samples.
Lins Neto MAF, Verdi GMX, Veras AO, Veras MO, Caetano LC, Ursulino JS.
Use of metabolomics to the diagnosis of inflammatory bowel disease

Several culture-independent analyses of the diversity of microbiota associated with IBD have shown that the groups *Clostridium coccoides* and *C. leptum* decrease when compared to healthy subjects\(^8\)\(^{-\text{10}}\). These bacterial groups are functionally important, as they are primarily responsible for the production of SCFA, which provides energy to intestinal cells and promotes epithelial growth\(^9\). This dysbiosis may occur due to the destruction of microbiota by the T lymphocytes, which are hyperactive against bacterial antigens and promote inflammatory destruction of the epithelial cells, which could lead to malabsorption of nutrients\(^9\)\(^{10}\). We found increased amounts of amino acids, including alanine, tyrosine, and leucine, in the feces of patients with CD and UC when compared to the HC. This effect may be the consequence of malabsorption caused by the inflammation.

After analyzing the data with the PLS-DA statistical methodology, the results showed that the metabolic analyses of fecal water by \(^1\)H-NMR spectrometry were useful in differentiating the patients with inflammatory bowel disease from healthy individuals. Although CD and UC share many epidemiological, immunological, therapeutic, and clinical characteristics, the results of metabolomics also have shown the formation of two distinct subtypes of IBD. The changes in metabolic profiles observed from UC subjects were not as marked as those observed with the CD group. It suggests that the metabolic consequences caused by inflammation are more severe in the CD group than in the UC group.

The literature reports distinctions between metabolic profiles within the disease, depending on the target affected\(^8\)\(^{-\text{10}}\)\(^{13}\). In individuals with CD, the studies have shown that the patterns affecting mainly ileum (CDI) and patterns primarily affecting the colon (CDC) are quite distinct. However, the CDC profile is similar to the profile of HC, and there is no such significant difference between HC and CDC\(^8\). Our study distinguishes through metabolomics, the metabolic profile among UC, CD, and HC. Therefore, the distinction of metabolomics in the disease at its different stages and locations is a promising future research in the metabolomics analysis of fecal samples of patients with inflammatory bowel disease.

Despite the small of the sample in this study, it was possible to infer that based on the results and responding to our objective, metabolomics is a valuable tool for investigation and determination of the distinction between Crohn’s disease and ulcerative colitis. This alternative proved to be exciting and corroborates our results with those found in the literature.

**DISCUSSION**

This study focuses on the use of \(^1\)H-NMR spectrometry in the diagnosis of inflammatory bowel disease. Several studies have shown that \(^1\)H-NMR metabolomics is promising. Patients with identical clinical evaluation and diagnosed with IBD may have different outcomes. In addition, the metabolites found in IBD when compared to the HC group, show changes that may imply response and/or resistance to therapy. Through the metabolomics it was not possible to discriminate the final products of the fermentation of SCFA (butyrate, propionate), having identified the increase of the intermediate compounds, suggesting a dysbiosis of the intestinal microbiota. The increase of intermediate metabolites, including lactate and succinate, appears as a prominent feature in CD patients if compared with healthy subjects. Studies suggest that the rise of such metabolites is due to a dysbiosis of the host intestinal microbiota\(^4\)\(^{-\text{9}}\). Bacteria of the genus *Phascolarctobacterium* behaves as exclusive consumers of succinate and propionate when associated with bacteria of the genus *Paraprevotella*\(^17\). Thus, the decrease of *Phascolarctobacterium* bacteria in IBD may reflect in a reduction in propionate production\(^17\). Studies also showed that propionate is formed from pyruvate, by decarboxylation of succinate or via the acrylate pathway, where lactate is reduced to propionate\(^3\). Thus, with the reduction of the succinate-consuming bacteria and the propionate-producing bacteria, the intermediate compounds of succinate and lactate raise considerably.

Several culture-independent analyses of the diversity of microbiota associated with IBD have shown that the groups *Clostridium coccoides* and *C. leptum* decrease when compared to healthy subjects\(^8\)\(^{-\text{10}}\). These bacterial groups are functionally important, as they are primarily responsible for the production of SCFA, which provides energy to intestinal cells and promotes epithelial growth\(^9\). This dysbiosis may occur due to the destruction of microbiota by the T lymphocytes, which are hyperactive against bacterial antigens and promote inflammatory destruction of the epithelial cells, which could lead to malabsorption of nutrients\(^9\)\(^{10}\). We found increased amounts of amino acids, including alanine, tyrosine, and leucine, in the feces of patients with CD and UC when compared to the HC. This effect may be the consequence of malabsorption caused by the inflammation.

After analyzing the data with the PLS-DA statistical methodology, the results showed that the metabolic analyses of fecal water by \(^1\)H-NMR spectrometry were useful in differentiating the patients with inflammatory bowel disease from healthy individuals. Although CD and UC share many epidemiological, immunological, therapeutic, and clinical characteristics, the results of metabolomics also have shown the formation of two distinct subtypes of IBD. The changes in metabolic profiles observed from UC subjects were not as marked as those observed with the CD group. It suggests that the metabolic consequences caused by inflammation are more severe in the CD group than in the UC group.

The literature reports distinctions between metabolic profiles within the disease, depending on the target affected\(^8\)\(^{-\text{10}}\)\(^{13}\). In individuals with CD, the studies have shown that the patterns affecting mainly ileum (CDI) and patterns primarily affecting the colon (CDC) are quite distinct. However, the CDC profile is similar to the profile of HC, and there is no such significant difference between HC and CDC\(^8\). Our study distinguishes through metabolomics, the metabolic profile among UC, CD, and HC. Therefore, the distinction of metabolomics in the disease at its different stages and locations is a promising future research in the metabolomics analysis of fecal samples of patients with inflammatory bowel disease.

Despite the small of the sample in this study, it was possible to infer that based on the results and responding to our objective, metabolomics is a valuable tool for investigation and determination of the distinction between Crohn’s disease and ulcerative colitis. This alternative proved to be exciting and corroborates our results with those found in the literature.

**CONCLUSION**

It was possible to differentiate between IBD patients and healthy volunteers with \(^1\)H-NMR metabolic profiles of fecal extracts. Our study identified patients with CD and UC using this approach. Metabolomics has shown promising as a novel and a noninvasive diagnostic technique for further disease recognition and surveillance. Ongoing studies are necessary to validate these findings.

**ACKNOWLEDGEMENTS**

The authors are thankful to the Brazilian Ministry of Health for partial financial support. The authors also thank the Chemical Department and Medical School of Federal University of Alagoas.
(UFAL), for support in the analysis of the fecal samples, and the Institute of Multidisciplinary Skills in Intestinal Microbiota of UFAL, for collecting and storing the fecal samples.

Authors’ contribution


RESUMO – Contexto – A doença de Crohn e retocolite ulcerativa são as principais doenças inflamatórias intestinais (DII), e sua patogênese está relacionada a fatores genéticos e ambientais. Atualmente, o diagnóstico de DII resulta em uma abordagem multidisciplinar e apresenta desvantagens significativas, como sua natureza invasiva, tempo gasto e o fato de 10% dos pacientes permanecerem sem classificação diagnóstica. No entanto, surgiram novas metodologias de análise que permitiram ampliar o conhecimento sobre a DII, como a metabolômica, o estudo dos metabólitos. A presença e a prevalência desses metabólitos podem ser úteis como biomarcadores no diagnóstico da DII. Objetivo – Avaliar as amostras fecais por análise metabolômica no diagnóstico de DII, diferenciando os perfis metabólicos entre doença de Crohn e retocolite ulcerativa. Métodos – Estudo observacional com 36 indivíduos (doença de Crohn 11, retocolite ulcerativa 10 e 15 controles saudáveis), todos com consentimento esclarecido. Os extratos fecais de todos os pacientes são submetidos a uma espectroscopia de alta resolução por ressonância magnética nuclear de hidrogênio ("H-NMR) combinada com técnicas de reconhecimento de padrões multivariados e univariados. Por meio da metabolômica utilizando extratos fecais, foi possível obter uma caracterização adequada das doenças inflamatórias intestinais através de uma abordagem não invasiva. Resultados – Foi possível identificar os seguintes metabólitos nos pacientes com doença de Crohn: lactato, succinato, alanina e tirosina e, no grupo retocolite ulcerativa encontreu-se leucina, alanina e tirosina. Todos os aminoácidos apresentaram covariância positiva para a doença. Conclusão – Os resultados demonstraram diferentes perfis metabólicos entre pacientes com DII e voluntários saudáveis, com base na análise por "H-NMR dos extratos fecais. Além disso, pacientes com doença de Crohn e retocolite ulcerativa também podem ser discriminados usando essa abordagem. A análise metabolômica é promissora como uma nova técnica não invasiva de diagnóstico para melhor reconhecimento das DII. Novos estudos são necessários para validar esses achados.


REFERENCES