



DOI:10.2478/rrlm-2022-0004

An insight into the fecal microbiota composition in Romanian patients with ankylosing spondylitis using high-throughput 16S rRNA gene amplicon sequencing

Mihaela Oprea^{1#}, Daniela Cristea^{2#}, Sorin Dinu¹, Simona Adriana Ciontea², Violeta Claudia Bojinca^{3,4}, Denisa Predeteanu³, Andra Rodica Balanescu^{3,4}, Codruta Romanita Usein^{1*}

1. Molecular Epidemiology for Communicable Diseases Laboratory, "Cantacuzino" National Military Medical Institute for Research and Development, Bucharest, Romania
2. Enteric Bacterial Infections Laboratory, "Cantacuzino" National Military Medical Institute for Research and Development, Bucharest, Romania
3. Department of Internal Medicine and Rheumatology, "Sf. Maria" Clinical Hospital, Bucharest, Romania
4. Internal Medicine and Rheumatology, "Carol Davila" University of Medicine and Pharmacy, Bucharest, Romania

Abstract

Introduction. Application of next-generation sequencing technology generated a massive amount of information on the gut microbiome composition used to understand its role in the healthy state and in various diseases. We aimed to provide information on the gut microbiota composition of Romanian subjects diagnosed with ankylosing spondylitis, an immune-mediated arthropathy linked to a genetic predisposition and gut dysbiosis. **Methods.** Stool samples collected from 25 patients with ankylosing spondylitis and 16 healthy controls were investigated using high-throughput DNA sequencing of 16S rRNA amplicons from seven different hypervariable regions and Ion Torrent PGM instrument. Microbial composition of metagenomic data was analyzed with QIIME software and differential abundance analysis of taxa encompassed linear discriminant analysis effect size (LEfSe). **Results.** Overall, 14 phyla, 114 families, 114 genera, and 275 species were identified across the 41 samples, the aggregated data revealing as most abundant the phyla Bacteroidetes, Firmicutes, and Proteobacteria, the families Bacteroidaceae, Prevotellaceae, and Ruminococcaceae, the genera Bacteroides, Prevotella, and Faecalibacterium, and Prevotella copri species. Using various cutoffs for abundance and prevalence, core taxonomic members were identified which in general were shared between the patients and controls. However, evidence was gained that the diversity in the microbiomes from the former cohort was lower than for controls and that certain taxa had significantly different abundance between the two groups. **Conclusion.** This study allowed an informative high-throughput 16S rRNA profiling of the gut microbiota needed to identify microbiome signatures of risk in the autochthonous population with AS.

Keywords: ankylosing spondylitis, next generation sequencing, gut microbiome, 16S rRNA amplicon sequencing

Received: 23rd November 2021; Accepted: 5th January 2022; Published: 12th January 2022

* **Corresponding author:** Codruta Romanita Usein, Molecular Epidemiology for Communicable Diseases Laboratory, "Cantacuzino" National Military Medical Institute for Research and Development, Bucharest, Romania E-mail: codrtausein@gmail.com

These authors contributed equally

Introduction

Ankylosing spondylitis (AS) is a chronic progressive inflammatory rheumatic disease which mainly affects the spine and sacroiliac joints and has as dominant clinical features back pain and progressive stiffness of the spine (1). According to epidemiological studies, AS ranges from 9 to 30 per 10,000 in the general population, the heterogeneity in estimates depending on geographic area, study population or data source, case definition, and ascertainment methods (2). Although progress has been made in understanding the etiology and pathogenesis of this handicapping disease there are still unclear aspects that need ongoing investigation. The genetic predisposition indicated by the striking association of this condition with the human major histocompatibility complex (MHC) class I allele HLA-B27 along with the microbiome and its interaction with the immune system are the major topics in AS pathogenesis (3). Concerning the microbial involvement, the dysbiosis of the intestinal microbiota is the hypothesis prioritized in research (4, 5).

Gut microbiota has been traditionally analyzed by culture methods which in spite of improving in time in terms of techniques for cultivation of anaerobic species provided a limited resolution of the analysis. The addition of the molecular approach conferred the possibility to increase its accuracy by allowing the identification of viable but non-culturable or other difficult-to-culture as well as unknown microorganisms (6). The further application of next-generation sequencing (NGS) technology generated a massive amount of information on the gut microbiome configuration used to progress in understanding the complexity of the microbial communities and their function in health and disease. However, with each microbiome fingerprint added, emerging priorities and challenges for research in this field arise (7). Currently, little has been published about the distinct microbiome configurations across healthy

and/or diseased Romanian population. In order to allow meaningful comparisons of autochthonous data with those accumulated for individuals from other regions, this report provides baseline microbiological information on the gut microbiota composition of Romanian subjects with AS diagnosis as determined through the culture-independent amplification of 16S ribosomal RNA (rRNA) gene and an NGS approach.

Material and methods

Study participants

This observational retrospective study was linked to a convenience cohort of adults who were admitted throughout a two-year period (i.e., January 2016 - December 2017) at Saint Maria Clinical Hospital in Bucharest, a well-known academic medical center with a large catchment area. The 41 participants enrolled into the cohort were 25 patients (20 males and 5 females, 22 - 68 years/ mean age 49 years) with AS diagnosis, defined according to the modified New York classification criteria for AS (8) and 16 male control subjects (21 - 77 years/ mean age 50 years) on no prescription medication. Sixteen patients from the AS group were positive for the HLA-B27 risk allele. The subjects forming the HC (healthy control) group were self-reported healthy individuals with no prior history of gastrointestinal, rheumatic disease, or other chronic disorders and did not take any antibiotics or immunosuppressive agents during the month prior to the inclusion in the study. All the participants to the study enrolled voluntarily and provided their written informed consent. The study protocol was approved by the Ethics Committee of the hospital (ref. 1442/13.07.2018).

Samples and DNA extraction

Participants were provided with sterile specimen jars for fecal collection and were informed of the need to retrieve spontaneously passed fresh stool

samples. Immediately upon receipt in the laboratory the stool samples were stored at -80°C until further processing.

Microbial DNA was extracted from each fecal sample using the commercial QIAamp DNA Stool Mini Kit (Qiagen) following the manufacturer's instructions and quantified on the Qubit 3.0 Fluorometer (Invitrogen) with Qubit dsDNA HS Assay kit (Invitrogen).

16S rRNA gene sequencing

Amplification of rRNA gene hypervariable regions was performed by using the Ion 16STM Metagenomics Kit (Thermo Scientific). DNA library was obtained following the protocol proposed by Ion Torrent, including Ion Plus Fragment Library kit (Thermo Scientific), Ion Xpress Barcode Adapters (Thermo Scientific), and Agencourt AMPure XP DNA (Beckman Coulter) for purification steps. DNA concentration of the libraries was estimated by qPCR using Ion Universal Library Quantitation Kit (Thermo Scientific). Emulsion PCR was carried out in OneTouch 2 instrument, using Ion PGMTM Hi-QTM View OT2 Kit (Thermo Scientific). The enriched libraries were loaded on a 318 Ion Torrent Chip and sequences were obtained with an Ion Torrent PGM system using Ion PGMTM Hi-QTM View Sequencing Kit (Thermo Scientific) with 400 bp chemistry.

Data processing

Sequences were processed by PGM Torrent Suite software which removed all low quality and polyclonal sequences. Quality approved, trimmed, and filtered resulting reads were then exported to Ion Reporter software pipeline which performed automated analysis, annotation, and taxonomic assignment via Ion Metagenomics 16S analysis workflow 1.1 version 5.12. Operational Taxonomic Units (OTUs) were reported based on the alignment to MicroSEQ ID and Greengenes 16S rRNA gene databases us-

ing Quantitative Insights into Microbial Ecology (QIIME) software package vs. 1.9.1. Alpha metrics for diversity within each sample (Chao1 and Shannon indexes) and beta diversity metrics (Bray-Curtis index) for diversity between samples were calculated using QIIME. Principal-coordinate analysis (PCoA) was performed with Bray-Curtis dissimilarities for further visualization of the differences between the samples.

Linear discriminant analysis (LDA) Effect Size (LEfSe) algorithm performed with the Galaxy tool available at the public Huttenhower Galaxy server [<http://huttenhower.sph.harvard.edu/galaxy/>] was used to identify the taxa that were differentially abundant between the groups of samples investigated with the default size-effect threshold of 2.0 on the logarithmic LDA score (9).

The significance of differences for microbiome characteristics between the subjects investigated were compared using Mann-Whitney U test and a p value of < 0.05 was accepted as indicating statistical significance.

The fastq files reported in this study were deposited in European Nucleotide Archive (ENA) database under PRJEB37736 project (ERS4423124 – ERS 4423164).

Results

Reads from libraries generated from the Ion 16STM Metagenomics Kit were assigned Operational Taxonomic Units (OTUs) allowing identification of 14 phyla, 114 families, 114 genera, and 275 species. A detailed overview of taxa identified in the study samples for the AS and HC groups is presented in the supplementary information included in Table S1 and Table S2.

At phylum level, the majority of the reads were assigned to *Bacteroidetes*, *Firmicutes*, and *Proteobacteria* phyla which aggregated constituted 97.84% in the microbiome data in AS patients and 96.31% in the healthy controls.

The mean abundance of the major phyla for the AS group indicated *Bacteroidetes* as the main phylum (41.28%), followed by *Firmicutes* (38.28%) and *Proteobacteria* (18.28%). Certain variations were observed in relationship with the HLA-B27 positivity within the AS group where *Bacteroidetes* dominated the microbiome composition in more HLA-B27-positive than HLA-B27-negative subjects (11/16 versus 5/9), while *Firmicutes* was the most abundant phylum in more subjects of the latter subset (3/9 versus 3/16). The data from the HC group showed that *Firmicutes* had the highest average abundance (41.75%), followed by *Bacteroidetes* (38.5%), and *Proteobacteria* (16.06%).

Other 11 phyla were identified, all with OTU counts < 1%: *Actinobacteria*, *Cyanobacteria*, *Defferibacteria*, *Elusimicrobia*, *Fusobacteria*, *Lentisphaera*, *Nitrospinae*, *Spirochaetes*, *Synergistetes*, *Tenericutes*, and *Verrucomicrobia*. Of them, *Tenericutes* and *Synergistetes* were more abundant in the microbiomes of the HC group by comparison with the AS group ($p < 0.05$).

At family level, 79 of the 114 taxa identified were shared and 36 taxa were exclusively present in either AS group (16 families) or HC group (20 families). A significantly lower number of taxa were identified among the AS microbiomes (range 17 – 59, average 28.5) by comparison with the HC microbiomes (range 23 – 55, average 37.5) ($p < 0.05$), an observation also reflected by the alpha diversity measures output by

QIIME showing a lower gut microbiome richness ($p < 0.01$) and evenness ($p < 0.05$) for the AS patients (Figure 1a).

The aggregated data indicated *Bacteroidaceae* (AS group, 21.38% and HC group, 14.19%), *Prevotellaceae* (AS group, 13.47% and HC group, 16.50%), and *Ruminococcaceae* (AS group, 8.98% and HC, 11.38%) as the most abundant families. In addition, other families had an average abundance that exceeded 5% in at least one of the studied groups, namely *Veillonellaceae* (HC group, 8.38%), *Enterobacteriaceae* (AS group, 7.42% and HC group, 7.42%), *Lachnospiraceae* (AS group, 6.91 and HC group, 7.43%), and *Sutterellaceae* (AS group, 5.2%).

To capture the characteristics of the gut microbiomes, a core of taxonomic members was defined at family level taking into consideration two parameters: the taxon abundance per sample and the proportion of the samples that the taxon of interest was detected in. Taxon abundance cut-offs were set at 5%, 0.5%, and 0.05% and related to 30%, 50%, and 75% prevalence, respectively. Eight core members were distinguished in both the AS group and HC group at the highest abundance and lowest prevalence cutoffs, their number increasing to 14 which exceeded the 0.05% abundance in the AS group and 16 in the HC group, respectively. Thirteen families remained ubiquitous and abundant enough to be considered core representative for both AS and HC groups (Table 1).

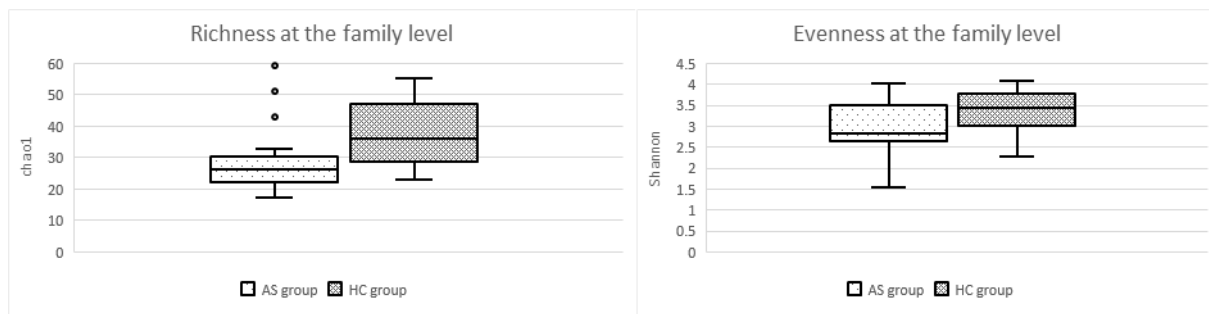


Fig. 1a. Box plots showing alpha diversity of the gut microbiome in AS and HC groups at family level

Table 1. Core members at selected abundance and prevalence cutoffs at family, genus, and species levels for AS and HC groups

Abundance	5%		0.5%		0.05%	
Prevalence	30%		50%		75%	
	Group					
Family	AS	HC	AS	HC	AS	HC
<i>Acidaminococcaceae</i>	+	-	+	+	-	+
<i>Bacteroidaceae</i>	+	+	+	+	+	+
<i>Clostridiaceae</i>	-	-	+	+	+	+
<i>Coriobacteriaceae</i>	-	-	-	-	+	-
<i>Desulfovibrionaceae</i>	-	-	+	+	+	+
<i>Enterobacteriaceae</i>	+	+	+	+	+	+
<i>Erysipelotrichaceae</i>	-	-	-	+	-	+
<i>Eubacteriaceae</i>	-	-	+	+	+	+
<i>Hyphomicrobiaceae</i>	-	-	+	+	+	+
<i>Lachnospiraceae</i>	+	+	+	+	+	+
<i>Lactobacillaceae</i>	-	-	+	+	+	+
<i>Oscillospiraceae</i>	-	-	-	+	-	+
<i>Porphyromonadaceae</i>	-	+	+	+	+	+
<i>Prevotellaceae</i>	+	+	+	+	+	+
<i>Rikenellaceae</i>	-	-	-	+	-	+
<i>Ruminococcaceae</i>	+	+	+	+	+	+
<i>Suterellaceae</i>	+	+	+	+	+	+
<i>Veillonellaceae</i>	+	+	+	+	+	+
Genus						
<i>Alistipes</i>	-	-	+	+	-	+
<i>Bacteroides</i>	+	+	+	+	+	+
<i>Barnesiella</i>	-	-	-	+	-	+
<i>Bilophila</i>	-	-	-	+	-	+
<i>Clostridium</i>	-	-	+	+	+	+
<i>Dialister</i>	-	+	-	+	-	-
<i>Eubacterium</i>	-	-	+	+	+	+
<i>Faecalibacterium</i>	+	+	+	+	+	+
<i>Gemmiger</i>	-	-	+	+	+	+
<i>Lactobacillus</i>	-	-	+	+	-	+
<i>Odoribacter</i>	-	-	-	-	-	+
<i>Parabacteroides</i>	-	-	+	+	+	+
<i>Phascolarctobacterium</i>	+	-	+	-	-	-
<i>Prevotella</i>	+	+	+	+	-	+
<i>Roseburia</i>	-	-	+	+	+	+
<i>Ruminococcus</i>	-	-	+	+	+	+
<i>Sutterella</i>	+	+	+	+	+	+
Species						
<i>Alistipes putredinis</i>	-	-	-	+	-	+
<i>Alistipes shahii</i>	-	-	-	-	-	+
<i>Bacteroides uniformis</i>	-	-	+	+	+	+
<i>Bacteroides massiliensis</i>	-	-	-	+	-	+
<i>Bacteroides vulgatus</i>	+	-	+	+	+	+
<i>Bilophila wadsworthia</i>	-	-	-	+	-	+
<i>Barnesiella intestinihominis</i>	-	-	-	+	-	+
<i>Eubacterium eligens</i>	-	-	-	+	-	+
<i>Faecalibacterium prausnitzii</i>	+	+	+	+	+	+
<i>Gemmiger formicilis</i>	-	-	+	+	+	+
<i>Lactobacillus rogosae</i>	-	-	-	+	-	+
<i>Odoribacter splanchnicus</i>	-	-	-	-	-	+
<i>Parabacteroides distasonis</i>	-	-	-	+	-	+
<i>Parabacteroides merdae</i>	-	-	-	-	-	+
<i>Prevotella copri</i>	+	+	+	+	+	-
<i>Roseburia faecis</i>	-	-	+	-	+	+
<i>Ruminococcus gnavus</i>	-	-	-	-	+	+

The families identified only in the AS group or HC group were mostly low abundant taxa identified in no more than 3 microbiomes within each group. In AS patients, 9 taxa were identified in single microbiomes (i.e., *Brachyspiraceae*, *Bradyrhizobiaceae*, *Cardiobacteriaceae*, *Desulfuromonadaceae*, *Entomoplasmataceae*, *Halobacteroidaceae*, *Phormidiaceae*, *Spirochaetaceae*, *Staphylococcaceae*), 5 in two microbiomes (i.e., *Actinomycetaceae*, *Bdellovibrionaceae*, *Corynebacteriaceae*, *Peptoniphilaceae*, *Rhodobacteraceae*), and one taxon was identified in three microbiomes (i.e., *Marinilabiliaceae*). With the exception of *Desulfuromonadaceae* with 16 mapped reads (0.0015%), the other taxa displayed abundances between 0.01% and 1.59%. In controls, 15 taxa with abundances between 0.01% and 0.77% were present in single microbiomes (i.e., *Alicyclobacillaceae*, *Alcaligenaceae*, *Colwelliaceae*, *Clostridiales* Family XII Incertae Sedis, *Cyclobacteriaceae*, *Elusimicrobiaceae*, *Geodermatophilaceae*, *Heliobacteriaceae*, *Magnetococcaceae*, *Microbacteriaceae*, *Mycobacteriaceae*, *Oligosphaeraceae*, *Piscirickettsiaceae*, *Thermoanaerobacteraceae*, *Thermoanaerobacterales* Family IV Incertae Sedis,) and 5 taxa in two microbiomes (i.e., *Beijerinckiaceae*, *Defluviitaleaceae*, *Flammeovirgaceae*, *Proteini-voraceae*, *Rhodothalassiaceae*). At genus level, the number of taxa identified within a microbiome ranged from 22 to 54 (me-

dian 31) and 28 to 49 (median 36) in the AS group and HC group, respectively. Seventy-three genera were shared by AS patients and controls of which *Bacteroides*, *Clostridium*, *Eubacterium*, *Faecalibacterium*, and *Roseburia* were present in all the microbiomes. Overall, the most abundant genera were *Bacteroides* (AS group, 24.86% and HC group, 20.22%), *Prevotella* (AS group, 15.57% and HC group, 17.76%), and *Faecalibacterium* (AS group, 8.89% and HC group, 11.46%). *Dialister* was an additional genus remarked as abundant but only for the HC group (HC group, 8.43% versus AS group, 2.52%).

The core composition of microbiomes changed at various taxon abundance and prevalence cut-offs, revealing similarities and differences between the samples of AS group and HC groups (Table 1). Alpha diversity metrics showed an increased diversity in samples from the HC group (Figure 1b).

PCoA plots based on Bray-Curtis dissimilarity index indicated no significant clustering in community composition throughout the samples investigated at family and genus levels between the AS and HC groups (data not shown).

At species level, out of the 275 species identified across the 41 sampled microbiomes, which belonged to 95 genera, 160 species were shared by the AS and HC groups, 92 were found only in the AS group, and 23 were unique to the HC group. Within the AS group, there were between 32 and

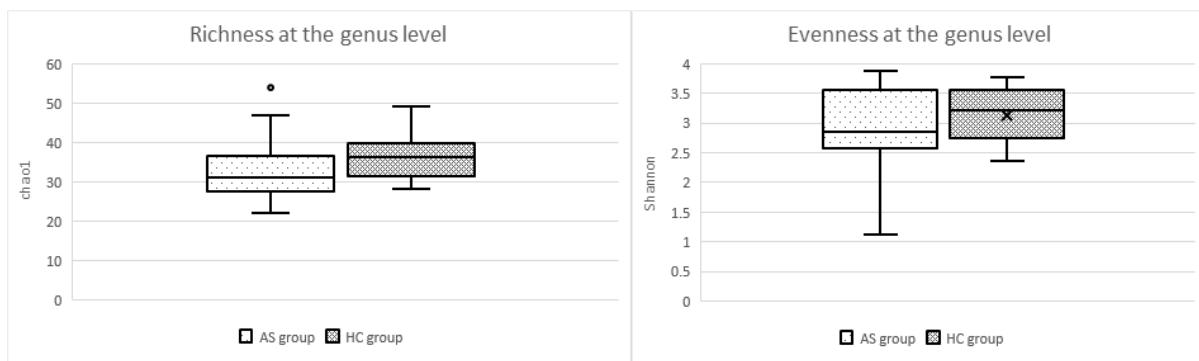


Fig. 1b. Box plots showing alpha diversity of the gut microbiome in AS and HC groups at genus level

98 species (median 53) assigned per microbiome whereas the HC group displayed from 39 to 77 species (median 54) per microbiome. The accuracy of the taxonomic assignment provided a lower resolution as the clustering of reads into OTUs did not hit a species name for 22 genera. The highest discrimination was achieved for *Bacterioides* (27 species) and *Clostridium* (22 species) genera. Of note, no canonical pathogens such as *Vibrio cholerae*, *Campylobacter jejuni*, and *Salmonella enterica* were detected at this level of sensitivity.

Overall, the aggregated data showed *Prevotella copri* (AS group, 15.32% and HC group, 15.93%) as being the most abundant species. Other conspicuous species were *Bacterioides vulgatus* (11.66%) and *Faecalibacterium prausnitzii* (9.15%) for the AS group and *Faecalibacterium prausnitzii* (11.62%), *Dialister invisus* (5.69%), and *Bacterioides vulgatus* (5.53%) for the HC group.

As the abundance cutoff was decreased, the number of species forming the core microbiome increased more for the HC group than for the AS group, nine taxa being added to the former and only four to the latter at 0.05% abundance cutoff (Table 1).

The histogram of the LDA scores used to detect features with significantly different abundance levels between assigned taxa showed 32 taxa differences between the AS and HC groups of which only 5 were more abundant in the former (Figure 2).

Moreover, LEfSe revealed abundance differences for 8 taxa within the AS group in relation to the HLA-B27-positive status as well (Figure 3).

Discussion

The knowledge gathered by consistent efforts devoted to the microbial community analysis is making room for introducing microbiota tailoring in daily clinical practice (10). Driven by the

interest to accelerate the adoption of gut microbiome profiling for research into autochthonous population diseases and given the emerging evidence that gut microbiome may play a role in AS pathogenesis, we aimed to investigate the gut microbiome composition of an adult population-based cohort comprising 25 individuals with AS and 16 subjects without this condition. To the best of our knowledge, this is the first Romanian pan-bacterial 16S rRNA gene microbiome survey of AS patients performed with massively parallel DNA sequencing technology. Recently, another Romanian study addressed the intestinal microbiota in AS which used realtime polymerase chain reactions to target a limited number of bacteria (11).

Concerns about the influence of the methodology used on the accuracy of the results and the confidence in their comparisons have been shared by research groups and studies designed to reduce the errors occurred along the NGS experimental flow, focusing on the different aspects, have been conducted to confirm or refute the data (12, 13). Biases introduced by the technology on which the sequencing platform is founded have been reported (14). We used the Ion Torrent's PGM, one of the main short read platforms in use in small to medium laboratories with high throughput capabilities which utilizes semiconductor sequencing technology that operates on acidity (pH) instead of light as do Illumina's platforms. Comparisons concerning capabilities in microbiome composition application of different platforms showed that Ion Torrent allowed valid biological conclusions (14 -16).

Another factor impacting the final output of 16S taxonomic profiling is represented by the differences in taxonomic informativeness of the 16S rRNA gene regions especially when high-throughput short-read sequencing technologies are used. We decided upon a mix of primers released by Life Technologies for the Ion Torrent platform for amplicon generation

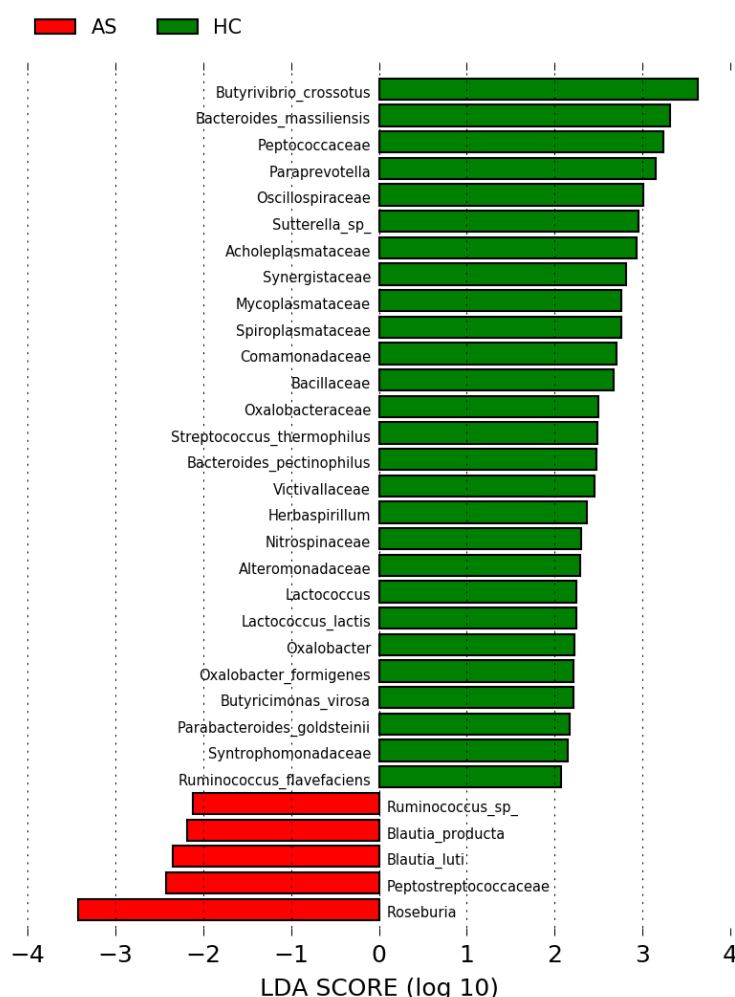


Fig. 2. Discriminant LEfSe analysis for identification of the taxa with the greatest differences in abundance between AS patients and HCs. Enriched taxa of AS group are indicated by a negative score (red) and enriched taxa of HC group by a positive score (green).

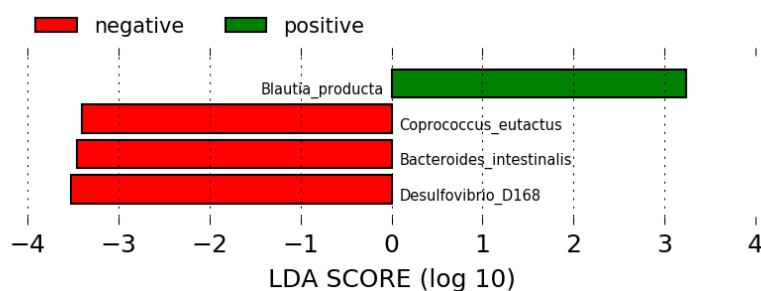


Fig. 3. Discriminant LEfSe analysis for identification of the taxa with the greatest differences in abundance between HLA B-27 positive patients and HLA B-27 negative patients. Taxa enriched in negative patients are indicated by negative score (red) and positive group's enriched taxa by a positive score (green).

which was reported to provide concordant results with home-brewed methods established for culture-free assessment of mixed bacterial populations (17). The two sets of primers of the kit covered 7 of the 9 variable regions of the 16S rRNA gene, generating amplicons which were recommended as most effective for short-read NGS (18, 19). The approach applied to the fecal specimens provided species-level taxonomic classification with few exceptions when OTUs were not assigned to a particular species among certain genera. An explanation for the certain degree of identification uncertainty can be the fact that the taxonomy assignment software QIIME used offers only one top-pick assignment eliminating the possibility to verify the 'second best' hit and make taxonomic decisions based on sequence data (20).

There are no standard protocols for sampling procedures of gut microbiota although a growing number of studies provides a view on the advantages and disadvantages of collection technologies (21). Currently, most of the studies rely on stool as proxies for gut microbiota even though it may rather be representative of the luminal microbiota of the colon and not the mucosal-associated bacterial communities (22). We also performed the microbiome analysis on stool samples because they were suitable for self-collection and easier to be handled by all the subjects enrolled in the study and we followed a collection protocol reported to be adequate for 16S RNA with sample storage at 4 °C for no more than 24 h before delivery at the lab and transport in an insulated container (23).

The human microbiota is known to display a significant degree of variation within and between individuals and the microbial community of the gut is one of the most complex in terms of species assemblages and functions, temporal, and spatial changes (24, 25). The data presented here should be viewed as a single-experiment based approach warranting replicating studies and also

a quantitative confirmation of what was inferred from the number of reads of the sequenced libraries in order to conclusively understand them, all data being made publicly available for those wishing to use them for hypothesis generation. Overall, our results indicated both similarities and differences between the taxonomic composition of fecal microbiomes from AS subjects and controls. At phylum level, both groups were dominated by *Firmicutes* and *Bacteroidetes*, the phyla known to normally prevail in the human gut, clades within them like *Bacteroides*, *Prevotella*, and *Ruminococcaceae* having been used to cluster individuals together into "enterotypes" to get an overview of species variation and functional composition (26). Considered less abundant and stable over time phylum than *Firmicutes* and *Bacteroidetes* and also as potentially indicative of a difficult ability to maintain a balance gut microbial, *Proteobacteria* was third in abundance in our study, exceeding 15% (27). Apparently, the expansion of *Proteobacteria* was due to the *Enterobacteriaceae* family in both diseased subjects and healthy controls. These findings deserve further investigations as literature reports support the pivotal role of *Enterobacteriaceae* and, in general, *Proteobacteria* in disease (28). Regarding differences observed at phylum level between AS and HC groups, it is worth mentioning the significant lower abundance of *Tenericutes* and *Synergistetes* in the diseased cohort, *Tenericutes* phylum being previously remarked as depleted in AS patients (29). Knowing that delineation of core microbiota is helpful in discrimination of the consistent members from those that may be associated only with specific states or restricted to specific environmental conditions (25), we also addressed the common bacterial core of the microbiota structure at lower taxonomic levels. As expected, we found that the microbial communities in gut samples broadly overlapped in the two groups in consistence with the existence of a preeminent

compositional phylogenetic core in the human gut pan-microbiome which may be essential for human health (30, 31). When pooling the AS and HC data sets, the collective core comprised *Bacteroidaceae*/ *Bacteroides*/ *Bacteroides vulgaris* and *Prevotellaceae*/ *Prevotella*/ *Prevotella copri* from *Bacteroidetes* phylum, and *Ruminococcaceae*/ *Faecalibacterium*/ *Faecalibacterium prausnitzii* from *Firmicutes* phylum as the most prevalent OTUs.

However, in terms of microbiome composition, a reduced microbiome diversity among the AS microbiomes was observed compared to the HC group, as indicated by both richness and evenness measures of alpha diversity which was in agreement with Cardoneanu et al. who also reported a decreased intestinal bacterial diversity in the Romanian subjects with AS compared to that noted in the control (11). Interestingly, by further using LEfSe, a microbiome-specific method (9), we evidenced marked differences in the abundance levels of certain taxa between the diseased and control cohorts as well. Some of the discriminative taxa for the AS group have been reported in previous studies. For instance, the increased abundance of *Blautia* and decreased abundance of *Ruminococcus* genera found by us in the Romanian AS samples by comparison with the controls' samples is in agreement with the data reported by a Chinese study which also used 16S barcoding and stool samples (32). Of note, in our study, one of the *Blautia* species with significantly higher abundance in the microbiomes of the AS, *B. producta* was linked to the HLA-B27-positive subset. Also, when focusing on the genera for differential abundance, although not ubiquitous enough to qualify for the core membership, *Dialister* was the genus with the largest effect size genus identified across the AS subjects. Tito et al. found that *Dialister* was increased in the microbiota of inflamed tissue from biopsies collected from the AS patients

and positively correlated with the disease score (5). Also, the microbiomes of the AS group were characterized by a preponderance of *Roseburia*, a finding also reported by Costello et al. in a study of microbiomes of terminal ileum biopsies from AS subjects (4).

In general, loss of the overall diversity, loss of commensal, beneficial organisms or excessive growth of potentially harmful ones, pathobionts, aspects not mutually exclusive, are criteria for diagnosing dysbiosis, a key term in the human microbiome field especially used to mediate the association of microbiome patterns to disease states that need to imply a fine-grained analysis (33, 34). Several studies focusing on the intestinal microbiome and how it relates to AS disease revealed alterations in the gut microbiome in patients with AS, but, to date, no distinct AS microbial signature has been identified (35). In spite of the tendency of greater pooled diversity of the HC group than of the AS group which further needs more refined consideration as some of the low abundant or not uniform across the cohort taxa may be a potential source of beneficial or pathogenic organisms, we have limited evidence of a significant difference between the gut microbiomes as a whole in AS patients and healthy individuals. Yet, our study has limitations that impaired the interpretation of the sequence data and must be addressed in the future. The main weaknesses were the lack of consistent metadata, including lifestyle and diet aspects, and the underrepresented control group with exclusively self-reported personal information that allowed for confounding factors which may have influenced the microbiome composition. The strength of this study is that it allowed for an informative high-throughput sequencing investigation of the gut microbiota needed to identify gut microbiome signatures of risk in the autochthonous population with AS.

Abbreviations

AS - Ankylosing Spondylitis
 NGS – Next-Generation Sequencing
 MHC - Major Histocompatibility Complex
 rRNA - 16S ribosomal RNA gene
 HC - Healthy Control
 OTU - Operational Taxonomic Unit
 QIIME - Quantitative Insights into Microbial Ecology
 PCoA - Principal-Coordinate analysis
 LDA - Linear Discriminant Analysis
 LefSe - LDA Effect Size

Acknowledgments

This work was funded through NUCLEU PN 19 14 02 02_47N/2019 and NUCLEU 16 39 01 04_46N/2016 from the Romanian Ministry of Research, Innovation, and Digitization. The funder had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Authors' contributions

Conceptualization: CRU, DC, MO; Supervision: CRU, MO; Formal analysis and investigation: CRU, MO, DC, SD, VCB, DP, ARB; Visualization: CRU, MO, SD; Writing – original draft: CRU, MO; Writing – review and editing: CRU, MO, SD, DC; Funding acquisition: MO, DC. All authors read and approved the final manuscript.

Conflict of Interest

The authors declare that they have no conflicts of interest.

References

1. Zhu W, He X, Cheng K, Zhang L, Chen D, Wang X, et al. Ankylosing spondylitis: etiology, pathogenesis, and treatments. *Bone Res.* 2019 7(22). DOI: 10.1038/s41413-019-0057-8
2. Wang R, Ward MM. Epidemiology of axial spondyloarthritis: an update. *Curr Opin Rheumatol.* 2018 March;30(2):137-43. DOI: 10.1097/BOR.0000000000000475
3. Ranganathan V, Gracey E, Brown MA, Inman RD, Haroon N. Pathogenesis of ankylosing spondylitis - recent advances and future directions. *Nat Rev Rheumatol.* 2017 Jun;13(6):359-67. DOI: 10.1038/nr-rheum.2017.56
4. Costello ME, Ciccio F, Willner D, Warrington N, Robinson PC, et al. Brief Report: Intestinal Dysbiosis in Ankylosing Spondylitis. *Arthritis Rheumatol.* 2015 Mar;67(3):686-91. DOI: 10.1002/art.38967
5. Tito RY, Cyper H, Joossens M, Varkas G, Van Praet L, Glorieu E, et al. Brief Report: Dialister as a Microbial Marker of Disease Activity in Spondyloarthritis. *Arthritis Rheumatol.* 2017 69(1):114-21. DOI: 10.1002/art.39802
6. Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, Stares MD, et al. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature.* 2016 May;533(7604):543-546. DOI: 10.1038/nature17645
7. Cullen CM, Aneja KK, Beyhan S, Cho CE, Woloszynek S, Convertino M, et al. Emerging Priorities for Microbiome Research. *Front Microbiol.* 2020 Feb;19(11):136. DOI: 10.3389/fmicb.2020.00136
8. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum.* 1984 27:361-8. DOI: 10.1002/art.1780270401
9. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011 Jun;12(6):R60. DOI: 10.1186/gb-2011-12-6-r60
10. Kashyap PC, Chia N, Nelson H, Segal E, Elinav E. Microbiome at the Frontier of Personalized Medicine. *Mayo Clin Proc.* 2017 Dec;92(12):1855-64. DOI: 10.1016/j.mayocp.2017.10.004
11. Cardoneanu A, Cozma S, Rezus C, Petrariu F, Burlui AM, Rezus E. Characteristics of the intestinal microbiome in ankylosing spondylitis. *Exp Ther Med.* 2021 Jul;22(1):676. DOI: 10.3892/etm.2021.10108
12. Pollock J, Glendinning L, Wisedchanwet T, Watson M. The madness of microbiome: attempting to find consensus "best practice" for 16S microbiome studies.

- Appl Environ Microbiol. 2018 84(7):e02627-17. DOI: 10.1128/AEM.02627-17
13. Bharti R, Grimm DG. Current challenges and best-practice protocols for microbiome analysis. *Brief Bioinform.* 2021 22(1):178-93. DOI: 10.1093/bib/bbz155
 14. Allali I, Arnold JW, Roach J, Cadenas MB, Butz N, Hassan HM, et al. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol.* 2017 Sep;17(1):194. DOI: 10.1186/s12866-017-1101-8
 15. Salipante SJ, Kawashima T, Rosenthal C, Hoogstraat DR, Cummings LA, Sengupta DJ, et al. Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Appl Environ Microbiol.* 2014 Dec;80(24):7583-91. DOI: 10.1128/AEM.02206-14
 16. Mas-Lloret J, Obón-Santacana M, Ibáñez-Sanz G, Guinó E, Pato ML, Rodríguez-Moranta F, et al. Gut microbiome diversity detected by high-coverage 16S and shotgun sequencing of paired stool and colon sample. *Sci Data.* 2020 Mar;7(1):92. DOI: 10.1038/s41597-020-0427-5
 17. Watts GS, Youens-Clark K, Slepian MJ, Wolk DM, Oshiro MM, Metzger GS, et al. 16S rRNA gene sequencing on a benchtop sequencer: accuracy for identification of clinically important bacteria. *J Appl Microbiol.* 2017 Dec;123(6):1584-96. DOI: 10.1111/jam.13590
 18. Soergel DA, Dey N, Knight R, Brenner SE. Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *ISME J.* 2012 Jul;6(7):1440-4. DOI: 10.1038/ismej.2011.208
 19. Mizrahi-Man O, Davenport ER, Gilad Y. Taxonomic classification of bacterial 16S rRNA genes using short sequencing reads: evaluation of effective study designs. *PLoS One.* 2013 8(1):e53608. DOI: 10.1371/journal.pone.0053608
 20. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods.* 2010 May;7(5):335-6. DOI: 10.1038/nmeth.f.303
 21. Tang Q, Jin G, Wang G, Liu T, Liu X, Wang B et al. Current Sampling Methods for Gut Microbiota: A Call for More Precise Devices. *Front Cell Infect Microbiol.* 2020 10:151. DOI: 10.3389/fcimb.2020.00151
 22. Lavelle A, Lennon G, O'Sullivan O, Docherty N, Balfe A, Maguire A, et al. Spatial variation of the colonic microbiota in patients with ulcerative colitis and control volunteers. *Gut.* 2015 Oct;64(10):1553-61. DOI: 10.1136/gutjnl-2014-307873
 23. Penington JS, Penno MAS, Ngui KM, Ajami NJ, Roth-Schulze AJ, Wilcox SA, et al. Influence of fecal collection conditions and 16S rRNA gene sequencing at two centers on human gut microbiota analysis. *Sci Rep.* 2018 Mar;8(1):4386. DOI: 10.1038/s41598-018-22491-7
 24. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. *Science.* 2009 Dec;326(5960):1694-7. DOI: 10.1126/science.1177486
 25. Berg G, Rybakova D, Fischer D, Cernava T, Vergès MC, Charles T, et al. Microbiome definition re-visited: old concepts and new challenges. *Microbiome.* 2020 Jun;8(1):103. DOI: 10.1186/s40168-020-00875-0
 26. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. *Nature.* 2011 May;473(7346):174-80. DOI: 10.1038/nature09944
 27. Shin NR, Whon TW, Bae JW. Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends Biotechnol.* 2015 Sep;33(9):496-503. DOI: 10.1016/j.tibtech.2015.06.011
 28. Baldelli V, Scaldaferrì F, Putignani L, Del Chierico F. The Role of Enterobacteriaceae in gut microbiota dysbiosis in inflammatory bowel diseases. *Microorganisms.* 2021 Mar;9(4):697. DOI: 10.3390/microorganisms9040697
 29. Chen Z, Qi J, Wei Q, Zheng X, Wu X, Li X, et al. Variations in gut microbial profiles in ankylosing spondylitis: disease phenotype-related dysbiosis. *Ann Transl Med.* 2019 Oct;7(20):571. DOI: 10.21037/atm.2019.09.41
 30. Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. *Science.* 2016 Apr;352(6285):560-4. DOI: 10.1126/science.aad3503
 31. Aguirre de Cárcer D. The human gut pan-microbiome presents a compositional core formed by discrete phylogenetic units. *Sci Rep.* 2018 Sep;8(1):14069. DOI: 10.1038/s41598-018-32221-8
 32. Zhang L, Han R, Zhang X, Fang G, Chen J, Li J, et al. Fecal microbiota in patients with ankylosing spondylitis: Correlation with dietary factors and disease activity. *Clin Chim Acta.* 2019 Oct;497:189-96. DOI: 10.1016/j.cca.2019.07.038

33. Hooks KB, O'Malley MA. Dysbiosis and its discontents. *mBio*. 2017 Oct;8(5):e01492-17. DOI: 10.1128/mBio.01492-17
34. Levy M, Kolodziejczyk AA, Thaïs CA, Elinav E. Dysbiosis and the immune system. *Nat Rev Immunol*. 2017 Apr;17(4):219-32. DOI: 10.1038/nri.2017.7
35. Harkins P, Burke E, Swales C, Silman A. 'All disease begins in the gut'-the role of the intestinal microbiome in ankylosing spondylitis. *Rheumatol Adv Pract*. 2021 Sep;5(3):rkab063. DOI: 10.1093/rap/rkab063

