
Gut microbiota analysis in systemic sclerosis according to disease characteristics and nutritional status

G. Natalello^{1,2}, S.L. Bosello^{1,2}, F. Paroni Sterbini^{3,4}, B. Posteraro^{5,6},
E. De Lorenzis^{1,2}, G.B. Canestrari^{1,2}, L. Gigante^{1,2}, L. Verardi^{1,2},
G. Ferraccioli^{1,2}, M. Sanguinetti^{3,4}, E. Gremese^{1,2}

¹Institute of Rheumatology, Università Cattolica del Sacro Cuore, Rome;

²Division of Rheumatology, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome;

³Institute of Microbiology, Università Cattolica del Sacro Cuore, Rome;

⁴Dipartimento di Scienze di Laboratorio e Infettivologiche, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome;

⁵Institute of Medical Pathology and Semeiotics, Università Cattolica del Sacro Cuore, Rome;

⁶Dipartimento di Scienze Gastroenterologiche, Endocrino-Metaboliche e Nefro-Urologiche, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy.

Gerlando Natalello, MD

Silvia Laura Bosello, MD, PhD

Francesco Paroni Sterbini, MSc

Brunella Posteraro, MD, Prof.

Enrico De Lorenzis, MD

Giovanni Battista Canestrari, MD

Laura Gigante, MD

Lucrezia Verardi, MD

Gianfranco Ferraccioli, MD, Prof.

Maurizio Sanguinetti, MD, Prof.

Elisa Gremese, MD, Prof.

Please address correspondence to:

Elisa Gremese,

Divisione di Reumatologia

Fondazione Policlinico Universitario

A. Gemelli IRCCS,

Presidio Columbus di Roma,

Università Cattolica del Sacro Cuore,

Via Giuseppe Moscati 31,

00168 Rome, Italy.

E-mail: elisa.gremese@unicatt.it

Received on March 16, 2020; accepted in revised form on June 12, 2020.

Clin Exp Rheumatol 2020; 38 (Suppl. 125): S73-S84.

© Copyright CLINICAL AND

EXPERIMENTAL RHEUMATOLOGY 2020.

Key words: systemic sclerosis, gut microbiota, nutritional status, 16S ribosomal RNA, microbial diversity

Competing interests: none declared.

ABSTRACT

Objective. Systemic sclerosis (SSc) is a rare multi-organ disorder with a prominent gastrointestinal (GI) involvement. Altered gut microbiota is now considered a pivotal factor associated with the development of immune-mediated and inflammatory diseases. We performed a 16S ribosomal RNA (rRNA) gene-sequencing analysis of fecal microbiota in a cohort of SSc patients and matched healthy controls (HCs), with the aim to obtain some hints about a possible role of dysbiosis in the onset, progression, and severity of the disease.

Methods. We analysed stool samples from 63 SSc patients with different disease duration, phenotype, and nutritional status and from 17 HCs through 16S ribosomal RNA (rRNA) gene-sequencing.

Results. Microbial richness was lower for patients with long-standing disease. A similar observation was made for patients with diffuse cutaneous SSc (dsSSc) compared to those with limited variant (lcSSc) and for patients who reported a recent weight loss. Consistent with previous reports, we noted a deviation of the intestinal microbial composition in patients with SSc compared to HCs, with a greater expression of *Lactobacillus* and *Streptococcus* and a depletion of *Sutterella*. Nutritional status, assessed using BMI as a surrogate, appeared to have a marked impact on the gut microbiota, with overweight patients showing lower richness compared both to underweight and normal-BMI patients.

Conclusion. Our findings expand the current knowledge of gut microbiota in SSc and could be useful to identify patients who would most benefit from treatments aimed at restoring the eubiosis.

Introduction

Systemic sclerosis (SSc) is a rare multi-organ disease whose pathological hallmarks are autoimmunity, vasculopathy, and fibrosis (1-3). The involvement of the gastrointestinal (GI) tract is frequent, affecting up to 90% of the patients with SSc, and virtually extending from the mouth to the anus (4, 5). GI symptoms are the presenting feature of the disease in about 10% of individuals and severe GI involvement is recognised as a major cause of morbidity and mortality in SSc, especially when associated with alteration of gut function that leads to intestinal bacterial overgrowth and malnutrition (6, 7). The scarcity of predictive factors of GI disease and valid assessment tools, partly related to the very heterogeneous manifestations, represents an unmet clinical need in SSc (8).

A model of damage characterised by progression from initial vascular and neural disorder, up to muscle dysfunction and fibrosis, is generally accepted but the exact etiopathogenesis of GI involvement in SSc is still elusive (9, 10). This is reflected in the lack of treatment strategies capable to impact on the progression of GI damage, with the majority of the established therapeutic efforts aimed at most at symptomatic relief (3). Recent progress in high-throughput and bioinformatics allowed to deeply explore the human microbiome. Many pieces of evidence suggest that the highly diverse and evolving microbes which inhabit the human gut and other mucosal interfaces might be implicated in both the homeostasis maintenance and the induction and progression of autoimmune diseases, by cross-talking with innate and adaptive immunity at multiple levels (11-14). Altered gut microbiota is now considered a pivotal

factor associated with the development and the activity of immune-mediated inflammatory diseases, such as inflammatory bowel diseases (IBDs) and spondyloarthritis (15-17), and new studies suggest an influence of microbial communities in the onset and modulation of the aberrant immune response in diseases like rheumatoid arthritis (RA) (18, 19) and systemic lupus erythematosus (SLE) (20, 21). Nevertheless, plausible mechanistic studies are relatively scarce in the field of rheumatic diseases and mostly limited to evaluating a single potential pathobiont rather than examining an interaction model that considers whole microbial communities (14).

Given the complexity of SSc, a pathology with multifactorial aetiology that frequently implies a subversion of the intestinal barrier, the interaction at the gut-lumen interface between the microbiota and the host immune and neuroendocrine systems is likely to have some influence on the disease onset and phenotypic expression.

Observational studies in SSc patients reported a high prevalence of dysbiosis (22) and described specific profiles in the gut microbial composition of either the colonic mucosal-luminal interface or the luminal content (faecal microbiota) (23-25). Increases in both the taxa associated with other inflammatory conditions and, surprisingly, genera typically considered beneficial such as *Lactobacillus* have been reported in small cohorts. Other studies tried to shed light on the functional implication of these alterations both evaluating patients with clinically evident gastrointestinal involvement and integrating plasma metabolomic data in relation to the microbial profile (26-28). However, due to the small sample size of most of the previous studies and considered the clinical heterogeneity of SSc, relatively few data are available about the possible variations of the microbial communities according to the different disease characteristics and the nutritional status of the patients.

Here we performed a 16S ribosomal RNA (rRNA) gene-sequencing analysis of fecal microbiota in a relatively large cohort of Italian patients, with the aim

of strengthening the current knowledge of gut microbial communities in SSc and to obtain some hints about a possible role of dysbiosis in the onset, progression, and severity of the disease. In particular, the objectives of the present study were: i) to elucidate if the characteristics of fecal microbiota in SSc patients may change according to the disease duration, ii) to define qualitative and quantitative differences in fecal microbiota composition between our SSc patients and matched healthy controls (HCs), iii) to investigate the variations of microbial communities across different subsets of the disease and according to the nutritional status.

Materials and methods

Study cohort

Eligible participants included Caucasian adults from the geographical region of central Italy, able to provide written informed consent. Pre-fixed exclusion criteria for all participants (SSc patients and HCs) were: i) comorbid inflammatory bowel disease or celiac disease, ii) presence of gastrointestinal tract (GIT) lesions such as ulcers and cancers, iii) history of GIT surgery, iv) treatment with antimicrobial or probiotic medications within one month prior to the fecal sample collection, v) known lactose intolerance, vi) alternative diet such as vegetarian or gluten-free diet, and vii) pregnancy. Patients who fulfilled the 2013 ACR/EULAR classification criteria for SSc (29), admitted to our Scleroderma Clinic from June 2015 to June 2017, were considered. Additional exclusion criterion for the SSc patients was a major therapeutic change (immunosuppressive and vasoactive drugs) within 2 months prior to the faecal sample collection, due to the progression of the disease. All participating subjects signed informed consent.

Thus, 20 consecutive SSc patients with a disease duration less than or equal to 3 years since the first non-Raynaud's phenomenon (non-RP) manifestation (early SSc) and 20 consecutive scleroderma patients with disease duration longer than 3 years (long-standing SSc) were enrolled. Twenty HCs with comparable mean age and sex ratio to all SSc patients served as a control group. The samples

of HCs were selected from our institutional Microbiology depository. These subjects were required to have a normal body mass index (BMI) according to the World Health Organization (18.5 to 25.0 kg/m²), not to have any known significant disease, not to take chronic therapy, not to have used probiotics or antibiotics in the recent period and not to practice alternative diet regimens.

Additionally, 14 consecutive overweight (OW) (BMI, >25.0 kg/m²) and 14 underweight (UW) SSc patients (BMI, ≤18.5 kg/m²) were also enrolled. For all participating subjects, a comprehensive medical history and a pharmacological and dietary anamnesis were collected. Weight and height were measured to calculate the BMI. All patients were characterised with regard to disease features and organ involvement. Disease duration was defined both in years from the onset of RP and in years from the first non-RP manifestation. The severity of cutaneous involvement was assessed with the modified Rodnan Skin Score (mRSS) by a trained physician (SLB) and patients were classified in limited (lcSSc) and diffuse (dcSSc) subsets according to LeRoy classification (30). The presence of SSc-related interstitial lung disease (ILD) and pulmonary arterial hypertension (PAH) was determined on the basis of historical diagnostic findings on high-resolution chest CT and right heart catheterisation, respectively. For all patients, we required to be available at least one instrumental determination for forced vital capacity (FVC), carbon monoxide diffusing capacity (DLCO) and pulmonary artery systolic pressure (PASP) in the previous 12 months and at least one historical determination for the presence of antinuclear antibodies (ANA), which was determined by indirect immunofluorescence using Hep-2 cells as substrates, and of autoantibodies against extractable nuclear antigens (ENAs) specificities, which was determined by ELISA (Shield, Dundee, UK). Positivity for anti-topoisomerase I (ATA) and anti-centromere (ACA) antibodies was registered.

All participants were instructed to collect a fresh stool in a standard sterile container and to keep the sample frozen

until reaching the clinic. Faecal samples were promptly frozen at -80°C upon arrival at our microbiology laboratory. On the same day of stool collection, each patient completed the Italian version of the UCLA SCTC GIT 2.0 questionnaire. The questionnaire consists of 34 items and 7 multi-item scales (reflux, distension/bloating, diarrhea, fecal soiling, constipation, emotional well-being, and social functioning) and has been shown to be a valid measure of GI tract symptoms severity in SSc patients (31, 32). The five-step Malnutrition Universal Screening Tool (MUST) was also performed to assess nutritional risk in patients (6, 33) and a blood sample was collected to measure levels of haemoglobin, glucose, albumin, ferritin, total cholesterol, HDL cholesterol, triglycerides, 25-hydroxyvitamin D, vitamin B12, folate according to normal practice of our institution clinical laboratory.

DNA isolation and

16S rRNA gene pyrosequencing

After thawing, 0.25 grams of each stool sample was homogenised using a mini-bead beater (Biospec Products, Bartlesville, OK, USA) at maximum speed for 1 minute to enhance the yield of DNA to be subjected to PCR amplification, particularly from robust bacteria (34). DNA was extracted from the samples by using the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany) according to manufacturer's protocol, and DNA concentrations were measured by a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Minneapolis, MN). Multi-tag pyrosequencing of the 16S rRNA gene V1-V3 variable regions was performed using the GS Titanium technology (Roche 454 Life Sciences, Branford, CT, USA) and the 454 adaptor-linked 28F (5'-TTTGATCNTG-GCTCAG) and 519R (5'-GTNTTAC-NGCGGCKGCTG) primers, as described elsewhere (35-38). The amplification reaction was prepared in 50 μL final volume containing 0.4 μM of each primer, 200 μM of each deoxynucleoside triphosphate, 1.8 μM of MgCl_2 , 5 μL of 10 \times PCR buffer, 50 ng of template DNA and 2.5 U of FastStart High-Fidelity Enzyme Blend (Roche). The PCR conditions were as follows: 95°C

for 2 minutes, followed by 35 cycles of 95°C for 20 seconds, 56°C for 45 seconds, and 72°C for 5 minutes, with final extension step at 72°C for 7 minutes. For each PCR reaction, one additional PCR negative control without DNA template was added. PCR products were purified with the Agencourt AMPure XP Kit (Beckman Coulter, Milan, Italy), visually inspected on agarose gels to observe bands of the specific size (approximately 600 bp) and quantified with the Quant-iT PicoGreen double-stranded DNA Assay Kit (Life Technologies, Monza MB, Italy). After clean-up and quality control, amplicons from all the samples were normalised, pooled, purified, and then unidirectionally sequenced on the 454 GS Junior platform using a GS Titanium Sequencing Kit according to the manufacturer's instructions.

Sequence data were processed in QIIME (Quantitative Insights into Microbiology Ecology) 1.8.0 (39). Sequence reads were demultiplexed and quality-filtered according to default parameters, and pyrosequencing errors were removed using the PyroNoise algorithm (40). The reads were then sorted and grouped into operational taxonomic units (OTUs) using the UCLUST algorithm (<http://www.drive5.com/usearch/>) at a distance-based similarity of 97% (41). Potentially chimeric sequences were removed using the UCHIME algorithm (42). The OTU representatives were aligned with PyNAST and assigned to different taxonomic levels (from phylum to species) using the UCLUST consensus taxonomy classifier and the Greengenes taxonomy reference database (version 13.8.0), with a sequence identity threshold of 97% (43, 44). Sequences that did not match any existing bacterial sequences in the database were pooled together as "unassigned". The phylogenetic tree was built with FastTree (45). Based on rarefaction tests, samples with a sampling depth of fewer than 739 sequences were excluded from downstream analysis, in order to control for differences in coverage and to limit the effects of uneven sampling. A complete list of samples with the number of reads can be found in the Supplementary file, Table S1.

Microbiome analysis and statistics

The analysis of 16S rRNA data was carried out through a comprehensive tool based on the backend MicrobiomeAnalystR package for R (46). For the purposes of removing low quality and likely uninformative features, *a priori* data filtering was set, including only features with at least 2 counts in at least 10% of samples. Data were rarefied to minimum library size and normalised through the Total Sum Scaling method. The profiling of microbial communities consisted in estimates of alpha diversity (within-sample diversity), which was measured by calculating the richness (Observed OTUs and Chao1) and evenness (Simpson and Shannon) indices at the OTU level, and estimates of beta-diversity (between-samples diversity), which was measured by calculating phylogenetic (UniFrac (47)) and non-phylogenetic (Bray-Curtis (48)) distances at the OTU level. Statistical comparison of alpha diversity between samples was performed using a *t*-test, analysis of variance (ANOVA) or a nonparametric test, as appropriate; beta-diversity matrices were visualised through the principle coordinate analysis (PCoA) ordination method and statistical significance verified using the permutational ANOVA test. In order to identify potential biomarkers related to disease state and disease subgroups, differentially abundant features were identified using the linear discriminant analysis (LDA) effect size (LEfSe) method (49) or the EdgeR algorithm (50), considering meaningful an adjusted $p < 0.1$ with a LDA score of at least 2.0 and an adjusted $p < 0.05$, respectively. Random Forests learning algorithm (based on an ensemble of classification trees) was also executed for identification and classification of features predictive of disease status. General data were analysed using SPSS Statistics v. 24.0 for Windows (IBM Corp. Released 2016. Armonk, NY). Continuous variables were reported as mean \pm SD while categorical variables were reported as numbers and percentages. $p > 0.05$ was considered non-significant (NS).

Table I. Demographic and clinical characteristics of early and long-standing SSc and HCs.

	Early SSc n=19	Long-standing SSc n=20	<i>p</i>	All SSc (normal BMI n=39)	Healthy Controls) n=17	<i>p</i>
Age, years, mean±SD	52.7 ± 16.7	52.4 ± 11.6	NS	52.5 ± 14.2	50.8 ± 14.3	NS
Female, n (%)	15 (78.9)	18 (90.0)	NS	33 (84.6)	12 (70.6)	NS
Body mass index, kg/m ² , mean±SD	22.2 ± 11.6	22.7 ± 1.7	NS	22.5 ± 1.7	22.6 ± 1.8	NS
Disease duration, years ^a , mean±SD	2.3 ± 2.0	10.8 ± 4.1	<.001	6.6 ± 5.4	-	-
ATA / ACA, n (%)	11 (57.9) / 2 (10.5)	10 (50.0) / 6 (30.0)	NS/NS	21 (53.8) / 8 (20.5)	-	-
Diffuse cutaneous SSc, n (%)	9 (47.4)	12 (60.0)	NS	21 (53.8)	-	-
Modified Rodnan skin score, mean±SD	9.8 ± 7.2	9.8 ± 8.7	NS	9.8 ± 8.8	-	-
Digital ulcers (ever), n (%)	7 (36.8)	10 (50.0)	NS	17 (43.6)	-	-
Interstitial lung disease (HRCT), n (%)	8 (42.1)	11 (55.0)	NS	19 (48.7)	-	-
FVC, %, mean±SD	92.6 ± 22.5	91.5 ± 25.9	NS	92.0 ± 26.1	-	-
DLCO, %, mean±SD	59.2 ± 19.3	57.0 ± 19.7	NS	58.1 ± 20.1	-	-
Pulmonary arterial hypertension, n (%)	2 (10.5)	4 (20.0)	NS	6 (15.4)	-	-
PASP, mmHg, mean±SD	33.2 ± 7.8	30.9 ± 7.6	NS	32.0 ± 7.7	-	-
UCLA SCTC GIT 2.0						
Total score, mean±SD	0.40 ± 0.24	0.40 ± 0.34	NS	0.40 ± 0.29	-	-
Reflux, mean±SD	0.56 ± 0.38	0.65 ± 0.78	NS	0.61 ± 0.61	-	-
Bloating, mean±SD	0.96 ± 0.79	0.78 ± 0.74	NS	0.87 ± 0.76	-	-
Fecal soilage, mean±SD	0.11 ± 0.32	0.35 ± 0.81	NS	0.23 ± 0.63	-	-
Diarrhea, mean±SD	0.26 ± 0.39	0.18 ± 0.41	NS	0.22 ± 0.39	-	-
Social functioning, mean±SD	0.24 ± 0.22	0.29 ± 0.34	NS	0.27 ± 0.28	-	-
Emotional well-being, mean±SD	0.28 ± 0.61	0.14 ± 0.23	NS	0.21 ± 0.45	-	-
Constipation, mean±SD	0.47 ± 0.46	0.54 ± 0.79	NS	0.51 ± 0.64	-	-
Moderate/severe GI symptoms ^b , n (%)	3 (15.8)	5 (25.0)	NS	8 (20.5)	-	-
MUST, score ≥1, n (%)	3 (15.8)	3 (15.0)	NS	6 (15.4)	-	-

SD: standard deviation; ATA: anti-topoisomerase I antibodies; ACA: anti-centromere antibodies; HRCT: high resolution computed tomography; ^afrom first non-Raynaud's phenomenon symptom; ^baccording to UCLA SCTC GIT 2.0 Total score (32).

Results

Data from a total of seven samples (1 early SSc, 1 underweight SSc, 3 overweight SSc, 2 HCs) were removed since they did not achieve the minimum number of read sequences considered to be adequate for downstream analysis. One further sample (HC) was not analyzed due to a technical problem that occurred in the DNA extraction phase. Overall, samples from 63 scleroderma patients and 17 healthy controls were considered. The characteristics of the subjects whose samples were included are shown in Table I (subjects with normal BMI) and Suppl. Table S2 (summary of treatments and blood chemistry parameters of subjects with normal BMI) and Suppl. Table S3 (underweight and overweight patients). Considering the patients with normal BMI, lcSSc and dcSSc were represented similarly in early and long-standing subgroups. Overall, these patients had a slightly higher than usually reported prevalence of ATA positivity (3), without significant difference between subgroups. No significant differences between early and long-standing patients

were noted about major organ involvement, history of digital ulcers and severity of reported GI symptoms. Sex ratio, age, and BMI were homogeneous. Almost all patients were on proton pump inhibitors (PPI) therapy (89.7%). Other common medications included iloprost (56.4%) and calcium-channel blockers (66.7%). Sixteen patients (41.0%) were on a stable dose of immunosuppressive drugs (mycophenolate mofetil, azathioprine, methotrexate), with or without low-dose oral glucocorticoids, for their internal organ involvement. A total of 11 patients (28.2%) were exposed to cyclophosphamide during their history. It should be noted that in this latter group of patients with normal BMI, the risk of malnutrition according to MUST was identified exclusively with an unexpected weight loss equal to or greater than 5% of body weight in the previous 3–6 months. Regarding patients with abnormal BMI, the two subgroups (UW and OW) were homogeneous for age, time from disease onset, distribution of dcSSc and ATA/ACA positivity. Subjects were all females except for one UW male

patient. Also, in this sub-cohort, the UCLA SCTC GIT 2.0 questionnaire did not bring out any significant difference in reported GI symptoms. UW patients showed significantly lower mean levels of albumin, fasting glucose and total cholesterol. Since a BMI <18.5 kg/m² is one of the determinants of MUST scoring, all UW subjects were considered at risk of malnutrition. Only 3 patients in the UW group had a MUST score ≥2, indicating further risk factors (mainly a recent and unexpected weight loss) other than BMI itself.

Gut microbial communities profiling

Initially, the microbial composition of samples from early and long-standing SSc patients was compared. These two groups exhibited a significant difference in alpha diversity according to Observed OTUs ($p=0.014$) and Chao1 index ($p=0.009$), with samples from early SSc patients showing greater richness compared to those with long-standing disease. Also, evenness measured by the Shannon index was significantly different ($p=0.043$) with early SSc accounting for higher values.

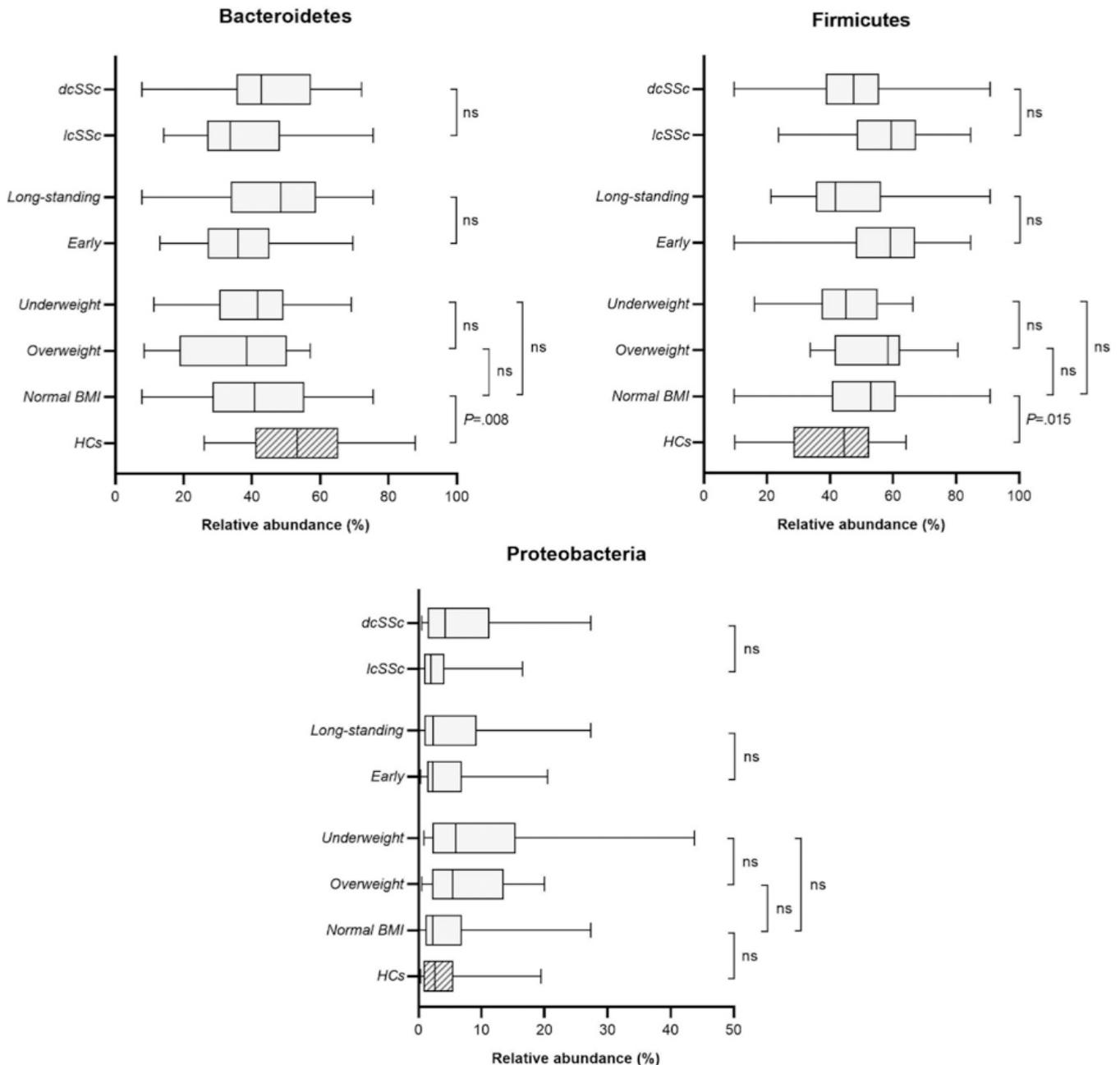


Fig. 1. Plots of relative abundances of the three most frequent bacterial *phyla* in HCs and SSc subgroups. Significant differences (*t*-test) are noted. The middle line in the box plot represents the median value, and the box is drawn from 25% to 75% quartiles; whiskers show the minimum and maximum values.

No meaningful difference was noted by the Simpson index ($p=0.061$). Beta diversity was then computed, showing a significant difference both according to Bray Curtis ($p=0.023$, $R^2=0.04$) and weighted Unifrac ($p=0.002$, $R^2=0.076$) distances, with apparently only a small part of the variation explained.

The microbiota of early and long-standing SSc as a whole group (SSc patients with normal BMI) was then compared to that from matched HCs. *Bacteroidetes* and *Firmicutes* represented the

two most abundant *phyla* in both SSc patients and HCs, accounting for 93.5% and 94.5% of all sequences, respectively. Interestingly, the ratio between *Bacteroidetes* and *Firmicutes* was inverted in the two groups, with the latter phylum predominant in samples from SSc patients. In univariate analysis, the difference in mean abundance was significant both for *Bacteroidetes* ($p=.008$) and *Firmicutes* ($p=0.015$). *Proteobacteria* represented the third most abundant phylum in both SSc patients and

HCs, accounting for 5.5% and 4.4% of all sequences, respectively. No significant difference was noted in mean abundance ($p=0.4$ in univariate analysis). Plots of relative abundance with the comparisons between subgroups are shown in Figure 1.

The samples from SSc exhibited similar alpha diversity based on the Chao1 ($p=0.25$), Observed OTUs ($p=0.79$), Shannon ($p=0.32$) and Simpson ($p=0.15$) indices compared to HCs, while beta diversity was significantly

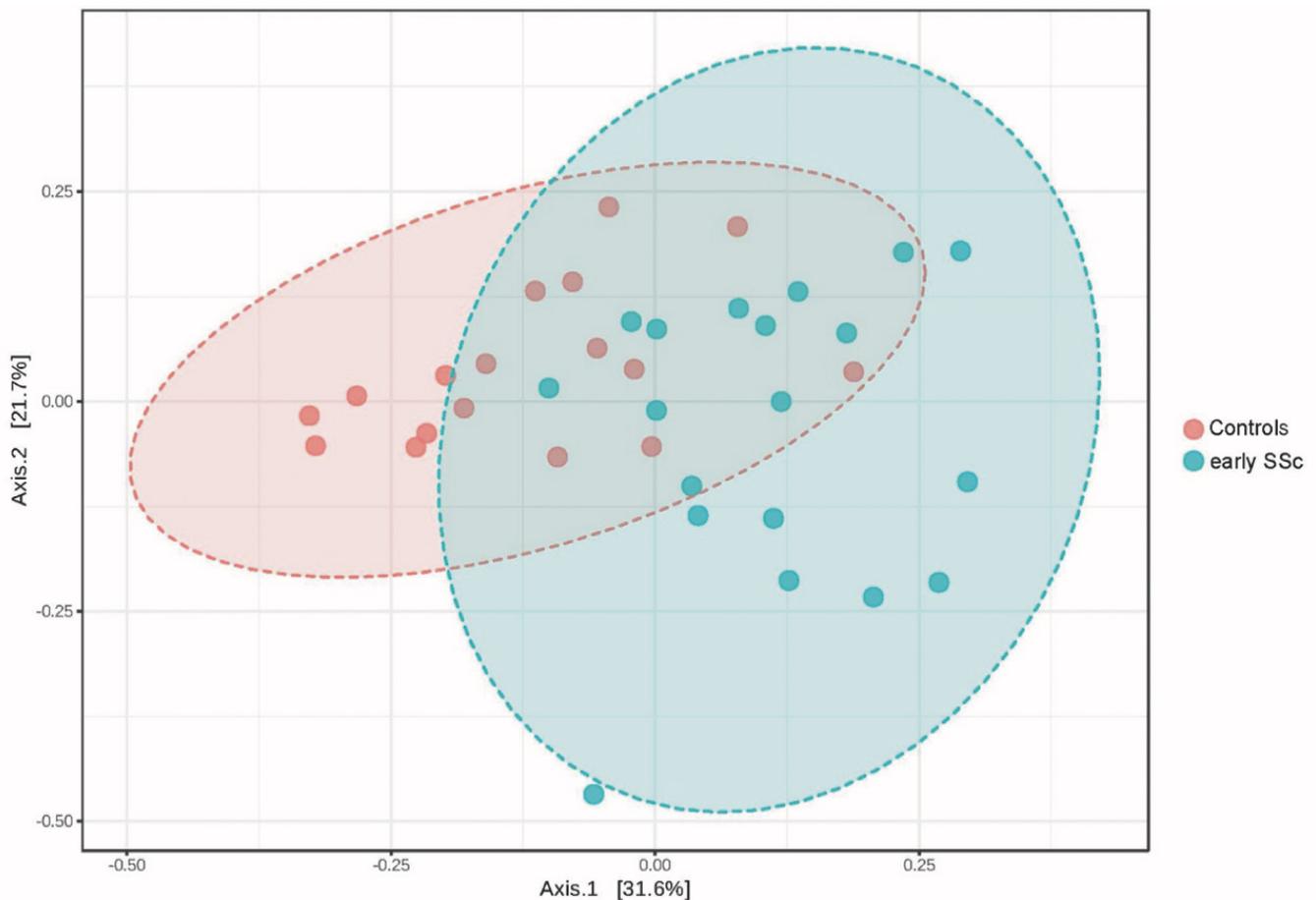


Fig. 2. Plots of PCoA ordinated beta diversity values (first two axes) of HCs compared to early SSc patients, according to weighted UniFrac ($p < 0.001$, $R^2 = 0.165$).

different when computed both with Bray-Curtis ($p < 0.001$, $R^2 = 0.047$) and weighted UniFrac ($p = 0.003$, $R^2 = 0.068$) distances. Furthermore, samples from early SSc and long-standing SSc group were compared separately against HCs, both maintaining most of the above-mentioned differences. Interestingly, when considering only early SSc data, the magnitude of difference in beta-diversity increased noticeably according to weighted UniFrac ($p < 0.001$, $R^2 = 0.165$) (Fig. 2), while long-standing SSc alone showed a significantly lower richness according to the Chao1 ($p = 0.023$) index and the number of Observed OTUs ($p = 0.029$) that was previously inapparent. The summary of the results is reported in Supplementary Table S4.

Gut microbial composition differences according to SSc patients' subsets

Given the clinical and prognostic significance of the extension of the skin involvement in SSc (LeRoy classifica-

tion), microbial communities from samples of lcSSc and dcSSc patients were compared. A summary of general and disease characteristics distribution according to this classification can be found in Supplementary Table S5. A lower richness was present in samples from dcSSc compared with lcSSc based on the Chao1 index ($p = 0.004$) and Observed OTUs ($p = 0.004$) and a lower evenness based on Shannon ($p = 0.03$) index, but not on the Simpson one ($p = 0.11$). Beta diversity was non-different between dcSSc and lcSSc according to Bray-Curtis ($p = 0.12$, $R^2 = 0.033$) and weighted UniFrac ($p = 0.14$, $R^2 = 0.038$) distances.

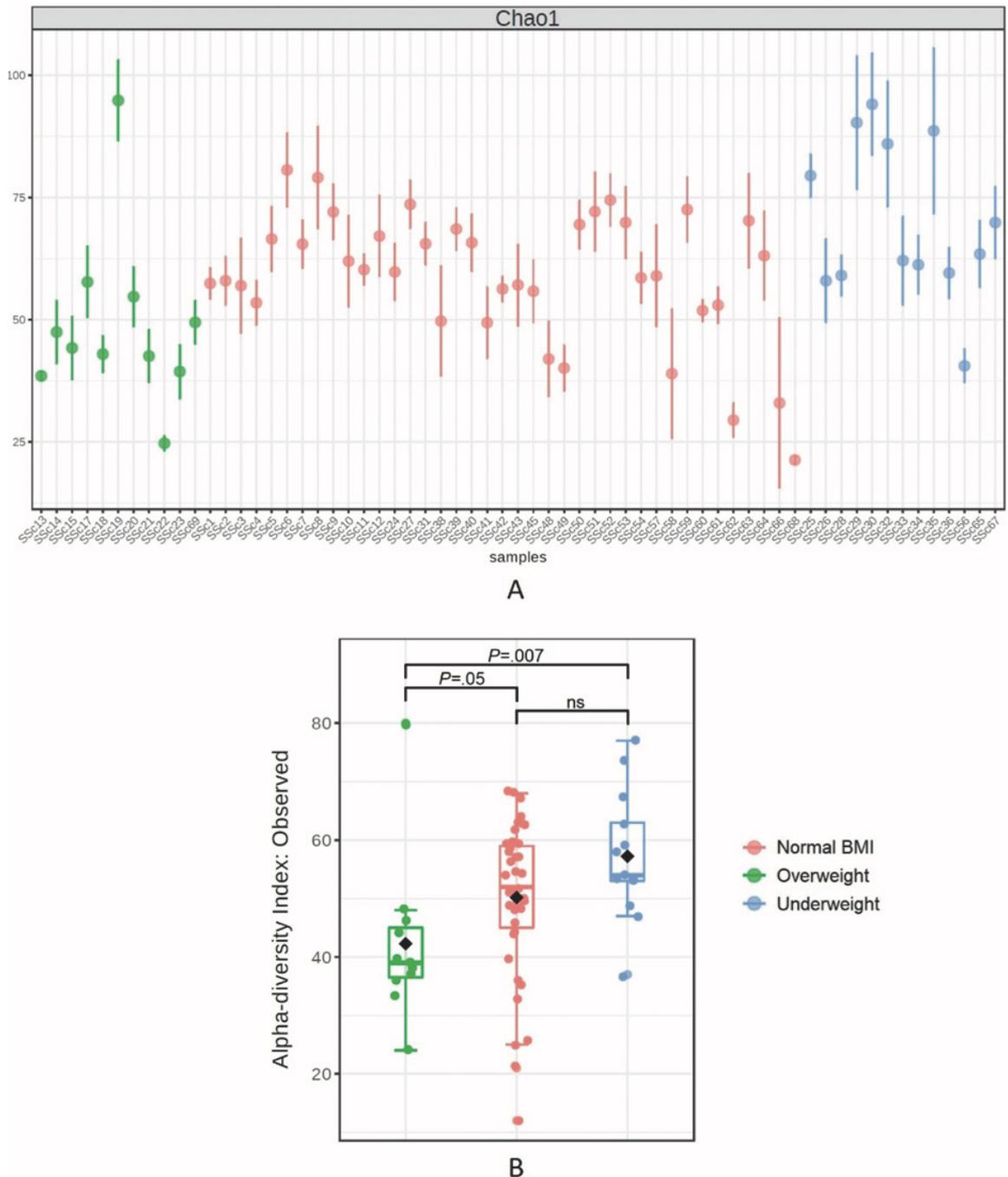
Comparisons of microbiota according to SSc-specific autoantibodies (ATA vs. ACA vs. others), presence of major disease involvement (SSc-ILD, SSc-PAH, history of SSc digital ulcers), GI symptoms by UCLA SCTC GIT 2.0 total score (none to mild vs. moderate to severe symptoms) and malnutrition risk according to MUST were also per-

formed for SSc patients with normal BMI. Interestingly, patients at risk for malnutrition showed lower richness according to both Chao1 ($p = 0.042$) and the number of Observed OTUs ($p = 0.035$). No significant differences emerged from alpha and beta diversity analysis among other subgroups. The summary of the results can be found in Supplementary Table S6.

Gut microbial composition differences according to patients' BMI

Data from UW and OW SSc patients were considered along with those from the normal BMI SSc group. Alpha diversity was significantly different according to richness measures, with the OW group showing the lower values and the UW group the higher ones both according to Chao1 (Kruskal-Wallis $p = 0.002$) and Observed OTUs (Kruskal-Wallis $p = 0.009$). However, this difference was mainly supported by the characteristics of UW and OW microbial communities, as the normal

Fig. 3. Plots of alpha diversity values of BMI subgroups according to Chao1 estimator (Kruskal-Wallis $p=0.002$) (A) and observed OTUs (Kruskal-Wallis $p=0.009$) (B). For the latter, p for multiple comparisons are noted.



BMI subgroup showed less evident differences with respect to the other two subgroups in multiple comparisons. A graphical summary is shown in Figure 3. OW patients showed also a lower evenness according to the Shannon index ($p=0.046$), with the comparison by Simpson method remaining non-significant ($p=0.12$). Regarding beta diversity, a significant difference emerged according to the Bray-Curtis distance ($p=0.037$, $R^2=0.044$) but not to the weighted UniFrac one ($p=0.42$,

$R^2=0.032$), with the magnitude of the former slightly increasing when considering only samples from UW and OW patients ($p=0.003$, $R^2=0.072$).

Biomarkers discovery

To identify specific bacterial genera accountable for differences in microbiota between SSc patients and matched HCs, we applied the LEfSe method. In samples from SSc patients, significantly greater expression of *Streptococcus* ($Q=0.005$, $LDA\ score=5.2$), *Lacto-*

bacillus ($Q=0.025$, $LDA\ score=5.2$), *Blautia* ($Q=0.025$, $LDA\ score=4.9$), *Ruminococcus* ($Q=0.067$, $LDA\ score=4.8$) and *Phascolarctobacterium* ($Q=0.067$, $LDA\ score=4.8$) genera was found, compared to HCs. Furthermore, a depletion of *Sutterella* ($Q=0.005$, $LDA\ score=-4.9$), *Bacteroides* ($Q=0.025$, $LDA\ score=-5.8$), *Odoribacter* ($Q=0.036$, $LDA\ score=-3.6$) and *Roseburia* ($Q=0.067$, $LDA\ score=-4.0$) genera was noted. Analysis through the random forests algorithm confirmed a

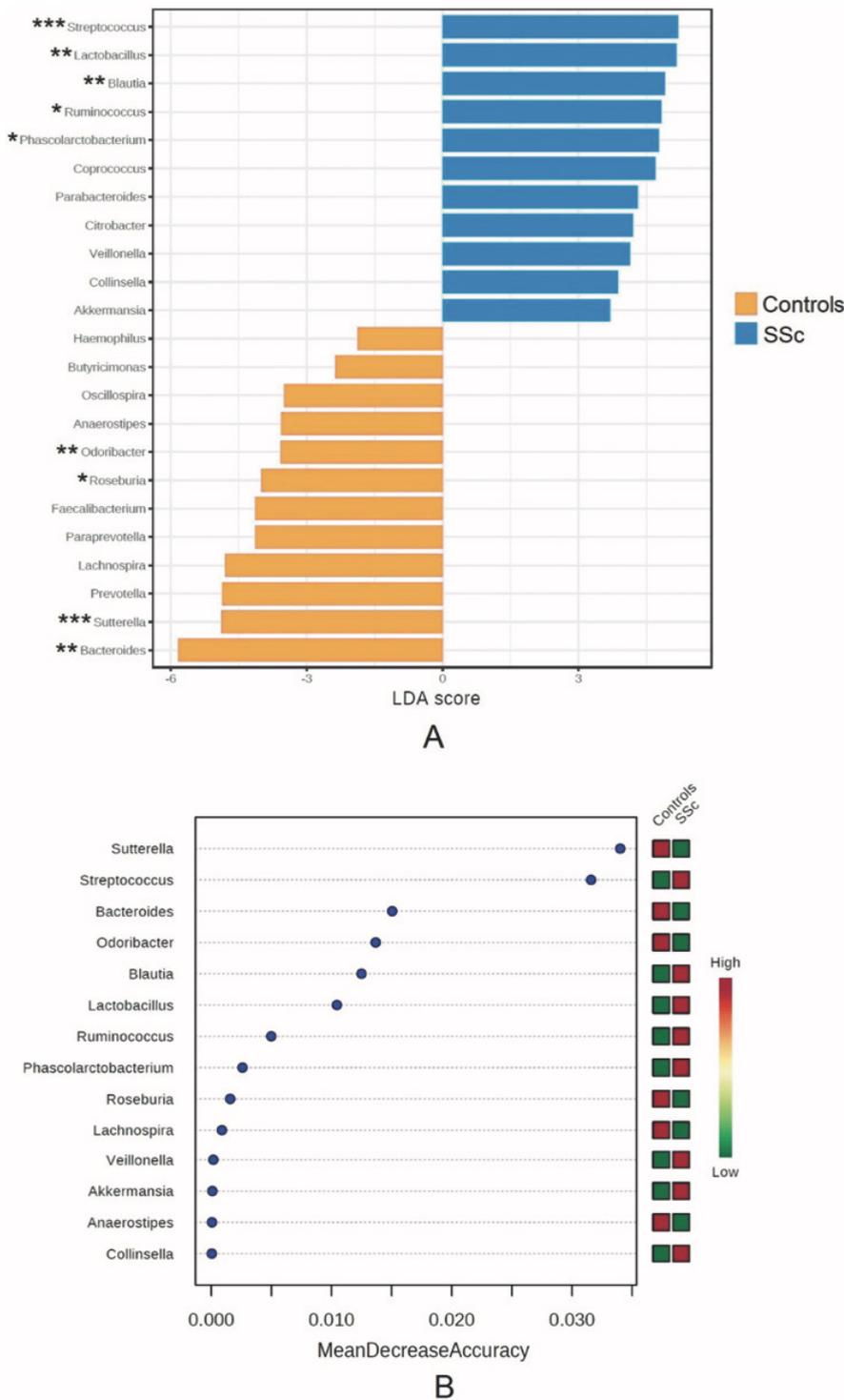


Fig. 4. Plots of top different features at genus level according to Linear Discriminant Analysis Effect Size (* $Q < 0.1$; ** $Q < 0.05$; *** $Q < 0.01$, in blue increased genera in SSc, in orange decreased genera in SSc) (A) and of top features for classification accuracy according to Random Plot Forest model (nTree=1000, OOB error=0.196) (B).

good power of the above-mentioned genera in discriminating the two groups, with an out-of-bag error (OOB) of 0.196. In particular, *Streptococcus* and *Sutterella* showed higher values for classification accuracy for SSc and

HCs, respectively. Plots of top discriminating features according to LEfSe and classification accuracy computed with random forests are shown in Figure 4. The report of cumulative error rates and classification performance of

random forests can be found in Supplementary Figure S1.

Subsequently, LEfSe multivariate analysis against HCs was repeated considering the individual SSc subgroups that had shown greater variability, namely early SSc, long-standing SSc, lcSSc, and dcSSc, in order to verify the validity of discriminating genera previously found. Overall, *Streptococcus* and *Lactobacillus* showed to be constantly increased and *Sutterella* to be constantly depleted through SSc subgroups compared to HCs, thus representing the strongest biomarkers of disease status. The summary of the results is shown in Table II.

LEfSe comparison between the main subgroups was also performed. Long-standing SSc patients showed no significant differences at the genus level compared to those with early disease, while in dcSSc patients a depletion of *Coprococcus* ($Q = 0.041$, $LDA\ score = -4.8$) was noted compared to those with lcSSc.

Furthermore, we applied the multivariate EdgeR method to identify differentially abundant features at different taxonomic levels based on the risk of malnutrition according to MUST score (at risk if score ≥ 1) and on the severity of lower gastrointestinal tract symptoms (categorised in none to mild vs. moderate to severe according to the single lower-GI items of the UCLA SCTC GIT 2.0 questionnaire). Notably, *Akkermansia muciniphila* appeared to be associated with more severe diarrhea and fecal soiling, while *Prevotella copri* was more abundant in samples from patients reporting more severe fecal soiling and recent weight loss but depleted in those reporting greater constipation. The summary of the results is shown in Table III.

Discussion

To our knowledge, the present is the largest cohort of SSc patients (63) whose fecal microbiota was characterised through a full 16S rRNA gene-sequencing protocol. In a case series of 98 SSc patients, the authors adopted a dysbiosis test based on a predetermined set of probes, limited to 54 bacterial rRNA specificities, therefore not intended to capture all the diversity of the samples

Table II. Summary of differentially expressed genera between subgroups patients and HCs based on relative abundances computed with LEfSe. Positive LDA score indicates genera more expressed in samples from SSc patients; negative LDA scores indicate depleted genera in samples from SSc patients.

Genus	All SSc (normal BMI) vs. HCs		Long-standing SSc vs. HCs		Early SSc vs. HCs		leSSc vs. HCs		dcSSc vs. HCs	
	LDA score	Q	LDA score	Q	LDA score	Q	LDA score	Q	LDA score	Q
<i>Bacteroides</i>	-5.8	0.025	-5.6	NS	-6.0	0.002	-5.8	0.094	-5.9	0.044
<i>Blautia</i>	4.9	0.025	4.6	NS	5.0	0.004	4.6	0.094	5.1	0.003
<i>Lactobacillus</i> ^a	5.2	0.025	4.9	0.041	5.2	0.019	5.2	0.051	3.2	0.097
<i>Odoribacter</i>	-3.6	0.036	-4.1	0.08	3.9	NS	-4.2	0.013	4.2	NS
<i>Phascolarctobacterium</i>	4.8	0.067	4.7	NS	4.8	NS	4.8	NS	4.8	NS
<i>Roseburia</i>	-4.0	0.067	-3.8	NS	-3.9	NS	-3.9	NS	-3.9	NS
<i>Ruminococcus</i>	4.8	0.067	5.1	NS	4.5	NS	-4.0	NS	5.2	0.022
<i>Streptococcus</i> ^a	5.2	0.005	5.1	0.041	5.4	0.002	5.1	0.013	5.2	0.003
<i>Sutterella</i> ^b	-4.9	0.005	-4.8	0.041	-5.0	0.002	-4.9	0.014	-4.9	0.035

^aConsistently increased genera in SSc patients; ^bconsistently depleted genera in SSc patients.

Table III. Summary of differentially abundant taxa, according to symptoms severity. The features significantly associated with more severe symptoms in each domain of UCLA SCTC GIT 2.0 (lower GI) and with the risk of malnutrition according to MUST are reported (no features were individuated for the distension/bloating domain, hence is not reported). Log2FC values indicate the fold-change of the specific feature abundance, with a minus sign indicating a decreased abundance. *Note that in this group of patients with normal BMI, the risk of malnutrition according to MUST has been identified exclusively with an unexpected weight loss equal to or greater than 5% of body weight in the previous 3-6 months.

	Taxon	Log2FC	Q
Faecal soiling			
Increased	<i>Akkermansia muciniphila</i> (s)	1.83	0.003
	<i>Prevotella copri</i> (s)	3.29	0.003
Diarrhoea			
Increased	<i>Akkermansia muciniphila</i> (s)	1.64	0.039
Constipation			
Increased	Enterobacteriaceae (f)	3.23	<.001
Decreased	<i>Prevotella copri</i> (s)	-3.7	0.02
Risk of malnutrition*			
Increased	<i>Prevotella copri</i> (s)	3.29	0.002

(22). Furthermore, we first evaluated overweight and underweight patients, using BMI measurement and evaluation through a widely validated malnutrition assessment tool as surrogates of nutritional status, to get some hints on how it interacts with the disease state, at gut microbiota level.

The presence of a group of HCs, with similar age and sex ratio and of the same ethnic and geographical extraction, allowed us to verify any deviations in microbial composition in our cohort of normal weight SSc patients. Given the complex interaction between intestinal bacteria and multiple host-intrinsic and host-extrinsic factors, it is necessary to bear in mind that it is unfeasible to eliminate all confound-

ing factors for this type of study design. In order to limit this issue, we adopted stringent inclusion criteria and precise characterisation of SSc patients. The pharmacological exposure remained a potential source of bias, given its heterogeneity and the consequent impossibility of performing meaningful subgroup analyses. However, the distribution of immunosuppressive, glucocorticoids and gastroprotective drugs was homogeneous between the main subgroups of SSc, and we included only patients in stable therapy to avoid the possible negative impact of newly introduced treatments on GI symptoms. We described a shift in microbiota characteristics between patients with early and long-standing disease. For the pur-

pose of this study, we decided to define early SSc as patients who fulfilled the latest ACR/EULAR classification criteria and with disease duration ≤ 3 years from first non-Raynaud's phenomenon symptom. This definition may not match with other definitions of early scleroderma in the literature (51, 52) as it was not intended to include patients with very early diagnosis of SSc or pre-scleroderma but to include patients with a well-defined disease yet with reduced aspects of chronicity and accumulated organ damage. The most remarkable difference we noted was a reduction in the richness and evenness of microbial communities in patients with long-standing disease. Notably, the reduction in alpha diversity is a common find in inflammatory and immune-mediated diseases such as IBDs (53, 54), SLE (55), and RA (19), in the latter case also correlating with disease duration. Although our data seem to suggest some degree of longitudinal variation with respect to the onset of disease, the cross-sectional nature of the present study is not suitable to establish this aspect with certainty. In a recent abstract, Volkman and colleagues reported stability of absolute and relative abundance of specific genera in stool specimens, over 12 months (56), but no precise information on the duration of disease at baseline are reported. In general, our data are in agreement with these observations, as the early and long-standing patients of our cohort did not show any significant differences regarding the distribution of the main genera. Indeed,

the major differences were highlighted with regard to alpha diversity measures (within-sample diversity).

A reduction in the richness was also observed in patients with dcSSc compared to those with lcSSc, and in patients at malnutrition risk according to MUST, for a recent unexpected weight loss equal to or greater than 5% of body weight. Weight loss in SSc is a poor prognostic factor, usually implying severe underlying intestinal disease and presumably a higher degree of fibrosis of the gut (57). Taken together, these findings suggest a relationship between lower microbial diversity and the degree of fibrotic burden. In the case of dcSSc, a further determinant of the difference with respect to lcSSc was found in the significant reduction of the *Coprococcus* genus. Notably, the observation of *Coprococcus* depletion in the context of reduced gut microbial richness has identified a subgroup of patients with non-alcoholic fatty liver disease (NAFLD) with a tendency to persistent activation of pro-inflammatory metabolic pathways and progression to liver fibrosis (58). On this basis, future studies could complement known measures of intestinal permeability as plasma zonulin (59) in patients with SSc, in the attempt to define a mechanistic link between intestinal microbiota, inflammation and fibrosis.

No major differences were observed with respect to the overall GI symptoms reported by the patients. In part, this could be due to the intrinsic limits of the evaluation method for GI symptoms, based on a questionnaire. In this regard, it will be useful to validate microbial communities' analysis against instrumental approaches able to identify and more strictly quantify aspects such as dysmotility and bacterial overgrowth. However, in-depth evaluation through multivariate analysis, considering specifically the questionnaire items for lower GI involvement, highlighted the association between particular microbial species and a more serious GI symptomatology. In particular, both *Prevotella copri* and *Akkermansia muciniphila*, two species implicated in pro-inflammatory state and fibrotic changes in other diseases (60-62), were

associated with increased severity of symptoms related to the lower GI tract. These last observations must be taken with caution given the reduced size of the sample and the predisposition of the statistical method to generate type I errors. Still, our data agree with those already reported in separate cohorts about the association of *Prevotella copri* (23) and *Akkermansia* (28) with worse GI symptoms, suggesting the possibility of identifying, within a complex dysbiotic pattern, single microbial biomarkers for lower GI tract clinical manifestations.

The comparison between patients with SSc and HCs has allowed us to identify some genera characterising the disease status. In particular, those constantly increased, were found to be *Lactobacillus* and *Streptococcus*, accompanied by a constant depletion of *Sutterella*. The latter is a member of Proteobacteria phylum, considered a widely prevalent commensal (63). Strikingly, the increased expression of the *Lactobacillus* genus is a common observation in several cohorts of SSc patients with different geographical origins and with different sample types (22, 23, 25, 28). This evidence suggests the need for a re-evaluation of the role of *Lactobacillus*, classically considered a beneficial commensal genus and inversely related to inflammatory states (64), in view of the peculiar dysbiotic mosaic of SSc. For instance, we cannot exclude that the increased abundance of *Lactobacillus*, could represent the effort to maintain a healthier intestinal interface when the disease modifies gut homeostasis. In general, the LEfSe multivariate analysis based on the relative abundances has shown highly overlapping results to what reported in two different SSc cohorts (Los Angeles and Oslo) for fecal samples (23). At the phylum level, SSc patients showed inversion of the ratio between two predominant phyla, compared to HCs. Specifically, Firmicutes, to which the aforementioned *Streptococcus* genus belongs, were significantly more represented in SSc patients, prevailing on Bacteroidetes. Notably, the deviation of this ratio could have a greater impact on the metabolism and human health (65, 66) and has been previously described in SSc (25, 28).

With regard to patients with altered nutritional status, the most evident data was the progressive reduction of the richness going from lean to overweight patients. It should be emphasised that UW patients did not show significantly more GI symptoms than OW in our cohort. Furthermore, only a minority showed additional risk factors of malnutrition in addition to the underweight itself. This would suggest, limited to this cohort, only a small influence of the disease on BMI, justifying an overlap with what is observed in the general population (67, 68). Ultimately, caution is needed in generalising observations when major factors, closely related to the metabolic state and dietary habits, such as BMI overlap with SSc state in determining the composition of the intestinal microbiota. To this regard, one of the limitations of this study is represented by the lack of direct measures targeted at the metabolic profiling of the intestinal environment. For instance, the evaluation of fecal short-chain fatty acids, as a product of commensal bacteria and one of the main substrates for the colonic epithelium (69), would help to shed light on the significance of specific patterns of microbial communities in presence of major deviations in body fat composition and deserve future studies.

In general, our data seem to suggest a model of dysbiosis characterised by the presence *ab initio* of specific genera peculiar of SSc. Progressively, the impairment of normal intestinal biodiversity could increase the interaction of these characteristic genera with immune and neuroendocrine systems of the intestine, especially in the more aggressive phenotypes of the disease, characterised by a higher degree of fibrosis and impairment of the intestinal wall. Indeed, alpha diversity seems altered toward a lower richness and evenness values in patients whose clinical characteristics are linked to a worse prognosis, namely the diffuse cutaneous variant and the risk of malnutrition due to abnormal weight loss. Whether the changes in the characteristics of the microbiota are primary or secondary with respect to the local and systemic immune aberrations is yet to be established.

The data acquired so far support the possibility of a therapeutic intervention aimed at re-establishing a state as close as possible to the eubiosis. Noteworthy, the negative results of two recent placebo-controlled trials (70, 71) of multistrain probiotic to ameliorate GI symptoms in SSc, requires the identification of new therapeutic strategies to impact intestinal flora. Given the complex nature of the relationship between intestinal microbial composition and SSc, with deviations both from a phylogenetic and quantitative point of view, it would be advisable to privilege interventions with a broad capacity to affect microbial communities, such as diet and fecal microbiota transplantation (FMT) (72, 73), rather than the mere administration of probiotic agents. Recently, a small pilot trial of FMT (cultivated single-donor type) in patients with SSc has shown encouraging signals regarding the ability to modify the characteristics of the microbiota and efficacy on reported symptoms for the lower digestive tract over 16 weeks (74). This kind of approach certainly deserves further development to establish the best modalities for administration and preparation of both donors and recipients (75). At the same time, potentially serious adverse events related to infectious risk and the invasive route of administration require greater efforts in patient selection. In this context, the preliminary evaluation of the gut microbial composition, based on the research of precise microbiological biomarkers, could allow identifying in a targeted way the patients who could most benefit from such interventions.

Conclusions

With the present study, we evaluated the faecal microbiota in a relatively large cohort of patients with SSc, a disease characterised by a prominent gastrointestinal involvement. Our data suggest a possible interaction between characteristics of the microbial communities and disease phenotype. In particular, a reduced microbial richness seems to be associated with conditions linked to a worse prognosis, such as the diffuse cutaneous variant and a precarious nutritional state with weight loss.

This alteration was also noted for patients with longer disease duration and needs to be confirmed with a longitudinal study involving patients with very early SSc. Furthermore, we observed that even in scleroderma patients the basal nutritional status appears to have a marked influence on the gut microbial balance. Consistently with previous studies, we noted a deviation of the intestinal microbial composition in patients with SSc compared to healthy subjects, identifying a potential microbiological signature that could be used to select patients who would most benefit from treatments aimed at restoring the eubiosis.

References

- DENTON CP, KHANNA D: Systemic sclerosis. *Lancet* 2017; 390: 1685-99.
- GABRIELLI A, AVVEDIMENTO EV, KRIEG T: Scleroderma. *N Engl J Med* 2009; 360: 1989-2003.
- BARSOTTI S, ORLANDI M, CODULLO V *et al.*: One year in review 2019: systemic sclerosis. *Clin Exp Rheumatol* 2019; 37 (Suppl. 119): S3-14.
- JAEGER VK, WIRZ EG, ALLANORE Y *et al.*: Incidences and risk factors of organ manifestations in the early course of systemic sclerosis: a longitudinal EUSTAR study. *PLoS One* 2016; 11: e0163894.
- MILLER JB, GANDHI N, CLARKE J, MCMAHAN Z: Gastrointestinal involvement in systemic sclerosis: an update. *J Clin Rheumatol* 2018; 24: 328-37.
- BARON M, HUDSON M, STEELE R *et al.*: Malnutrition is common in systemic sclerosis: results from the Canadian scleroderma research group database. *J Rheumatol* 2009; 36: 2737-43.
- DOMSIC RT, NIHTYANOVA SI, WISNIEWSKI SR *et al.*: Derivation and external validation of a prediction rule for five-year mortality in patients with early diffuse cutaneous systemic sclerosis. *Arthritis Rheumatol* 2016; 68: 993-1003.
- EMMANUEL A: Current management of the gastrointestinal complications of systemic sclerosis. *Nat Rev Gastroenterol Hepatol* 2016; 13: 461-72.
- McFARLANE IM, BHAMRA MS, KREPS A *et al.*: Gastrointestinal manifestations of systemic sclerosis. *Rheumatology* (Sunnyvale) 2018; 8: 235.
- SJOGREN RW: Gastrointestinal motility disorders in scleroderma. *Arthritis Rheum* 1994; 37: 1265-82.
- FORBES JD, VAN DOMSELAAR G, BERNSTEIN CN: The gut microbiota in immune-mediated inflammatory diseases. *Front Microbiol* 2016; 7: 1081.
- WEINSTOCK GM: Genomic approaches to studying the human microbiota. *Nature* 2012; 489: 250-6.
- XU J, MAHOWALD MA, LEY RE *et al.*: Evolution of symbiotic bacteria in the distal human intestine. *PLoS Biol* 2007; 5: e156.
- KONIG MF: The microbiome in autoimmune rheumatic disease. *Best Pract Res Clin Rheumatol* 2020 Feb 7.
- RIZZO A, FERRANTE A, GUGGINO G, CICCIA F: Gut inflammation in spondyloarthritis. *Best Pract Res Clin Rheumatol* 2017; 31: 863-76.
- MANICHANH C, BORRUEL N, CASELLAS F, GUARNER F: The gut microbiota in IBD. *Nat Rev Gastroenterol Hepatol* 2012; 9: 599-608.
- BELKAID Y, HAND TW: Role of the microbiota in immunity and inflammation. *Cell* 2014; 157: 121-41.
- KALINKOVICH A, LIVSHITS G: A cross talk between dysbiosis and gut-associated immune system governs the development of inflammatory arthropathies. *Semin Arthritis Rheum* 2019; 49: 474-84.
- CHEN J, WRIGHT K, DAVIS JM *et al.*: An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome Med* 2016; 8: 43.
- RUIZ L, LÓPEZ P, SUÁREZ A, SÁNCHEZ B, MARGOLLES A: The role of gut microbiota in lupus: what we know in 2018? *Expert Rev Clin Immunol* 2018; 14: 787-92.
- HEVIA A, MILANI C, LÓPEZ P *et al.*: Intestinal dysbiosis associated with systemic lupus erythematosus. *MBio* 2014; 5: e01548-14.
- ANDREASSON K, ALRAWI Z, PERSSON A, JÖNSSON G, MARSAL J: Intestinal dysbiosis is common in systemic sclerosis and associated with gastrointestinal and extraintestinal features of disease. *Arthritis Res Ther* 2016; 18: 278.
- VOLKMANN ER, HOFFMANN-VOLD A-M, CHANG Y-L *et al.*: Systemic sclerosis is associated with specific alterations in gastrointestinal microbiota in two independent cohorts. *BMJ Open Gastroenterol* 2017; 4: e000134.
- VOLKMANN ER: Intestinal microbiome in scleroderma: recent progress. *Curr Opin Rheumatol* 2017; 29: 553-60.
- VOLKMANN ER, CHANG Y-L, BARROSO N *et al.*: Association of systemic sclerosis with a unique colonic microbial consortium. *Arthritis Rheumatol* 2016; 68: 1483-92.
- BELLOCCHI C, VOLKMANN ER: Update on the Gastrointestinal Microbiome in Systemic Sclerosis. *Curr Rheumatol Rep* 2018; 20: 49.
- BELLOCCHI C, FERNÁNDEZ-OCHOA Á, MONTANELLI G *et al.*: Microbial and metabolic multi-omic correlations in systemic sclerosis patients. *Ann NY Acad Sci* 2018; 1421: 97-109.
- PATRONE V, PUGLISI E, CARDINALI M *et al.*: Gut microbiota profile in systemic sclerosis patients with and without clinical evidence of gastrointestinal involvement. *Sci Rep* 2017; 7: 14874.
- VAN DEN HOOGEN F, KHANNA D, FRANSEN J *et al.*: 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. *Ann Rheum Dis* 2013; 72: 1747-55.
- LEROY EC, BLACK C, FLEISCHMAJER R *et al.*: Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol* 1988; 15: 202-5.

31. GUALTIEROTTI R, INEGNOLI F, TWO R *et al.*: Reliability and validity of the Italian version of the UCLA Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument in patients with systemic sclerosis. *Clin Exp Rheumatol* 2015; 33 (Suppl. 91): S55-60.
32. KHANNA D, HAYS RD, MARANIAN P *et al.*: Reliability and validity of the University of California, Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. *Arthritis Rheum* 2009; 61: 1257-63.
33. ELIA M: The 'MUST' report. Nutritional screening for adults: a multidisciplinary responsibility. Development and use of the 'Malnutrition Universal Screening Tool' (MUST) for adults. 2003.
34. YANG I, NELL S, SUERBAUM S: Survival in hostile territory: the microbiota of the stomach. *FEMS Microbiol Rev* 2013.; 37: 736-61.
35. BINLADEN J, GILBERT MTP, BOLLEBACK JP *et al.*: The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS One* 2007; 2: e197.
36. HAMADY M, WALKER JJ, HARRIS JK, GOLD NJ, KNIGHT R: Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* 2008; 5: 235-7.
37. PARONI STERBINI F, PALLADINI A, MASUCCI L *et al.*: Effects of proton pump inhibitors on the gastric mucosa-associated microbiota in dyspeptic patients. *Appl Environ Microbiol* 2016; 82: 6633-44.
38. SOERGEL DAW, DEY N, KNIGHT R, BRENNER SE: Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *ISME J* 2012; 6: 1440-4.
39. CAPORASO JG, KUCZYNSKI J, STOMBAUGH J *et al.*: QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010; 7: 335-6.
40. REEDER J, KNIGHT R: Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. *Nat Methods* 2010; 7: 668-9.
41. GEVERS D, COHAN FM, LAWRENCE JG *et al.*: Opinion: Re-evaluating prokaryotic species. *Nat Rev Microbiol* 2005; 3: 733-9.
42. EDGAR RC, HAAS BJ, CLEMENTE JC, QUINCE C, KNIGHT R: UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011; 27: 2194-200.
43. CAPORASO JG, BITTINGER K, BUSHMAN FD, DESANTIS TZ, ANDERSEN GL, KNIGHT R: PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 2010; 26: 266-7.
44. EDGAR RC: Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; 26: 2460-1.
45. PRICE MN, DEHAL PS, ARKIN AP: FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 2010; 5: e9490.
46. DHARIWAL A, CHONG J, HABIB S, KING IL, AGELLON LB, XIA J: MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res* 2017; 45: W180-88.
47. LOZUPONE C, HAMADY M, KNIGHT R: UniFrac—an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* 2006; 7: 371.
48. BEALS EW: Bray-Curtis Ordination: an effective strategy for analysis of multivariate ecological data, in advances in ecological research volume 14. 1984, 1-55.
49. SEGATA N, IZARD J, WALDRON L *et al.*: Metagenomic biomarker discovery and explanation. *Genome Biol* 2011; 12: R60.
50. ANDERS S, MCCARTHY DJ, CHEN Y *et al.*: Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc* 2013; 8: 1765-86.
51. BELLANDO-RANDONE S, MATUCCI-CERINIC M: Very early systemic sclerosis and pre-systemic sclerosis: definition, recognition, clinical relevance and future directions. *Curr Rheumatol Rep* 2017; 19: 65.
52. VALENTINI G, MARCOCCIA A, CUOMO G *et al.*: Early systemic sclerosis: marker autoantibodies and videocapillaroscopy patterns are each associated with distinct clinical, functional and cellular activation markers. *Arthritis Res Ther* 2013; 15: R63.
53. MANICHANH C, RIGOTTIER-GOIS L, BONNAUD E *et al.*: Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 2006; 55: 205-11.
54. LI J, BUTCHER J, MACK D, STINTZI A: Functional impacts of the intestinal microbiome in the pathogenesis of inflammatory bowel disease. *Inflamm Bowel Dis* 2015; 21: 139-53.
55. AZZOUC D, OMARBEKOVA A, HEGUY A *et al.*: Lupus nephritis is linked to disease-activity associated expansions and immunity to a gut commensal. *Ann Rheum Dis* 2019; 78: 947-56.
56. VOLKMANN EP, HOFFMANN-VOLD A-M, CHANG Y-L *et al.*: OP0085 - Longitudinal analysis of the gastrointestinal microbiota in systemic sclerosis. *Ann Rheum Dis* 2017; 76 (Suppl. 2): 87.
57. STEEN VD, MEDSGER TA, JR: Severe organ involvement in systemic sclerosis with diffuse scleroderma. *Arthritis Rheum* 2000; 43: 2437-44.
58. KIM HN, JOO EJ, CHEONG HS *et al.*: Gut microbiota and risk of persistent nonalcoholic fatty liver diseases. *J Clin Med* 2019; 8: 1089.
59. TAJIK N, FRECH M, SCHULZ O *et al.*: Targeting zonulin and intestinal epithelial barrier function to prevent onset of arthritis. *Nat Commun* 2020; 11: 1995.
60. SCHER JU, SCZESNAK A, LONGMAN RS *et al.*: Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *Elife* 2013; 2: e01202.
61. MAEDA Y, KURAKAWA T, UMEMOTO E *et al.*: Dysbiosis contributes to arthritis development via activation of autoreactive T cells in the intestine. *Arthritis Rheumatol* 2016; 68: 2646-61.
62. KUGATHASAN S, DENSON LA, WALTERS TD *et al.*: Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* 2017; 389: 1710-8.
63. HIIPPALA K, KAINULAINEN V, KALLIOMÄKI M, ARKKILA P, SATOKARI R: Mucosal prevalence and interactions with the epithelium indicate commensalism of *Sutterella* spp. *Front Microbiol* 2016; 7: 1706.
64. WALTER J: Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. *Appl Environ Microbiol* 2008; 74: 4985-96.
65. MARIAT D, FIRMESSE O, LEVENEZ F *et al.*: The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiol* 2009; 9: 123.
66. LEY RE, TURNBAUGH PJ, KLEIN S, GORDON JI: Microbial ecology: human gut microbes associated with obesity. *Nature* 2006; 444: 1022-3.
67. PETERS BA, SHAPIRO JA, CHURCH TR *et al.*: A taxonomic signature of obesity in a large study of American adults. *Sci Rep* 2018; 8: 9749.
68. LE CHATELIER E, NIELSEN T, QIN J *et al.*: Richness of human gut microbiome correlates with metabolic markers. *Nature* 2013; 500: 541-6.
69. DE LA CUESTA-ZULUAGA J, MUELLER NT, ÁLVAREZ-QUINTERO R *et al.*: Higher fecal short-chain fatty acid levels are associated with gut microbiome dysbiosis, obesity, hypertension and cardiometabolic disease risk factors. *Nutrients* 2018; 11: 51.
70. MARIGHELA TF, ARISMENDI MI, MARVULLE V, COLÓ BRUNIALTI MK, SALOMÃO R, KAYSER C: Effect of probiotics on gastrointestinal symptoms and immune parameters in systemic sclerosis: a randomized placebo-controlled trial. *Rheumatology (Oxford)* 2019; 58: 1985-90.
71. LOW AHL, TENG GG, PETERSSON S *et al.*: A double-blind randomized placebo-controlled trial of probiotics in systemic sclerosis associated gastrointestinal disease. *Semin Arthritis Rheum* 2019; 49: 411-9.
72. LOMBARDI VC, DE MEIRLEIR KL, SUBRAMANIAN K *et al.*: Nutritional modulation of the intestinal microbiota; future opportunities for the prevention and treatment of neuroimmune and neuroinflammatory disease. *J Nutr Biochem* 2018; 61: 1-16.
73. VOLKMANN ER, HOFFMANN-VOLD AM: Gastrointestinal tract microbiota modifications in systemic sclerosis. *Eur J Rheumatol* 2019 Dec 19.
74. FRETHEIM H, CHUNG BK, DIDRIKSEN H *et al.*: Fecal microbiota transplantation in systemic sclerosis: A double-blind, placebo-controlled randomized pilot trial. *PLoS One* 2020; 15: e0232739.
75. HOFFMANN-VOLD AM, FRETHEIM H, DIDRIKSEN H, MOLBERG Ø: The potential of fecal microbiota transplantation in systemic sclerosis. *Expert Rev Clin Immunol* 2020; 16: 117-8.