Research Article

Disordered intestinal microbes are associated with the activity of Systemic Lupus Erythematosus

Yao Li1,2,*, Hai-Fang Wang1,*, Xin Li1,*, Hai-Xia Li1, Qiong Zhang1, Hong-Wei Zhou3, Yan He3, Pan Li3, Chen Fu1, Xiao-He Zhang1, Yu-Rong Qiu1,4 and Ji-Liang Li2,5,6

1Laboratory Medicine Center, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, P.R. China; 2Institute of Antibody Engineering, School of Laboratory Medicine and Biotechnology, Southern Medical University, Guangzhou 510515, Guangdong, P.R. China; 3Division of Laboratory Medicine, Microbiome Medicine Center, Zhujiang Hospital, Southern Medical University, Guangzhou 510282, Guangdong, P.R. China; 4Huayin Medical Laboratory Center Co., Ltd., Guangzhou 510515, Guangdong, P.R. China; 5Wenzhou Medical University School of Biomedical Engineering and Eye Hospital, Wenzhou Institute of Biomaterials and Engineering, Wenzhou 325035, Zhejiang, P.R. China; 6Institute of Translational and Stratified Medicine, Plymouth University Faculty of Medicine and Dentistry, Plymouth PL6 8BU, U.K.

Correspondence: Yurong Qiu (qyr@smu.edu.cn) or Ji-Liang Li (Ji-liang.li@plymouth.ac.uk)

Intestinal dysbiosis is implicated in Systemic Lupus Erythematosus (SLE). However, the evidence of gut microbiome changes in SLE is limited, and the association of changed gut microbiome with the activity of SLE, as well as its functional relevance with SLE still remains unknown. Here, we sequenced 16S rRNA amplicon on fecal samples from 40 SLE patients (19 active patients, 21 remissive patients), 20 disease controls (Rheumatoid Arthritis (RA) patients), and 22 healthy controls (HCs), and investigated the association of functional categories with taxonomic composition by Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt). We demonstrated SLE patients, particularly the active patients, had significant dysbiosis in gut microbiota with reduced bacterial diversity and biased community constitutions. Amongst the disordered microbiota, the genera Streptococcus, Campylobacter, Veillonella, the species anginosus and dispar, were positively correlated with lupus activity, while the genus Bifidobacterium was negatively associated with the disease activity. PICRUSt analysis showed metabolic pathways were different between SLE and HCs, and also between active and remissive SLE patients. Moreover, we revealed that a random forest model could distinguish SLE from RA and HCs (area under the curve (AUC) = 0.792), and another random forest model could well predict the activity of SLE patients (AUC = 0.811). In summary, SLE patients, especially the active patients, show an apparent dysbiosis in gut microbiota and its related metabolic pathways. Amongst the disordered microflora, four genera and two species are associated with lupus activity. Furthermore, the random forest models are able to diagnose SLE and predict disease activity.

Introduction

The gut microbe population, known as ‘gut microbiota’ is heterogeneous and complex, and is composed of more than 1000 different bacterial species [1]. Intestinal mucosal immunity have clarified the correlation between the gut microbiota and the host immune system [2]. Microbial abnormalities, also known as ‘dysbiosis’, is thought to be correlated with various diseases, including chronic kidney disease, obesity, type 2 diabetes, atherosclerosis, and nonalcoholic fatty liver disease [3–6]. Systemic Lupus Erythematosus (SLE) is a heterogenic autoimmune disease promoted by a combination of genetic and environmental factors that bring about an intolerance toward self-antigens [7]. Although the etiology of SLE remains unclear, hormonal, environmental, and genetic factors are thought to be of importance. Recently, dysbiosis of the gut microbial community in the development of SLE has attracted attention.
Multiple evidences have shown a lower Firmicutes/Bacteroidetes (F/B) ratio and decreased abundance of some families in Firmicutes phylum may be involved in remissive SLE [8–10] However, such alterations are also discovered in Intestinal Mucositis and Crohn’s disease [11]. It was reported that the presence of Lactobacillus spp. in gut could attenuate kidney inflammation in lupus-prone mice in a sex hormone-dependent manner [12], suggesting the gut microbiome may be a possible therapeutic target for SLE. The association of the gut microbiome with different diseases has been shown to be diverse due to many factors such as host’s age, sex, genotype, diet, and geography [13–16]. Therefore, alterations of gut microbiome associated with SLE should be variable in SLE patients in Guangdong Province, China compared with other locations.

So far, there are limited studies on SLE and gut microbiota. Only one study observed the gut microbiota in active SLE patients, other studies have just focussed on SLE patients in remission [8–10,17]. Whether the gut microbiota is associated with the disease activity still remains unclear. In the present study, we recruited both active and remissive SLE patients to investigate the characteristics of intestinal microbes that are associated with disease activity. Since Rheumatoid Arthritis (RA) is another common autoimmune disease, we also included RA patients as the disease control to define the specificity of the SLE-associated gut microbiome. We found that the gut microbiota in SLE patients, especially in active SLE patients, had a distinct dysbiosis in microbiota and its related metabolic pathways. Four disordered genera and two species were revealed to be closely associated with SLE activity. Furthermore, the results suggested the gut microbiota were validated to have strong diagnostic potential for SLE, and even predict the disease activity through random forest analysis.

Materials and methods
Research participants and sample collection
A total of 40 SLE patients, 20 RA patients, and 22 healthy controls (HCs) were consecutively recruited from Nanfang Hospital, Southern Medical University during 2017. All SLE and RA patients fulfill the American College of Rheumatology (ACR) classification criteria for SLE or RA disease [18–20]. All patients with acute intercurrent illnesses or infections and those who used probiotics or antibiotics within 1 month before admission were excluded [6]. The gender- and age-matched HCs who had no known history of autoimmune diseases were also recruited from the Health Examination Centre of Nanfang Hospital. All the participants were female. Average age of SLE, RA, and HC group was 37.46 ± 14.17, 44.00 ± 6.53, and 37.18 ± 14.67 respectively (P=0.142).

Based on the SLE disease activity index (SLEDAI) [21], all the SLE patients were divided into the active SLE patients (A) (SLEDAI ≥ 8) (n=19) and remissive SLE patients (R) (SLEDAI < 8) (n=21). Exceptions of the age and gender distribution, patients in group A showed many significant differences from that of group R, having more severe symptoms, including anemia, hypocomplementemia, impaired renal functions, and increased autoantibodies, all of which are consistent with the clinical characteristics of SLE (Table 1).

For all participants, fresh fecal samples were frozen at −80°C immediately after collection. Ethics approval was granted by the Ethics Committee of Nanfang Hospital, and all the methods used were in accordance with the approved guidelines. Written informed consent was required from all patients and healthy volunteers in the study.

Illumina Miseq sequencing of 16S rRNA gene-based amplicons and data processing
Total DNA was extracted from thawed fecal samples using the LONGSEE STOOL DNA KIT (Longsee med Bio edicai., Ltd., Guangdong, China) following the manufacturer’s instructions. All the individually processed human fecal DNA extractions were amplified by polymerase chain reaction (PCR). The forward primer (5′-ACT CCT ACG GGA GGC AGC AG-3′) and reverse primer (5′-GGA CTA CHV GGG TWT CTA AT-3′) were used to amplify the 16S rRNA gene V3–V4 variable region from the bacteria by PCR as described previously [22]. Briefly, amplifications were performed using a step cycling protocol consisting of 98°C for 30 s, 35 cycles of 98°C for 10 s, 54°C for 30 s, and 72°C for 45 s, ending with the final elongation at 72°C for 10 min. PCR products were purified using an AxyPrep PCR Cleanup Kit (Axygen, California, U.S.A.).

For the sequencing of 16S rRNA gene-based amplicons, the amplicon library was prepared using a TruSeq Nano DNA LT Library Prep Kit (Illumina Inc, CA, U.S.A.). The sequencing reaction was conducted using Illumina MiSeq platforms and the data were analyzed by the Quantitative Insights Into Microbial Ecology platform (QIIME, www.qiime.org) using the default parameters [23]. The raw sequence data for 16S rRNA gene sequencing datasets was available from the Sequence Read Archive (SRA) database (http://www.ncbi.nlm.nih.gov/sra) at accession number PRJNA493726.
## Table 1 Characteristics of the active SLE patients and remissive SLE patients

<table>
<thead>
<tr>
<th></th>
<th>A ((n=19))</th>
<th>R ((n=21))</th>
<th>(P)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>34.05 (13.92)</td>
<td>40.57 (13.63)</td>
<td>0.143</td>
</tr>
<tr>
<td>WBC, (10^9/\text{l})</td>
<td>7.40 (5.32–8.53)</td>
<td>6.72 (5.72–8.22)</td>
<td>0.955</td>
</tr>
<tr>
<td>RBC, (10^{12}/\text{l})</td>
<td>4.17 (3.49–4.35)</td>
<td>4.33 (4.00–4.58)</td>
<td>0.0389²</td>
</tr>
<tr>
<td>HGB, (\text{g/l})</td>
<td>106 (95–117.50)</td>
<td>126 (115–133.25)</td>
<td>&lt;0.001⁴</td>
</tr>
<tr>
<td>PLT, (10^9/\text{l})</td>
<td>205 (152.50–308.00)</td>
<td>235.50 (202.25–272.50)</td>
<td>0.558</td>
</tr>
<tr>
<td>C3, (\text{g/l})</td>
<td>0.68 (0.32)</td>
<td>0.93 (0.23)</td>
<td>0.001³</td>
</tr>
<tr>
<td>CRP, (\text{mg/l})</td>
<td>1.68 (0.65–4.66)</td>
<td>1.56 (0.57–4.07)</td>
<td>0.765</td>
</tr>
<tr>
<td>ESR, (\text{mm/h})</td>
<td>21.50 (11–53.50)</td>
<td>19.00 (5.00–60.00)</td>
<td>0.457</td>
</tr>
<tr>
<td>Pyuria, (%)</td>
<td>6 (31.58)</td>
<td>17 (80.95)</td>
<td>0.002³</td>
</tr>
<tr>
<td>- (negative)</td>
<td>13 (68.42)</td>
<td>4 (19.05)</td>
<td></td>
</tr>
<tr>
<td>+ (positive)</td>
<td>3 (15.79)</td>
<td>15 (71.43)</td>
<td>0.000⁴</td>
</tr>
<tr>
<td>Albuminuria, (%)</td>
<td>16 (84.21)</td>
<td>6 (28.57)</td>
<td></td>
</tr>
<tr>
<td>- (negative)</td>
<td>4 (21.05)</td>
<td>19 (90.48)</td>
<td>0.000⁴</td>
</tr>
<tr>
<td>+ (positive)</td>
<td>15 (78.95)</td>
<td>2 (9.52)</td>
<td></td>
</tr>
<tr>
<td>24-UTP, (\text{mg/24 h})</td>
<td>0.68 (0.22–2.29)</td>
<td>0.17 (0.11–0.46)</td>
<td>0.048²</td>
</tr>
<tr>
<td>Anti-dsDNA, UI/ml⁻¹</td>
<td>65.64 (16.24–156.66)</td>
<td>20.07 (2.88–62.20)</td>
<td>0.016³</td>
</tr>
<tr>
<td>Lupus nephritis, (%)</td>
<td>15 (78.95)</td>
<td>9 (42.88)</td>
<td>0.20</td>
</tr>
<tr>
<td>SLEDAI</td>
<td>12 (9.5–14.0)</td>
<td>4 (1.5–6.0)</td>
<td>&lt;0.001⁴</td>
</tr>
<tr>
<td>Medication use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydroxychloroquine</td>
<td>12</td>
<td>10</td>
<td>0.119</td>
</tr>
<tr>
<td>Glucocorticoid</td>
<td>17</td>
<td>16</td>
<td>0.527</td>
</tr>
<tr>
<td>Cyclophosphamide</td>
<td>3</td>
<td>3</td>
<td>0.574</td>
</tr>
<tr>
<td>Biological agent</td>
<td>5</td>
<td>2</td>
<td>0.894</td>
</tr>
</tbody>
</table>

Abbreviations: A, the active SLE patient; Anti-dsDNA, anti-double stranded DNA; CRP, C-reactive protein; C3, complement component 3; ESR, erythrocyte sedimentation rate; HGB, hemoglobin; PLT, platelet; R, the remissive SLE patient; RBC, red blood cell; WBC, white blood cell; 24-UTP, 24-h urine protein.

Data represent the mean (standard deviation), the data and (1) represents the median (interquartile range). The \(P\)-values were calculated by Mann–Whitney \(U\)-test or Chi-square test.

\(²P<0.05.\)  
\(³P<0.01.\)  
\(⁴P<0.001.\)

Before assembly, sequence reads were first filtered to remove low-quality or ambiguous reads, including reads lacking exact matching with the primer, sequences with mismatch ratio sequences higher than 0.05 in the overlap region, and raw reads shorter than 100bp with Trimmomatic v.0.32 software [24]. Paired-end clean reads were merged using FLASH [25] according to the relationship of the overlap between the paired-end reads when at least 10 of the reads overlapped the read generated from the opposite end of the same DNA fragment, the maximum allowable error ratio of an overlap region of 0.2, and the spliced sequences were called raw tags.

High-quality sequences with a distance-based similarity of 97% or greater were grouped into operational taxonomic units (OTUs) using the Vsearch algorithm. Representative sequence was then extracted from each OTU. Next, the chimeric sequences were detected and removed. To assign taxonomy information to each clustered feature, extracted representative sequences were subjected to similarity search against Greengenes sequence and taxonomy database using RDP classifier algorithm (uccluster approach with default settings) and the classify-sklearn plugin within QIIME software (version 1.9.1). The phylogenetic relationships were determined based on a representative sequence alignment using Fast-Tree [26]. Computation of \(\alpha\)-diversity metrics and \(\beta\)-diversity metrics were performed on all samples within the feature table with Qime diversity \(\alpha/\beta\) plugin. Rarefaction curve plots the number of individual’s sample versus the number of species, which was done with Qiime diversity \(\alpha\)-rarefaction plugin. Rank abundance curve portray relative abundance and species diversity within a community by plotting relative abundance of species (y-axis) against their rank in abundance (x-axis), which was plotted using QIIME v.1.9.1 software.

Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) is a bioinformatics software package designed to predict metagenome functional content from marker gene surveys and full genomes.
PICRUSt analysis was performed to identify Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathways, and determine functional categories associated with taxonomic composition [27]. Comparisons of relative abundance of taxa between groups were performed using Linear discriminant analysis effect size (LEfSe), a non-parametric Mann–Whitney U-test applied to detect features with significant differential abundance with respect to the groups compared, followed by a Linear Discriminant Analysis (LDA) to estimate the effect size of each differentially abundant feature in Linux platform [28].

**Statistical analyses**

We used the mean (± S.D.) to express measurement data that obeyed a normal distribution, the median (interquartile range) to express measurement data that obeyed a skewed distribution, and a percentage to express enumeration data. Mann–Whitney U-test or Student's t test was performed to compare the variables of two sample groups. Multiple group comparisons were made by the Kruskal–Wallis test or one-way analysis of variance. False discovery rate (FDR) correction for multiple comparisons was employed, and the statistical power was analyzed via power and sample size calculation in R software [29,30], then the FDR q-value was calculated.

The α-diversity determines the species richness and evenness within bacterial populations. The α-diversity metrics include: Observed species and Chao1 (microbial richness), and Shannon index and Simpson index (microbial diversity) [31]. The β-diversity determines the shared diversity between bacterial populations. Different distance metrics reveal distinctive views of community structure. UniFrac distances measure the shared phylogenetic diversity between communities. A smaller UniFrac distance between two samples indicates a higher similarity amongst the two microbial communities [32]. Principal coordinates analysis (PCoA) was plotted using the package in R software (version 3.4.4). The Wilcoxon rank sum test was used to determine significance in α-diversity and β-diversity.

We used Spearman algorithm to analyze the relationship amongst microbiota, predicted pathways, and SLE activity index. The Random Forest models were trained by ‘randomForest’ package with default parameters in R, then the performance of the model was assessed with a ten-fold cross-validation approach and measured by area under the receiver-operating characteristic (ROC) [33]. All tests were performed using GraphPad Prism (v6.0) (GraphPad Software, Inc., CA, U.S.A.), SPSS Statistics (V. 24.0.0.0) (SPSS Inc., Chicago, U.S.A.) or R software (version 3.4.4).

**Results**

**Characteristics of 16S rRNA sequences**

A total of 82 samples were subjected to 16S rRNA sequencing. These samples were composed of three groups including 22 healthy individuals, 20 RA patients, and 40 SLE patients. We obtained 2182143 16S rRNA sequencing reads from stool samples of SLE patients, 976140 reads from RA patients, and 1277858 reads from HC, which belong to 714 kinds of OTUs. The parameters, including Chao1 rarefaction curves, Shannon rarefaction curve, and rank abundance of OTUs, were evaluated to confirm the reliability of the sequencing data (Supplementary Figure S1).

**Difference of the gut microbiota in SLE patients from those of controls**

The α-diversity between two groups was compared using Chao1, Observed species, Shannon index, and Simpson diversity indices. Overall, the α-diversity metrics Chao1 and Observed species were significantly higher in HCs than in SLE patients (P=0.038; P=0.004, respectively), indicating that the gut microbiome in SLE patients exhibited a lower richness than HCs (Figure 1A,B and Supplementary Table S1). However, no difference in Shannon and Simpson index (P=0.089; P=0.092, respectively) was observed between SLE patients and healthy individuals, suggesting that the evenness of the gut microbiome of the two groups had no significant difference (Supplementary Table S1). There were no associations between α-diversity and drug treatments, such as Hydroxychloroquine, Glucocorticoid, Cyclophosphamide, and Biological agent (Supplementary Figure S2).

To measure the extent of the similarity of fecal microbial communities, β-diversity was calculated using UniFrac distances. PCoA based on weighted and unweighted UniFrac distance matrix were used for visualizing sample relationships, and ADONIS analysis was used to test the homogeneity of dispersion amongst different groups. Our results suggested that there were no associations between β-diversity and medicine treatments, including Hydroxychloroquine, Glucocorticoid, Cyclophosphamide, and Biological agent (Supplementary Table S2 and Figure S3), however, the unweighted UniFrac distance analysis of β-diversity difference demonstrated that the structure of microbiota of SLE patients differed from HCs (ADONIS analysis, P<0.001, R² = 0.054) (Figure 1C and Supplementary Table S3). Thus, the microbial diversity was significantly different between SLE group and HCs.

We then analyzed the phylum-level profiles of feces between SLE patients and HCs. The phylum level profiles for gut microbiota of SLE patients and controls were fairly similar, except for reads from the phyla Fusobacteria.
Figure 1. The different microbial diversity between SLE patient group and HCs

(A,B) Significantly different richness of \( \alpha \)-diversity between the gut microbiota of SLE and HC. (C) PCoA illustrating the grouping patterns of SLE and HC group based on the unweighted UniFrac distances. Each closed circle represented a sample. Distances between any pair of samples represented their dissimilarities. (D) The average relative abundances of the predominant bacterial taxa at the phylum level in the SLE patients and HC group. (E,F) The significantly different phyla in SLE patients compared with HCs \((P=0.027)\) and Tenericutes \((P=0.002)\) (Figure 1D–F). A lower F/B ratio was reported in the feces of remissive SLE patients compared with HCs [34]. However, we showed the ratio of F/B the feces of SLE had a decreasing trend but no significant difference compared with HC group (Supplementary Table S4).

To further determine the phylogenetic clustering pattern between these two groups, the logarithm LDA was performed (Figure 2). The phylum Tenericutes, along with Mollicutes and RF39, were significantly reduced in the intestinal flora of SLE patients compared with HCs. In addition, patients with SLE exhibited a significant decrease in the genus Faecalibacterium alongside its species prausnitzii, while the taxonomic clade Cryptophyta and genus Roseburia were reduced in the gut microbiota of SLE group. On the contrary, the taxonomic clade Bacilli from the phylum Firmicutes showed clustered differences, while Streptococcaceae and Lactobacillales were expanded in the feces of SLE patients compared with HCs. Moreover, the genera Streptococcus and Lactobacillus, along with their species Streptococcus anginosus and Lactobacillus mucosae were enriched in the intestinal flora of SLE group compared with HC group. In addition, the feces of SLE patients showed an increase in genus Megaphaera (significant taxa \([P<0.005, \text{Kruskal–Wallis test}]\) with LDA score > 2 were shown). Taken together, sequence profiling of the gut microbiota revealed an apparent dysbiosis of the gut microbiota in SLE patients, which was characterized by reduced bacterial \( \alpha \)-diversity and biased community constitutions. These results demonstrated the gut microbiota of patients with SLE differed from those of HCs.

To investigate whether the disordered intestinal microbes were specific to SLE patients, we further compared the intestinal microbiota distribution between SLE and RA patients. There were no significant differences in \( \alpha \)-diversity (Supplementary Table S1) and \( \beta \)-diversity (Supplementary Table S3 and Figure S4C,D) between two groups. LEfSe analysis showed the different microbiota between SLE group and RA group (Supplementary Figure S5). The taxonomic clade EB1017, Ellin6529, and Anaerofilum were increased in the intestinal flora of RA patients, while the Lactobacillales from the Bacilli, with its genus Streptococcus were enriched in the feces of SLE patients compared with RA patients. In addition, the phylum Fusobacteria, along with its taxonomic clade Fusobacteriia, Fusobacteriales, Fusobacteriaceae, and Fusobacterium were increased in the feces of SLE patients. The genus Megaphaera and Veillonella, with its species Veillonella dispar, were also enriched in the feces of SLE group compared with RA group (significant taxa \([P<0.05, \text{Kruskal–Wallis test}]\) with LDA score > 2.5 were shown).

Collectively, these results demonstrated that the gut microbiota of SLE patients differed from healthy individuals, however, there was no significant difference in gut microflora diversities between SLE and RA patients. The genera...
Figure 2. The differentially abundant taxa between the feces of SLE patients and HCs
LEfSe analysis was performed to identify differentially abundant taxa, which are highlighted by the phylogenetic tree in cladogram format (A) and the LDA scores (B). Green color indicates an increased taxa in the feces of SLE compared with HC, while the red color indicates an increased taxa in the feces of HC compared with SLE (significant taxa [\(P<0.005\), Kruskal–Wallis test] with LDA score > 2 were shown).

*Streptococcus* and *Megasphaera* were specifically increased in the feces of SLE patients compared with HCs and RA patients.

**Difference of microbiota profiling in active SLE patients from remissive SLE patients**

Given that the gut microbiota was significantly different between SLE patients and HCs, we next investigated whether the gut microbiota was associated with disease activity of SLE. First, we compared 16S rRNA sequences of A group (active SLE patients) with R group (remissive SLE patients). The unweighted UniFrac distance analysis of \(\beta\)-diversity difference demonstrated that the structure of the microbiota of A group differed from R group (ADONIS analysis, \(P=0.047\), \(R^2 = 0.039\)) (Figure 3A and Supplementary Table S3), while no obvious difference was observed in \(\alpha\)-diversity (Supplementary Table S1), suggesting that the community constitutions in A group were distinctly different from R group, but no difference was found in microbial diversity.

As shown in Figure 3, LEfSe analysis further demonstrated that *Actinomycetales* and *Bifidobacteriales* from phylum *Actinobacteria* showed clustered differences, and the genus *Bifidobacterium* was increased in the feces of remissive SLE patients compared with active SLE patients. In addition, the species *Ruminococcus gnavus* was reduced in the feces of active SLE patients, whereas Lactobacillales from the *Bacilli*, along with its genus *Streptococcus* and species *S. anginosus*, were enriched in the feces of A group compared with R group. Moreover, the genus *Oribacterium* was increased in the intestinal flora of active SLE patients. Furthermore, active SLE patients exhibited a
Figure 3. The different microbial diversity between the feces of active SLE patients and remissive SLE patients

(A) PCoA illustrating the grouping patterns of the feces of A group and R group based on the unweighted UniFrac distances. Each closed circle represented a sample. Distances between any pair of samples represented their dissimilarities. (B) The significantly different phyla in the feces of A group compared with R group. (C,D) LEfSe analysis was performed to identify differentially abundant taxa, which are highlighted by the phylogenetic tree in cladogram format (C) and the LDA scores (D). Significantly discriminative taxa amongst the active patients (red), remissive patients (blue), and HCs (green) were determined using LEfSe (significant taxa \( P<0.005 \), Kruskal–Wallis test) with LDA score > 2 were shown.)
Figure 4. The significantly different predicted metabolic pathways between SLE patients and HCs
(A) Significantly two predicted metabolic pathways were increased in SLE compared with HCs. (B) Four predicted metabolic pathways were decreased in SLE patient group compared with HCs. It was analyzed by Kruskal–Wallis, and the FDR q-value was then calculated, and q-values < 0.1 were considered significant.

remarkable enrichment of the taxa Epsilonproteobacteria from phylum Proteobacteria, along with its Campylobacteria and Campylobacter (significant taxa [P < 0.005, Kruskal–Wallis test] with LDA score > 2 were shown). Finally, the ratio of F/B in remissive SLE patient group had a decreasing trend but no significant difference compared with HC group (Supplementary Table S4). Altogether, these results indicated that the gut microbiota profiling of active SLE patients were markedly different from that of remissive SLE patients.

Aberrant microbiome-associated pathway existed in SLE patients
Another emphasis of our study was to disclose the functional variation in the SLE gut microbiota community. Therefore, we predicted the microbiota-derived pathways using the PICRUSt algorithm with the KEGG database and compared functional abundances amongst the SLE, RA, and HC groups. In total, we characterized six different pathway categories between SLE group and HC group (Figure 4). The pathways of Apoptosis and Purine metabolism were significantly increased in SLE patient group compared with HC group (Figure 4A), while four pathways, including Pathways in cancer, Bacterial chemotaxis, Bacterial motility proteins, and Flagellar assembly, were decreased in SLE patients (Figure 4B). In addition, nine different functional pathways were identified between A group and R group (Figure 5). Five were related to Synthesis and degradation of ketone bodies, Apoptosis, Lipid metabolism, Secretion system, and Staphylococcus aureus infection, which were significantly higher in active SLE patients than remissive patients (Figure 5A). Conversely, Alanine aspartate and glutamate metabolism, Carbohydrate metabolism, Primary
bile acid biosynthesis, and Secondary bile acid biosynthesis, were obviously increased in remissive SLE patients compared with active patients (Figure 5B). However, there was no different pathway between SLE and RA group (data not shown).

We further examined correlations amongst SLE/HC-associated taxa and disordered functional pathway to obtain an overview of how specific taxa act during metabolic dysfunction in patient gut. For SLE patients, we characterized a positive correlation between the enrichment of *Streptococcus* and increased Apoptosis pathway ($r = 0.807$, $P < 0.000$, FDR < 0.000) and a negative correlation between *Streptococcus* and Pathways in cancer ($r = -0.550$, $P < 0.000$, FDR < 0.000) (Figure 6A). Further analysis also revealed the active SLE patient-enriched genus *Streptococcus* was negatively associated with pathways of Alanine aspartate and glutamate metabolism, Primary and secondary bile acid biosynthesis ($r = -0.680; r = -0.437; r = -0.434$, $P < 0.01$, FDR < 0.05, respectively) (Figure 6B), but positively associated with five increased pathways, including Synthesis and degradation of ketone bodies, Apoptosis, Lipid metabolism, Secretion system, and *Staphylococcus aureus* infection ($r = 0.574; r = 0.829; r = 0.406; r = 0.486; r = 0.903$, $P < 0.01$, FDR < 0.05, respectively) (Figure 6C).

Thus, several aberrant microbiome-associated gut metabolic pathways were associated with SLE using PICRUSt analysis. Interestingly, the SLE-enriched genus *Streptococcus* was positively associated with the pathways of Apoptosis, the metabolism of lipid, amino acid and bile acid, Secretion system, and pathogenic bacteria infection.

Figure 5. The significantly different predicted metabolic pathways between active SLE patients and remissive SLE patients (A) Significantly five predicted metabolic pathways were increased in active SLE compared with remissive SLE patients. (B) Four predicted metabolic pathways were enriched in remissive SLE patients compared with active SLE patients. A, the active SLE patient group; R, the remissive SLE patient group. It was analyzed by Kruskal–Wallis, and the FDR q-value was then calculated, and q-values < 0.1 were considered significant.
**Figure 6.** The associations between the abundance of genus *Streptococcus* and disordered metabolic pathways

(A) The predicted metabolic pathway Apoptosis and Pathways in cancer were correlated with SLE-enriched genus *Streptococcus*. (B) The predicted metabolic pathway of Alanine aspartate and glutamate metabolism, Primary bile acid biosynthesis and Secondary bile acid biosynthesis were negatively associated with active SLE-enriched genus *Streptococcus*. (C) Five pathway categories, which were higher in active SLE patients compared with remissive SLE patient group, were positively associated with active SLE-enriched genus *Streptococcus*. They were analyzed by Spearman ranks tests, and the FDR was calculated for multiple testing.

**Association of disordered microbiota and aberrant microbiome-associated pathway with activity of SLE**

SLEDAI, Complement C3, C-reactive protein (CRP), Erythrocyte Sedimentation Rate (ESR), and anti-double stranded DNA (anti-dsDNA) were commonly used to indicate the disease activity of SLE patients [21,35,36].

At genus and species levels, *Lactobacillus*, *Streptococcus*, *Megaphaera*, *Fusobacterium*, *Veillonella*, *L. mucosa*, *S. anginosus*, and *V. dispar* were increased in the feces of SLE patients compared with HCs or RA patients. Meanwhile, *Streptococcus*, *Oribacterium*, *Campylobacter*, and *S. anginosus* were enriched, but *Bifidobacterium* and *Ruminococcus gnavus* were reduced in the gut microbiota of active SLE patients compared with that of remissive SLE patients. Except *Bifidobacterium*, five changed genera were positively associated with disease activity (Figure 7A–F and Supplementary Table S5). For example, the abundance of *Streptococcus* was positively correlated to SLEDAI ($r = 0.492$, FDR $q = 0.008$), while negatively associated with Complement C3 ($r = -0.502$, FDR $q = 0.008$) (Figure 7A). *Campylobacter* and *S. anginosus* also showed a positive correlation with SLEDAI ($r = 0.470$, FDR $q = 0.009$; $r = 0.388$, FDR $q = 0.040$, respectively) (Figure 7). Moreover, the abundance of *Veillonella* and its species *V. dispar* showed negative correlations with Complement C3 ($r = -0.475$, FDR $q = 0.008$) (Figure 7).
Figure 7. Associations amongst disease activity, disordered genera, and predicted pathways in the gut microbiota of SLE patients

(A–C) Three genera, *Streptococcus*, *Campylobacter*, and *Veillonella*, were positively correlated with lupus activity. (D, E) Two species, *S. anginosus*, and *V. dispar*, were positively correlated with lupus activity. (F) The genus *Bifidobacterium* was negatively related to lupus activity. (G–I) The aberrant microbiome-associated pathways, Alanine aspartate and glutamate metabolism, Secondary bile acid biosynthesis and Lipid metabolism, had a positive association with the activity of SLE patients. It was analyzed by Spearman ranks tests, and the FDR q-value was then calculated for multiple testing. Abbreviation: C3, Complement component 3.

The genus *Streptococcus*, which was specifically associated with the activity of SLE, was related to eight aberrant microbiome-associated pathways (Figure 6). We further explored whether these eight disordered pathways were also related to the activity of SLE (Supplementary Table S5). Alanine aspartate and glutamate metabolism, Secondary bile acid biosynthesis, and Lipid metabolism were closely associated with SLEDAI (r = −0.376; r = −0.382; r = 0.318, FDR q < 0.001, respectively) (Figure 7G–I). As such we hypothesized that the genus *Streptococcus* might play an important role in the disease progression of SLE through these three pathways.

**Potentials of gut microbiota for SLE diagnosis or disease activity monitoring**

Given that the gut microbiota in SLE patients, especially in active SLE patients, had a distinct dysbiosis in microbiota, we next addressed the potential diagnostic value of the gut microbiota as potential biomarkers for SLE by ROC curve analyses. Due to its non-parametric assumptions, random forest was used to detect linear and nonlinear effects and potential taxon–taxon interactions, to identify taxa that could differentiate SLE subjects from control subjects (HCs and RA patients), and to discriminate active SLE patients from remissive patients. We used ten-fold cross-validation approach to evaluate the performance of model, and predictive power was scored in ROC analysis. We first made the
mode to differentiate the SLE patients from HCs and RA patients based on the genus and species levels. We showed that the area under the curve (AUC) was 0.792 (95% confidence interval (CI): 0.750−0.835) (Supplementary Table S6 and Figure 8A), suggesting that the gut microbiota had the potential to diagnose SLE from healthy and disease controls (RA patients). We observed that in the model, out of the top ten genera and species, eight belonged to the phylum Firmicutes, one belonged to Fusobacteria, and one belonged to Actinobacteria. Of the eight genera in the Firmicutes phylum, five were part of Clostridia class, and three were Bacilli (Supplementary Table S6). Furthermore, amongst the ten genera and species, the mucosa, Lactobacillus, Megasphaera, and Streptococcus were significantly enriched, while Faecalibacterium was decreased in the feces of SLE patients compared with HCs. In addition, both Veillonella and Fusobacterium were increased in the gut microbiota of SLE patients than RA patients (Figure 2; Supplementary Figure S4 and Table S5). Accordingly, most of the genera and species in the model were the disordered genera in the feces of SLE group compared with HCs and RA patients.

We further built another model to distinguish active SLE patients from remissive patients based on the genus and species levels. In this model, the AUC was 0.811 (95% CI: 0.754–0.869) (Supplementary Table S7 and Figure 8B), suggesting that the gut microbiota had the potential to monitor the activity of SLE. Anti-dsDNA was reported to be reasonably sensitive and specific in the diagnosis of SLE, and raised titers of anti-dsDNA along with hypocomplementemia were associated with the activity of SLE [37]. We showed that the AUC value for combination of Complement C3 and anti-dsDNA was only 0.773 (95% CI: 0.597–0.949) (Supplementary Figure S6). These results indicated that the combination of the gut microbiota might have a better surveillance value for SLE activity than the combination of Complement C3 and anti-dsDNA. Moreover, as shown in the model, out of the top ten genera and species, five belonged to the phylum Firmicutes, four belonged to Actinobacteria, and one belonged to Proteobacteria. Amongst the five genera in the Firmicutes phylum, three were from the Clostridia class, one was Erysipelotrichi, and one was Bacilli (Supplementary Table S7). In this case, the Campylobacter, Streptococcus, and Orbibacterium were enriched, while the gnavus and Bifidobacterium were reduced in active SLE patients compared with remissive SLE patients (Figure 3 and Supplementary Table S7). Altogether, a great part of the genera and species in the model were disordered genera in the feces of active SLE patients, suggesting that the disordered intestinal flora might have potential to diagnose SLE, even monitor disease activity.

Discussion

SLE is an autoimmune disease that affects multiple tissues, and causes joint pain, renal disease, muscle pain, fever, poor circulation, inflammation, fatigue, loss of appetite, and other symptoms [38]. Though the cause of SLE still remains unclear, it is thought to be involved with hormonal, genetic, and environmental factors [39]. The gut microbiome was believed to be a key factor in influencing predisposition to autoimmunity diseases [40]. Recent studies further supported that gut microbiome dysbiosis could act as an important factor in promoting chronic inflammation into autoimmune diseases [2,41,42]. However, there were only limited works in exploring the potential relationship of gut
microbiome with SLE [8–10,17,39]. In the present study, we have provided new evidence about the gut microbiome dysbiosis in female SLE patients by fecal bacteria sequencing. Importantly, we for the first time explored whether the gut disordered microbes were associated with the activity of SLE.

We investigated the profiling of the gut microbiota and showed a distinct dysbiosis of the gut microbiota in SLE patients, which was characterized by reduced bacterial α-diversity and biased community compositions. Most of the patients in our study were currently on various immunosuppressants and glucocorticoids treatments. Chen et al. have demonstrated that RA patients using methotrexate (MTX) and hydroxychloroquine exhibited an increase in species richness and diversity [43]. However, our results showed no significant relationship between drug treatments and the abundance diversity of gut microbiota in the SLE patients, which might be because most of the enrolled patients were treated with steroids or immunosuppressants, while only four patients did not use any drugs.

Phyla Firmicutes together with Bacteroidetes usually account for more than 90% of all phylogenetic species, were involved in host metabolism and immunity [44]. In our study, the Phyla Firmicutes and Bacteroidetes occupied the most abundant microorganisms, consistent with the typical human intestinal microbiome structures. It was reported that the F/B ratio was significantly lower in the feces of SLE patients in remission [8]. However, there was no significant difference for our cohort of remissive SLE patients and HCs (P>0.05). Also, there was no significant difference in the F/B ratio between active SLE and HCs (P>0.05), consistent with the available data [17]. The changes in the genera in SLE patients of our study were only partly consistent with previous studies [8,10,17,45], which might partially be due to the sample size and geographical locations of patients. It is well known that cohorts with different patient characteristics, including disease stage, geographical locations, diet and status, might exhibit different gut microbiota profiling [15,16,46–48]. Therefore, the alterations of gut microbiome associated with SLE should display differences amongst different geographical locations and disease status.

In the present study, we found that the abundance of pathogenic genus Streptococcus, with its species anginosus, and genus Megasphaera were significantly enriched in the feces of SLE patients compared with HCs; genus Streptococcus and its species anginosus were positively correlated to the activity of SLE. In addition, the genus Veillonella and its species dispar were significantly increased in the gut microbiota of SLE patients compared with RA patients and had a positive association with the activity of SLE. The association of these disordered genera with the activity of SLE was most striking, and to our knowledge, this is the first study to describe such a significant relationship with SLE. The genera Streptococcus and Megasphaera were reported to be closely related to the intestinal disturbance of autoimmune disorders. For example, Streptococcus and Megasphaera were enriched in primary biliary cirrhosis [49] and Pediatric Autoimmune Neuropsychiatric Disorders [50]. Also, Streptococcus was relatively increased in RA patients [43]. It was demonstrated that S. anginosus rarely caused infections in healthy individuals, but caused infections in the immunodeficient individuals [51]. As reported, genera Streptococcus and Veillonella had pro-inflammatory effects. For example, the combination of Streptococci with Veillonella appeared to negate IL-12p70 production, while augment IL-8, IL-6, IL-10, and tumor necrosis factor α (TNF-α) response [52].

The SLE patients, especially the active patients, had an increased population of oral bacteria, which is an interesting phenomenon that occurred in the intestinal flora of SLE. However, the gut microbiome of liver cirrhosis, colorectal cancer, RA, and ACVD patients also showed an increase in the abundance of oral bacteria in gut microbiota [53], and only RA and ACVD have been epidemiologically associated with periodontitis. Interestingly, our results suggested that the abundance of genus Streptococcus was enriched in active SLE patients, suggesting that the oral microbiota might be overrepresented in the lower gastrointestinal populations of patients with active SLE. Besides, more severe forms of periodontitis were found in SLE subjects that had higher bacterial loads [54], resulting in an increase in oral bacteria entering the intestine.

Furthermore, our data showed that many beneficial commensal microbes, such as Roseburia, Faecalibacterium and its species prausnitzii were depleted in SLE patients. Meanwhile, the genus Bifidobacterium was adversely correlated with activity of SLE. These microbes belonged to the phylogenetic core of the intestinal microbiota [55,56], which can produce short chain fatty acids (SCFAs), especially butyrate, to play multiple critical roles in the maintenance of human health, including producing energy components and intestinal epithelial nutrition [57], reducing the severity of inflammation [58], maintaining intestinal barrier functions [59], and enhancing colon motility functions [60].

Moreover, we observed an increased abundance of beneficial commensal genus Lactobacillus and its species L. mucosae in the feces of SLE cohort compared with HCs. Supportive of a role for Lactobacilli in the pathogenesis of lupus, taxa in this genus were found to be enriched in female NZB/W F1 mice, the model of systemic lupus. In the present study, Lactobacillus spp. were associated with more severe disease, whereas they were reduced as disease is controlled with dexamethasone [17]. As reported [61] Lactobacillus reuteri increased over time in the feces of mice from both lupus models as their disease progress, in addition, Lactobacillus spp. were increased in a longitudinal
cohort of SLE patients compared with HCs. In the present study, the pDC/IFN-promoting properties of *L. reuteri* in the context of a lupus-prone host suggest a paradigm in which a bacterium that is normally considered a probiotic may become harmful under certain genetic or environmental conditions. We also observed that *Lactobacillus* were enriched in feces of SLE patients, suggesting a potential role for these taxa in SLE pathogenesis, which need further research in the future.

Our study has demonstrated that some pro-inflammatory bacteria in genera *Streptococcus*, and *Campylobacter* expanded, while some anti-inflammatory bacteria in genera *Roseburia*, *Faecalibacterium*, and *Bifidobacterium* reduced in the feces of SLE patients, especially the active patients, resulting in the release of inflammatory factors, then aggravating the systemic inflammation level. Some pro-inflammatory pathogens increased accompanied with the intestinal mucosal barrier compromised, which lead to more bacterial LPS transferring into lymph nodes and blood to stimulate the TOLL-like pathway of the host cells, and produce inflammatory cytokine [62]. SLE patients generally used massive immunosuppressive agents and glucocorticoids during the active period, which could inhibit the immune system and might cause a large increase in opportunistic pathogens [63,64]. Notwithstanding, it is questionable whether such changes in gut bacterial profile are a cause or consequence of SLE. However, to posit further on this, is beyond the scope of the present study, and we will focus on this in the future research.

In addition, several aberrant microbiome-associated gut metabolic pathways were revealed to be associated with SLE using PICRUSt analysis. We found that SLE patients were enriched in multiple metabolic pathways containing gene functions of Apoptosis, Purine metabolism, and the Apoptosis were positively associated with the genus *Streptococcus* that was highly enriched in SLE patients and especially in active patients. As reported, Apoptosis pathway played an important role in the pathogenesis of SLE [65]. Besides, amongst these altered pathways, the alanine, aspartate and glutamate metabolism, Secondary bile acid biosynthesis, and Lipid metabolism were not only related to the disease activity, but also significantly associated with *Streptococcus*. The alanine, aspartate and glutamate metabolism, which was identified to be increased in remissive SLE patients in our study, had been previously reported to play a pivotal role in resting or activated T cells [66]. Lipid metabolism participates in the regulation of many cellular processes such as cell growth, proliferation, differentiation, survival, apoptosis, inflammation, motility etc [67]. In active SLE patients, the dyslipidemia was more prevalent, suggesting that inflammation may be related to lipid metabolism [68]. Thus, *Streptococcus* might play an important role in the pathogenesis of SLE through these pathways.

Due to the heterogeneous presentation of SLE patients and their unpredictable disease course, there is a great need for accurate assessment of disease activity. Several immunologic markers including anti-dsDNA antibody and complement are commonly used in laboratory monitoring of disease activity, however, these traditional biomarkers are better related to certain clinical manifestations of the disease, especially nephritis, rather than to the activity of the disease itself [69]. Currently, disease activity in SLE can be assessed using composite disease activity indices, such as SLEDAI score and British Isles Lupus Assessment Group (BILAG) score [70]. However, the composite disease activity indices depend on differential organ involvement and physical assessments [71,72]. Besides, they could be complex for use in routine clinical practice. Thus, there is a great urgent for the identification of new biomarkers that can quantitate disease activity [73,74].

Moreover, due to the existence of a remarkable difference in microbiota between SLE status, the random forest models were built in the present study to examine whether microbiota composition could identify their disease status. Of note, a random forest model was identified for diagnosing SLE from HCs and RA patients with an AUC value of 0.792. To be mentioned, another random forest predictive model showed to be a suitable model for the prediction of disease activity of SLE with the AUC of 0.811, which was higher than the combination of Complement C3 and anti-dsDNA (AUC = 0.773). Accordingly, our results suggested that the gut microbiota might be potential biomarkers for diagnosis of SLE and even monitoring SLE activity in a non-invasive method. However, the sample size enrolled in our study was relatively small, therefore, more samples are needed to evaluate the performance of the disordered genera in the future.

In summary, these disordered bacteria and related metabolic pathways might provide clues for studying the SLE pathogenesis, and searching for suitable biomarkers for the diagnosis SLE or monitoring SLE activity in a non-invasive method. Specific microbial clades might be viable targets for the therapeutic manipulation by dietary interventions, prebiotics, probiotics, and specifically tailored antibiotics. Determining the functions of the microbial clades that expand or contract in SLE will contribute to developing effective strategies to target them. However, the key role of microbiota in SLE pathogenesis and prospective mechanic studies still need to be further investigated.
Conclusions
In the present study, we reveal that intestinal dysbiosis and aberrant metabolism pathways are existed in SLE patients, especially in active SLE patients. Notably, there are four disordered genera and two species that are associated with the clinical disease activity in our patient cohort. Furthermore, the two kinds of genera-panels can be the indicators for diagnosing or monitoring disease activity of SLE by random forest algorithm. However, we also recognize the limitations of our study. Since the results are deduced by a single-center study with a relatively small sample size, larger and prospective cohort studies will be required to verify and validate this predictive model.

Clinical perspectives
- There are limited studies on SLE and gut microbiota, and whether the gut microbiota is associated with the disease activity still remains unclear.
- The SLE patients, especially the active SLE patients, had a distinct dysbiosis in microbiota and its related metabolic pathways. Four disordered genera and two species were revealed to be closely associated with SLE activity. Furthermore, the gut microbiota was validated to have strong diagnostic potential for SLE, and even predict the disease activity.
- Our work identified the association between gut microbiota and activity of SLE, which provided new insights into the mechanisms of lupus. Moreover, it also demonstrated that intestinal microbes could diagnose SLE and monitor the disease activity, which provided experimental evidence for searching new biomarkers for SLE.

Acknowledgements
We thank the Bioinformatics Consulting from the team of Prof. Hongwei Zhou in Division of Laboratory Medicine, Microbiome Medicine Center, Zhujiang Hospital, Southern Medical University. We also thank Miss Carly Bunston (University of Plymouth Faculty of Medicine and Dentistry, Plymouth, U.K.) for constructive comments and editing the manuscript.

Competing interests
The authors declare that there are no competing interests associated with the manuscript.

Funding
This work was supported by the Natural Science Foundation of Guangdong Province of China [grant number 2016A030313525]; the Guangzhou Science and Technology Program Key Projects [grant number 201604040003]; and the Science and Technology Planning Project of Guangzhou [grant number 201607010015]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contribution
Yao Li, Hai-Fang Wang, Xin Li, Ji-Liang Li, and Yu-Rong Qiu were responsible for the overall design and interpretation of the study. Yu-Rong Qiu and Ji-Liang Li conceived and designed the study. Yao Li, Hai-Fang Wang, Xin Li, Hai-Xia Li, Qiong Zhang, and Chen Fu performed the experiments. Yao Li, Hai-Fang Wang, Xin Li, and Xiao-He Zhang contributed to the experimental data collection and analysis. Hong-Wei Zhou, Yan He, and Pan Li contributed to the bioinformatics analysis. Yao Li, Hai-Fang Wang, and Xin Li wrote the manuscript. All authors read and approved the final manuscript.

Abbreviations
ACR, American College of Rheumatology; ACVD, acute cardiovascular disease; ADONIS, also known as permutational MANOVA; A group, active SLE patients group; anti-dsDNA, anti-double stranded DNA; AUC, area under the curve; CI, confidence interval; CRP, C reactive protein; C3, complement 3; ESR, Erythrocyte Sedimentation Rate; FDR, false discovery rate; F/B, Firmicutes/Bacteroidetes; HC, healthy control; HGB, Firmicutes/Bacteroidetes; IL, Interleukin; KEGG, Kyoto Encyclopedia of Genes and Genomes; LDA, linear discriminant analysis; LEfSe, linear discriminant analysis effect size; NZB/W F1, the hybrid progeny of NZW and NZB mice; OTU, operational taxonomic unit; PCoA, principal coordinates analysis; pDC/IFN, Interferon-derived plasmacytoid Dendritic Cells; PICRUSt, Phylogenetic Investigation of Communities by Reconstruction of Unobserved States; QIIME, Quantitative Insights Into Microbial Ecology; RA, rheumatoid arthritis; RBC, red blood cell; R group,
remissive SLE patients group; ROC, receiver-operating characteristic; SCFAs, short-chain fatty acids; S.D., standard Deviation; SLE, Systemic Lupus Erythematosus; SLEDAI, SLE disease activity index; WBC, white blood cell; 24-UTP, 24-h urine protein.

References


