BRITISH JOURNAL OF HOSPITAL MEDICINE

ARTICLE

Analysis of blood microbiota in patients with adult-onset Still's disease and sepsis by metagenomic next-generation sequencing

Quanquan Fang[1]

Jingzhi Xie[1]

Songlou Yin[1]

Hanqiu Yin[1]

Dongmei Zhou[1]

Author details can be found at the end of this article

Correspondence to:
Dongmei Zhou
(zdm@xzhmu.edu.cn)

Abstract

Aims/Background Adult-onset Still's disease (AOSD) shares similar clinical symptoms with sepsis. Thus, differentiating between AOSD and sepsis presents a great challenge while making diagnosis. This study aimed to analyse the changes in blood microbiota related to AOSD and sepsis using metagenomic next-generation sequencing (mNGS), identify potential biomarkers that distinguish AOSD from sepsis, and explore the diagnostic value of mNGS in differentiation between these two pathological conditions.

Methods Clinical data of four AOSD patients and four sepsis patients treated in the Department of Rheumatology and Immunology, The Affiliated Hospital of Xuzhou Medical University between October 2021 and February 2022 were collected. The mNGS diagnostic records of these patients were analysed for microbial correlations in terms of species taxonomic structure and beta diversity by comparing blood microbiota between AOSD and sepsis. The biomarkers with the strongest capability in distinguishing the subgroups were screened using a random forest algorithm.

Results There was no statistically significant differences between AOSD patients and sepsis controls in terms of gender and age ($p > 0.05$). A total of 91 operational taxonomic units (OTUs) were obtained. At the level of phylum, *Proteobacteria*, *Ascomycota* and *Basidiomycota* were present in high abundances in both groups (79.76%, 14.18% and 3.30% vs 54.03%, 32.77% and 5.81%). At the genus level, the abundances of *Parainfluenzae*, *Aspergillus* and *Ralstonia* were the top three highest in the AOSD group (73.88%, 10.92% and 5.48%), while *Ralstonia*, *Aspergillus* and *Malassezia* were ranked as the top three in the sepsis group in term of abundance (48.69%, 27.36% and 5.52%). In beta-diversity analysis, there were advances shown in visual principal coordinates analysis (PCoA) and non-metric multidimensional scaling (NMDS) between the AOSD group and sepsis group ($p < 0.05$), with little significant differences in the analysis of similarities (Anosim) ($p > 0.05$). Linear discriminant analysis effect size (LEfSe) showed that *Mucoromycota*, *Saccharomycetes*, *Moraxellales*, *Mucorales*, *Xanthomonadales, Saccharomycetales*, *Acinetobacter*, *Stenotrophomonas*, *Yarrowia*, *Apophysomyces*, *Acinetobacter johnson*, *Yarrowia lipolytica*, *Apophysomyces variabilis* and *Stenotrophomonas maltophilia* were more enriched in sepsis group ($p < 0.05$). The top five variables with the strongest capability in distinguishing between AOSD and sepsis were *Acinetobacter johnsonii*, *Apophysomyces variabilis*, *Propionibacterium acnes*, *Stenotrophomonas maltophilia* and *Yarrowia lipolytica*.

Conclusion The blood microorganisms in AOSD were different from sepsis, and mNGS was potential to distinguish between AOSD and sepsis.

Key words: Adult-onset Still's disease; Blood microbiota; Metagenomic next-generation sequencing; Sepsis

Submitted: 26 March 2024; Revised: 29 April 2024; Accepted: 10 May 2024

How to cite this article:
Fang Q, Xie J, Yin S, Yin H, Zhou D. Analysis of blood microbiota in patients with adult-onset Still's disease and sepsis by metagenomic next-generation sequencing. Br J Hosp Med. 2024. https://doi.org/10.12968/hmed.2024.0121

Introduction

Adult-onset Still's disease (AOSD) is a rare systemic disorder characterised by high-spiking fever, skin rash, polyarthritis, sore throat, hyperferritinemia, and neutrophilia (Wang et al, 2019). AOSD is associated with life-threatening complications, such as macrophage activation syndrome (Dunger-Baldauf et al, 2022), which has been reported to occur in up

to 23% of AOSD patients and result in a mortality rate of 20% to 42%. The complexity and heterogeneity of AOSD add further challenges to the diagnostic process, which requires making a differential diagnosis to rule out infections, malignancies and other rheumatic diseases (Hu et al, 2019a), thereby consistently contributing to delayed treatment in patients with AOSD. Therefore, to maximise the recovery chances, early treatment for AOSD is of utmost clinical importance, underscoring the need for biomarkers to facilitate differential diagnosis of AOSD.

In clinical practice, it is often difficult to distinguish systemic inflammatory response syndrome, which is presented as the first symptom in patients with AOSD, from conditions and diseases of infectious origin, especially septicaemia. Sepsis is defined as the presence of bacteria or other toxic products in the blood, accompanied by corresponding clinical manifestations (Srzić et al, 2022). Microbes entering the bloodstream can lead to septicaemia, accompanied by fever, increased white blood cell (WBC) count, and circulatory failure, requiring timely identification and treatment. The occurrence of sepsis is due to the excessive activation of the host defence mechanisms against systemic infection reactions, rather than the direct action of microbes (Nahm et al, 2008). In 2016, the revised Third International Consensus Definition for Sepsis and Septic Shock defined sepsis as 'life-threatening organ dysfunction caused by a dysregulated host response to infection' (Rello et al, 2017). Sepsis represents the immune system's response to injury, following a state of high inflammatory response (an immunosuppressive phase), during which the patient experiences multiple organ dysfunctions and is prone to hospital-acquired infections. It is often challenging to clinically differentiate between AOSD and sepsis, especially in the early stage since fever is the typical, initial clinical manifestation, a clinical feature shared by both AOSD and sepsis, which is also known as 'subacute septicemia.'

Both AOSD and sepsis also have overlapping similarities in clinical features and laboratory findings. Thus, identifying diagnostic markers for these two diseases is crucial for the accurate diagnosis and effective management of AOSD (Tian et al, 2021). At present, the traditional detection method for AOSD mainly lies in pinpointing the pathogenesis. Because of the nature of sampling and detection, the conventional culture method is a time-consuming process and yields less accurate results, especially for thick-walled bacteria, rare bacteria, fungi and viruses. In 40% of sepsis patients, the microbiological culture results are negative, increasing the difficulty in differentiating between the two diseases. In recent years, the emergence and gradual improvement of metagenomic next-generation sequencing (mNGS) has lent tremendous support to the clinical diagnosis of infectious diseases through molecular and genetic approaches. Genetic testing allows for rapid detection of viruses, bacteria, fungi and other pathogens all at the same time (Aletaha et al, 2010), without requiring sequence amplification using specific gene primers, thus shortening the processing time (Maeda et al, 2016). Besides, mNGS yields higher positive detection rate than traditional blood culture (67.74% vs 19.35%) (Sun et al, 2022), making it an important detection technique in the realm of medical microbiology.

The exact aetiology of AOSD remains unknown, but several microorganisms, especially viruses, have been implicated in the pathogenesis of the juvenile form of Still's disease and AOSD. It has been found that cytomegalovirus infection may trigger the onset or relapses of AOSD (Jia et al, 2019). The implication of intestinal flora in the development and regulation of the host immune system has been gaining increasing attention in recent years, and evidence has pointed to the dysregulation of the gut microbiome linked to the disease activity of rheumatic diseases (Bao et al, 2020). Intestinal microorganisms can induce the initiation and regulation of inflammation, playing a crucial role in the pathogenesis of autoimmune diseases (Vétizou et al, 2015). The gut microbiota is more susceptible to various factors, exerting a lesser impact on the blood microbiota compared to the gut microbiota. Additionally, studies on blood microbiota and its association with AOSD remain relatively limited.

In this study, we explored the microbial species composition in the blood of AOSD patients and sepsis controls by means of mNGS. We also further analysed the differences in species composition between the two groups, so as to identify blood microbial markers of AOSD (**Figure 1**). The findings of this study may provide new insights, from the microbiological perspective, on the diagnosis and treatment strategies for AOSD.

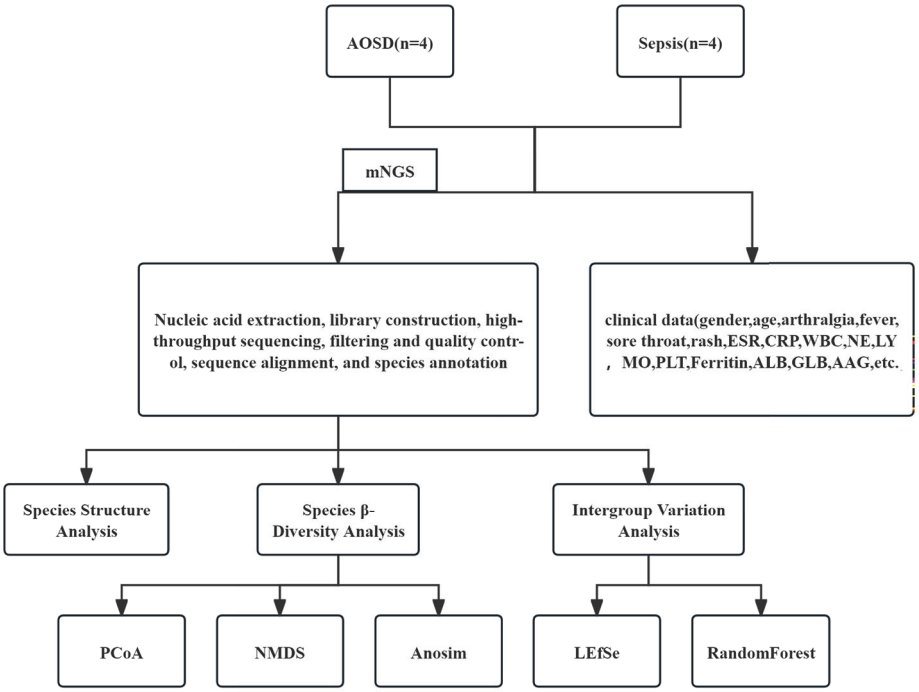

Figure 1. Flowchart depicting the experimental flow of this study. Abbreviations: AOSD: Adult-onset Still's disease; mNGS: metagenomic next-generation sequencing; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; WBC: white blood cell; NE: neutrophil; LY: lymphocyte; MO: monocyte; PLT: platelet; ALB: albumin; GLB: globulin; AAG: α1-acid glycoprotein; PCoA: principal coordinates analysis; NMDS: non-metric multidimensional scaling; Anosim: analysis of similarities; LEfSe: linear discriminant analysis effect size.

Methods

Study participants

Four AOSD patients (cases) and four sepsis patients (controls) who were admitted to the Department of Rheumatology and Immunology, The Affiliated Hospital of Xuzhou Medical University from October 2021 to February 2022 were included.

The AOSD patients who met the following criteria were included: Inclusion criteria for: (1) aged ≥ 18 years; (2) meeting Yamaguchi diagnostic criteria (Yamaguchi et al, 1992); and (3) having no other comorbid infectious and immune diseases. To meet the Yamaguchi diagnostic criteria (Yamaguchi et al, 1992), at least five criteria including two or more primary criteria should be met. The primary criteria include: (1) fever >39°C lasting at least 1 week; (2) arthralgia or arthritis lasting at least 2 weeks; (3) typical rash; and (4) leukocyte count ≥10 × 10^9/L consisting of at least 80% granulocytes. The secondary criteria include: (1) sore throat; (2) splenomegaly/lymphadenopathy; (3) lack of rheumatoid factors or antinuclear antibodies; and (4) impaired liver function. AOSD patients with the characteristics described in the following were excluded from this study: (1) aged < 18 years; (2) having cancer and other autoimmune diseases that may affect haematological parameters or confounding factors of treatment; and (3) having received chemotherapy and glucocorticoids for treatment.

The sepsis patients who met the following criteria were included: (1) aged ≥ 18 years; (2) having symptoms of infectious disease such as fever (temperature > 38°C), headache, chills, meningeal irritation signs, vomiting, convulsions, focal neurological deficits, altered consciousness, or drowsiness; (3) having infectious fever confirmed by blood culture during hospitalization; and (4) having no combined immunodeficiency. Sepsis patients with the following characteristics were excluded: (1) aged < 18 years; (2) having critical illness and acute sepsis; (3) having comorbid cancer and other autoimmune diseases that may affect

haematological parameters or confounding factors of treatment; and (4) having received chemotherapy and glucocorticoids for treatment.

All participants provided written informed consent in accordance with the principles of the Helsinki Declaration. This study was approved by the Institutional Ethics Committee of the Affiliated Hospital of Xuzhou Medical University (XYFY2022-KL275).

Data collection

The clinical data of the hospitalized patients were gleaned through the electronic medical record system, encompassing age, erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), white blood cell (WBC) count, neutrophil (NE) count, lymphocyte (LY) count, monocyte (MO) count, platelet (PLT), thrombocytocrit (PCT); albumin (ALB), globulin (GLB), ferritin, and α1-acid glycoprotein (AAG).

Sample collection

Fresh blood (3–6 mL) of each hospitalized patient was collected in a sterile collection tube containing a nucleic acid protectant and transported to MedcareDx (Shenzhen, China). metagenomic next-generation sequencing processes, covering nucleic acid extraction, library construction, high-throughput sequencing, filtering and quality control, sequence alignment, and species annotation, were all completed by a hospital-affiliated company, and the reports were then returned to the hospital system. The sample data were obtained from hospital reports.

Statistical analysis

IBM SPSS Statistics (Version 27.0, International Business Machines Corporation, Armonk, NY, USA) and R software (Version 4.3.1, Lucent, Vienna, Austria) were used for analysing the data. Data following a normal distribution are expressed as mean ± standard deviation. The *t*-test was used for comparing the data between groups. Non-normally distributed continuous variables are expressed as medians and interquartile ranges, and the Mann-Whitney U test is used. Species structure grouping analysis from macrogenome sequencing feedback data was performed using the online MicrobiomeAnalyst system (Version 2.0, Xia Lab, Canada, available at https://www.microbiomeanalyst.ca/MicrobiomeAnalyst/home.xhtml), and beta-diversity analysis was performed on the R package vegan (Version 2.6-4, WU Vienna, San Francisco, CA, USA). Results with $p < 0.05$ were considered statistically significant. The most important species categories capable of discriminating AOSD from sepsis were selected by random forest package (Version4.7-1.1, WU Vienna, San Diego, CA, USA).

Results

Clinical data

The patients in the AOSD group were aged 46.75 ± 12.09 years, while those in the sepsis group were aged 61.75 ± 15.04 years; there was no statistical difference in age between the two groups ($p > 0.05$). In the AOSD group, the AAG levels were significantly higher than those in the sepsis group ($p < 0.05$). Table 1 presents the comparisons of clinical characteristics between AOSD and sepsis groups.

Operational taxonomic units and species annotation

Whole-genome sequencing of the collected samples was performed on the Illumina/UW platform, and quality control optimisation was performed after obtaining the raw data. The filtered sequences were compared with the reference database to remove chimeric sequences to obtain clean data, where species operational taxonomic units (OTUs) clustering analysis and species classification annotation were performed. Since the inter-individual diversity was too great for direct comparison, the number of sequences after algorithmic standardisation by introducing parameters such as relative abundance, number of sequences and coverage made the number of sequences of the same broad class of pathogens detected between

different patients and at different times comparable. All sequences were homologously aligned and clustered into OTUs according to a similarity of 97%. Species annotation analysis for each OTU representative sequence was conducted based on the SILVA ribosomal RNA database. In this study, a total of 5476 read counts that were used to construct OTUs were obtained for both groups, with an average count of 684 per sample. A total of 91 OTUs were obtained (Figure 2).

Species structure analysis

The taxonomic results of all samples were identified mainly at the phylum and genus levels. Differences in species composition between samples were compared using cumulative histograms based on the top seven most abundant species.

Table 1. Clinical characteristics between AOSD and sepsis groups

	AOSD (n=4)	Sepsis (n=4)	*t*/Z	*p*
Age (year)	46.75 ± 12.09	61.75 ± 15.04	1.554	0.171
WBC (10^9/L)	12.28 ± 4.04	12.18 ± 4.35	0.034	0.974
NE (10^9/L)	10.05 ± 3.54	9.45 ± 4.49	–0.209	0.841
LY (10^9/L)	1.53 ± 0.61	1.55 ± 0.33	0.072	0.945
MO (10^9/L)	0.67 ± 0.52	0.93 ± 0.21	0.922	0.392
PLT (10^9/L)	209.00 (192.75, 351.25)	188.00 (83.75, 513.50)	–0.289	0.773
PCT (%)	0.25 ± 0.07	0.26 ± 0.20	0.094	0.928
Ferritin (ng/mL)	2076.50 (1087.40, 12183.50)	674 (280.50, 4274.50)	–1.443	0.149
AAG (mg/dL)	241.75 ± 21.65	126.50 ± 90.56	–2.476	0.048
CRP (mg/L)	144.28 ± 93.80	101.25 ± 60.83	–0.770	0.471
ESR (mm/h)	95.75 ± 32.01	48.00 ± 45.48	–1.717	0.137
ALB (g/L)	35.40 ± 5.12	29.30 ± 2.13	–2.200	0.070
GLB (g/L)	27.78 ± 4.63	30.13 ± 5.95	0.624	0.556

Notes: Data following a normal distribution are presented as mean ± standard deviation, while data following a non-normal distribution are expressed as median (interquartile range). Abbreviations: ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; WBC: white blood cell; NE: neutrophil; LY: lymphocyte; MO: monocyte; PLT: platelet; PCT: thrombocytocrit; ALB: albumin; GLB: globulin; AAG: α1-acid glycoprotein.

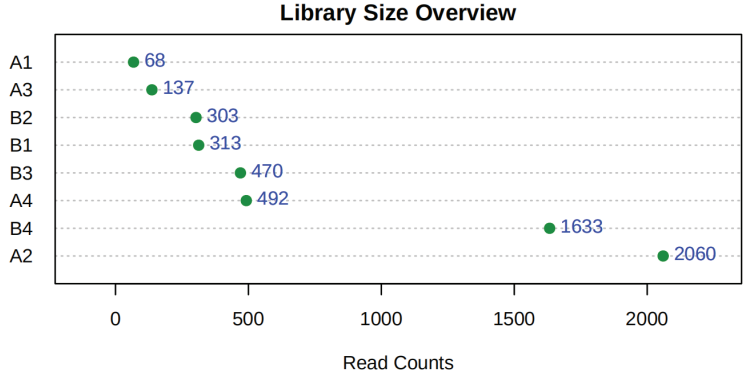

Figure 2. Overview of sample gene sequence counts. Note: Samples are labelled as A1–4 and B1–4.

The relative abundance of species distribution at the phylum level

We selected species with an abundance greater than 0.4% at the phylum level. Through the online MicrobiomeAnalyst system, we identified the top seven species with the highest relative abundance at the phylum level for both groups: *Proteobacteria* (67.23%), *Ascomycota* (23.37%), *Basidiomycota* (4.42%), *Firmicutes* (1.80%), *Bacteroidetes* (1.72%) *Actinobacteria* (1.03%), and *Mucoromycota* (0.44%). Among them, *Proteobacteria*, *Ascomycota*, and *Basidiomycota* claimed the top three places in this ranking for both groups. They accounted for 79.76%, 14.18%, and 3.30% in the AOSD group and 54.03%, 32.77%, and 5.81% in the sepsis group (Figure 3).

In addition, as shown in Figure 4, the relative abundances of the eight samples also showed huge individual differences. The relative abundances of *Proteobacteria* fluctuated from 0.00% to 98.74%, *Ascomycota* from 0.44% to 73.48%, and *Basidiomycota* from 0.00% to 50.00%. The relative abundance of *Bacteroidetes* fluctuated from 0.00% to 26.07%, *Firmicutes* from 0.00% to 11.68%, *Actinobacteria* from 0.00% to 8.82%, and *Mucoromycota* from 0.00% to 1.60%.

The relative abundance of species distribution at the genus level

A total of 54 species were identified at the genus level, and species with abundance greater than 0.1% at the genus level were screened. Through the online MicrobiomeAnalyst system, we identified the top eight species with relative abundance greater than 1.00%: *Parainfluenzae* (37.69%), *Ralstonia* (26.18%), *Aspergillus* (19.21%), *Malassezia* (4.46%), *Debaryomyces* (2.59%), *Streptococcus* (1.81%), *Myroides* (1.42%), and *Cutibacterium* (1.04%). *Parainfluenzae*, *Aspergillus*, and *Ralstonia* were the top three in the ranking for

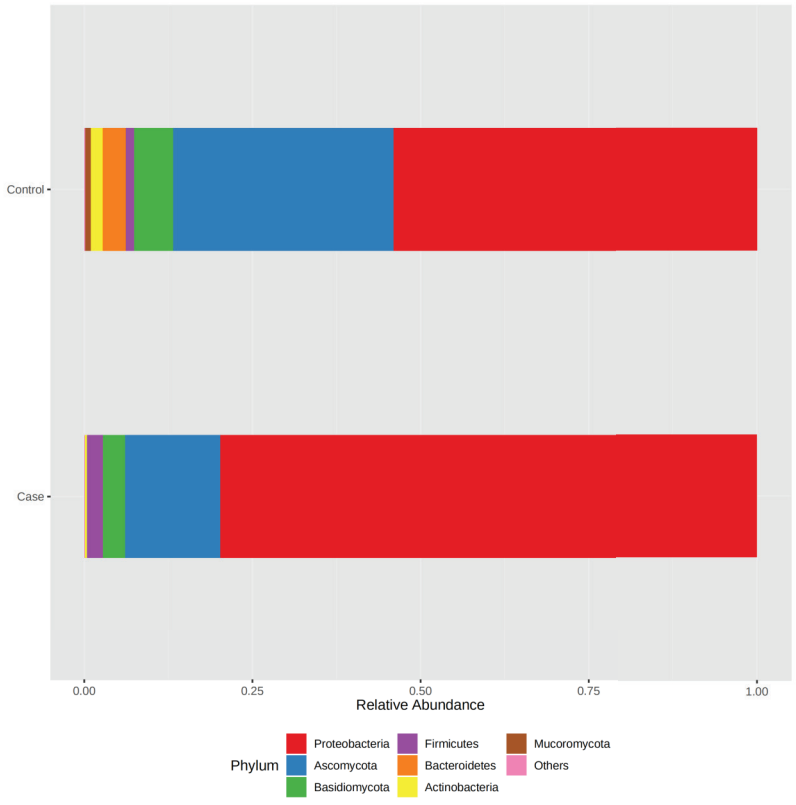

Figure 3. Comparison of the relative abundance between AOSD and sepsis groups at the phylum level. Note: Case, AOSD group; Control, sepsis group.

the AOSD group, accounting for 73.88%, 10.92%, and 5.48%, respectively, while the top three species in terms of relative abundance in the control group were *Ralstonia* (48.69%), *Aspergillus* (27.36%), and *Malassezia* (5.52%) (Figure 5).

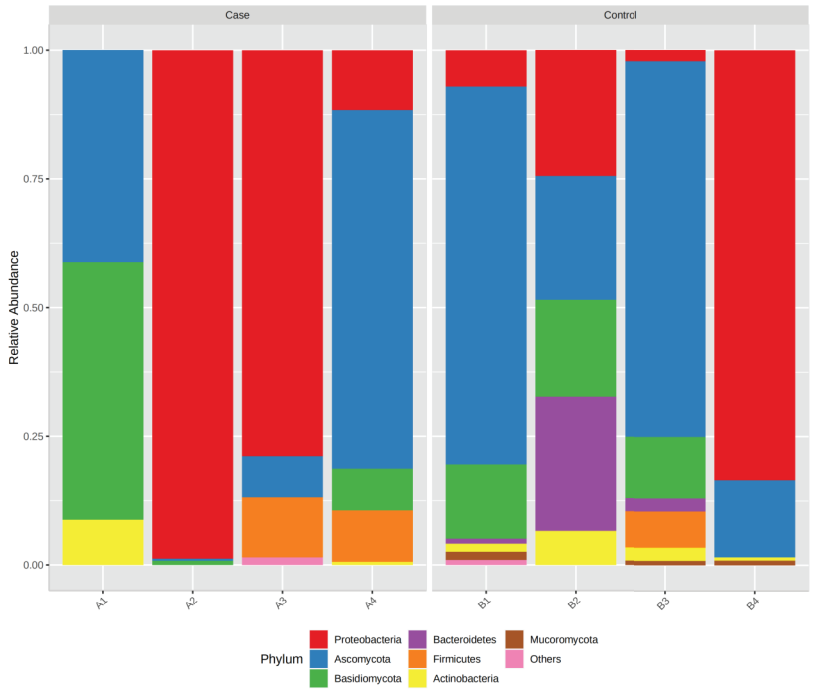

Figure 4. Relative abundance of the eight samples at the phylum level. Note: Samples are labelled as A1–4 and B1–4.

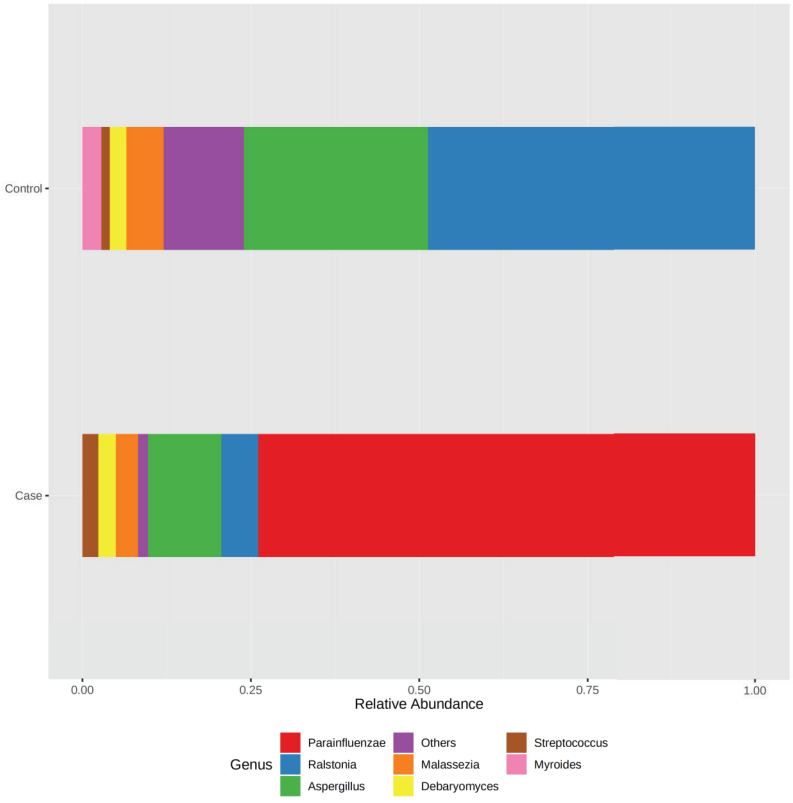

Figure 5. Comparison of the relative abundance between groups at the genus level. Note: Case, AOSD group; Control, sepsis group.

In addition, the relative abundance varied greatly across the samples. Among the top seven species, the relative abundance of *Parainfluenzae* fluctuated from 0.00% to 98.54%, *Ralstonia* from 0.00% to 80.77%, *Aspergillus* from 0.00% to 64.47%, *Malassezia* from 0.00% to 50.00%, *Debaryomyces* from 0.00% to 35.29%, *Streptococcus* from 0.00% to 11.68%, and *Myroides* from 0.00% to 21.45% (Figure 6).

Species beta-diversity analysis

Principal coordinates analysis

Principal coordinates analysis (PCoA) is a data dimensionality reduction processing method based on the distance matrix between samples. This method can present the visual coordinates of similarities or differences in research data and achieve the quantitative transformation of qualitative data (Wang et al, 2022). PCoA plots can be used to observe the differences between individuals or groups and reveal the principal components and contribution rates that contribute to sample variability. A short distance between two groups of samples indicates high structural similarity, whereas a long distance indicates high community variability (Lu et al, 2023). The Canberra distance is a weighted version of the Manhattan distance, introduced and refined in, 1967 by Lance, Williams and Adkins. It is sensitive to values close to zero when data are scattered around an origin. Using the 'Canberra' method for analysis, we obtained an explanation degree of 31.84% in the horizontal coordinate and 16.75% in the vertical coordinate. As shown in Figure 7, both AOSD and sepsis groups can be distinguished from each other, due to the differences in the species richness and diversity of blood microbiota between the two groups.

Non-metric multidimensional scaling

Non-metric multidimensional scaling (NMDS) is a technique used in statistics and data visualisation to analyse similarities or dissimilarities in data. It is a dimensionality reduction method that aims to represent the structure of the data in a lower-dimensional space, typically in two or three dimensions, while preserving the original pairwise distances or dissimilarities between data points as much as possible. To evaluate the quality of NMDS analysis results, the stress coefficient is used. The NMDS plot usually provides the stress

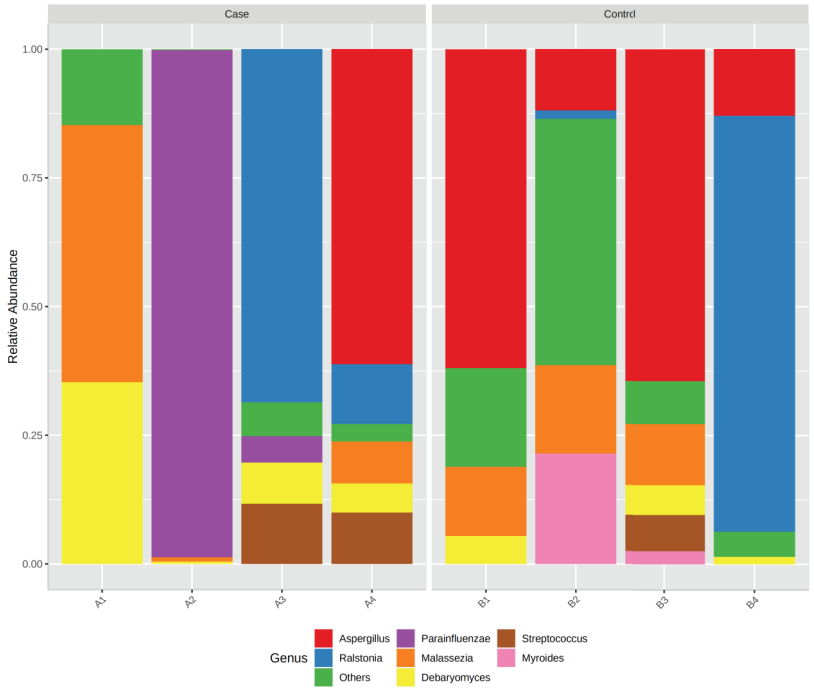

Figure 6. Relative abundance of the eight samples at the genus level. Note: Samples are labelled as A1–4 and B1–4.

value of the model, which is used to determine whether the plot can accurately reflect the true distribution of data rankings. It is generally considered that when stress < 0.2, the results can be represented using a two-dimensional point plot in NMDS, which has certain interpretive significance. When stress < 0.05, it is considered to have good representativeness. In this study, the stress was 0.0456 (< 0.05) for the two groups based on Canberra distance. As shown in Figure 8, the two groups can be distinguished from each other.

Analysis of similarities

Analysis of similarities (Anosim) is a non-parametric test based on permutation test and rank sum test, used to test whether the differences between groups are significantly greater than the differences within groups, thereby determining the significance of grouping. Since the above two visualisation downscaling methods intuitively showed a difference between groups, Anosim was further performed to explore whether the difference between groups is significant. Canberra distance was used to calculate the distance between each pair of samples, then sort all distances from small to large, and calculate the R and *p* values. R value is used to determine if there are differences between different groups, while *p* value is used to indicate the presence of significant differences. Unfortunately, the difference

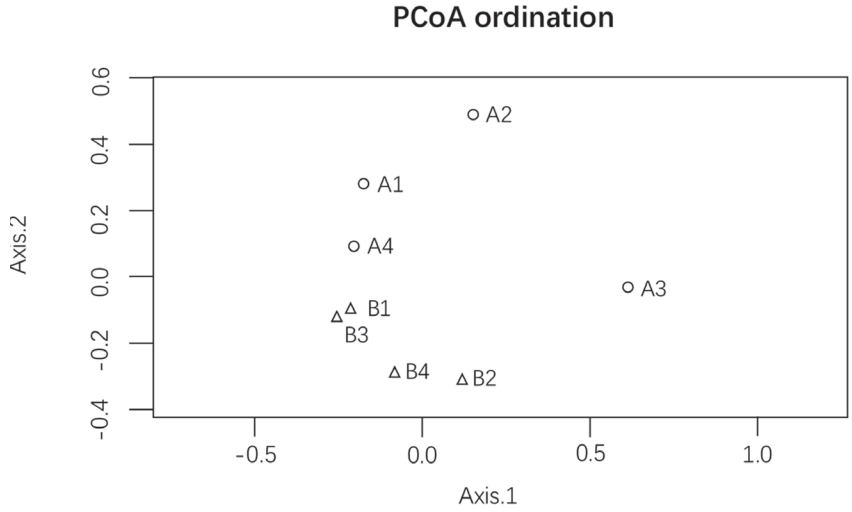

Figure 7. Principal coordinates analysis analysis based on Canberra distance. Note: A1-4, adult-onset Still's disease group; B1-4, sepsis group. Axis.1: Principal Component 1 (PC1), elucidates the greatest variability present in the dataset; Axis.2: Principal Component 2 (PC2), explains the next largest proportion of the remaining variation.

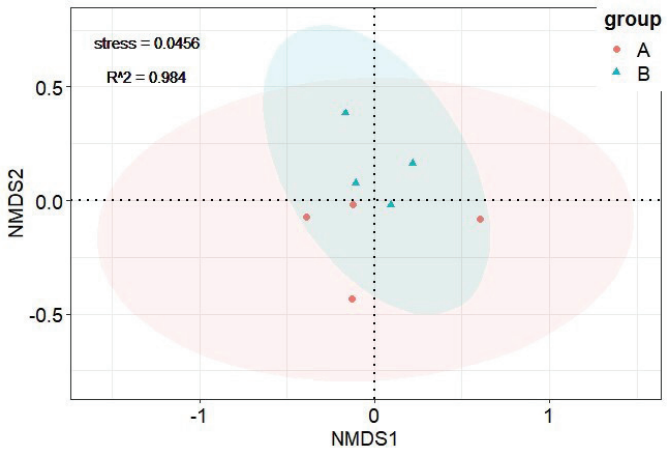

Figure 8. Non-metric multidimensional scaling analysis based on Canberra distance. Note: A, AOSD group; B, sepsis group; NMDS, non-metric multidimensional scaling.

in species abundance between the AOSD group and the spesis group was not statistically significant (R=0.198, p=0.051).

Analysis of intergroup variation in species composition

Linear discriminant analysis effect size (LEfSe)

Linear discriminant analysis effect size (LEfSe) was used for the comparison of colony differences between two or more groups of multiple subgroups for discovering and interpreting biomarkers of high-dimensional data. Species with statistical differences between groups were determined as biomarkers. Linear discriminant analysis effect size analysis in the case and control groups (LDA > 4.0) revealed differences in blood microbiota structure between the two groups. At the phylum level, *Mucoromycota* in the control group were enriched (p=0.047) (Figure 9); at the class level, *Mucoromycetes* and *Saccharomycetes* were enriched in the control group (p=0.047, p=0.047) (Figure 9). More enrichment of *Moraxellales*, *Mucorales*, *Xanthomonadales* and *Saccharomycetales* were found among

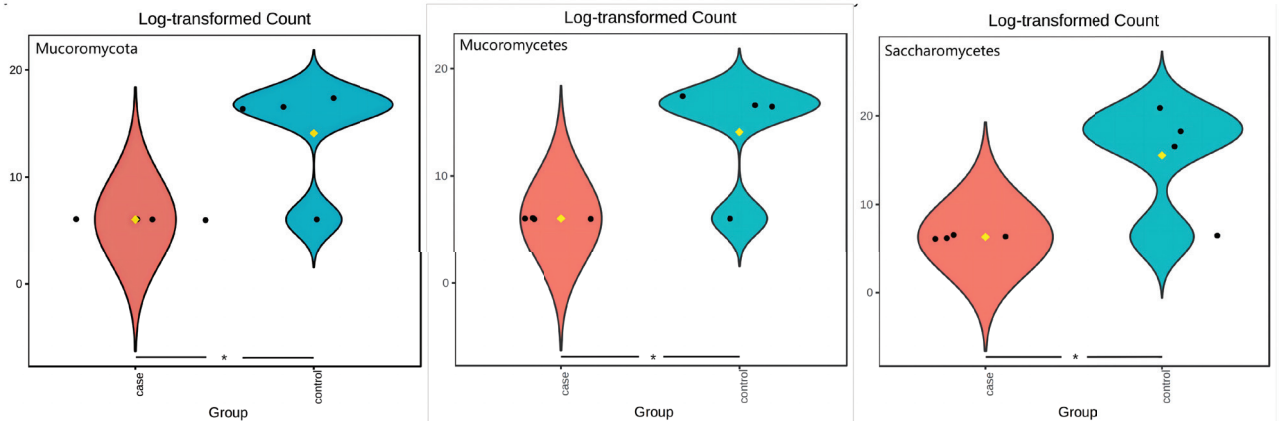

Figure 9. Violin plots of LEfSe analysis of differences at the phylum and class levels. Note: The left panel shows the differences at the phylum level, and the middle and right panels show the differences at the class level. *p < 0.05.

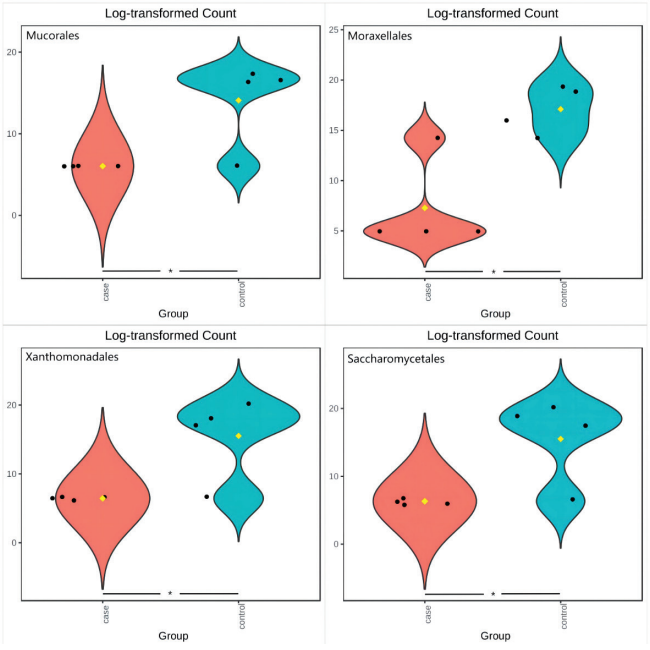

Figure 10. Violin plots of LEfSe analysis of differences at the order level. *p < 0.05.

sepsis patients at the order level (p=0.038, p=0.047, p=0.047, p=0.047) (Figure 10). Intuitively, *Moraxellaceae*, *Mucoraceae*, *Xanthomonadaceae* and *Dipodascaceae* at the family level were enriched in the sepsis patients (p=0.038, p=0.047, p=0.047, p=0.047) (Figure 11). At the genus level, the control group had enriched *Acinetobacter*, *Stenotrophomonas*, *Yarrowia*, and *Apophysomyces* (p=0.038, p=0.047, p=0.047, p=0.047) (Figure 12). The abundances of *Acinetobacter johnsonii*, *Yarrowia lipolytica*, *Apophysomyces variabilis*, and *Stenotrophomonas maltophilia* species also significantly increased in the sepsis group (p=0.014, p=0.047, p=0.047, p=0.047, respectively) (Figure 13). At different taxonomic levels, all differential micriobes exhibited higher enrichment in the sepsis

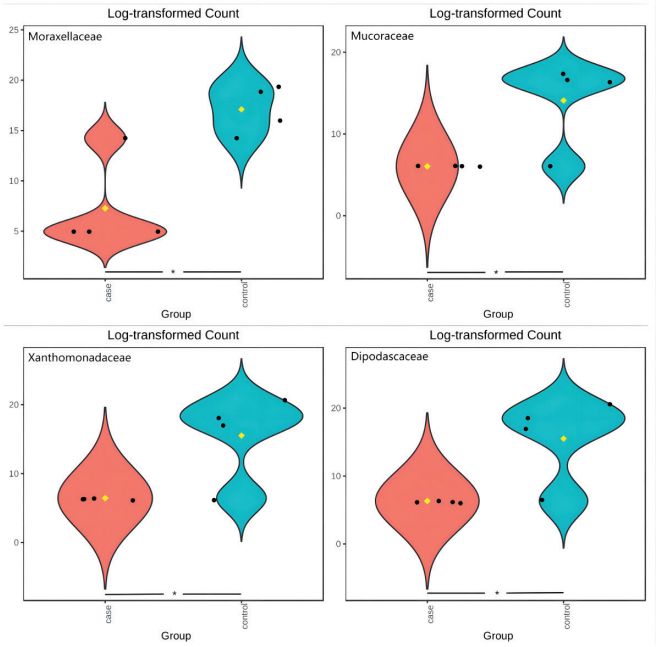

Figure 11. Violin plots of LEfSe analysis of differences at the family level. *p < 0.05.

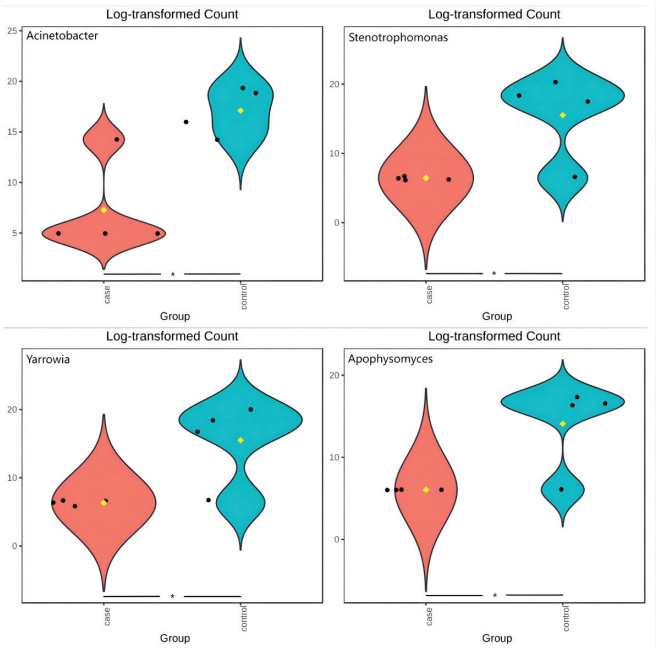

Figure 12. Violin plots of LEfSe analysis of differences at the genus level. *p < 0.05.

group. However, a univariate comparison of all bacterial species was unable to pinpoint any species exhibiting significant differences among the AOSD patients.

Random forest analysis of species with differential flora between groups

Random forest classification is a bagging method for integrated learning that can handle a large number of input variables to produce highly accurate classifiers. The classifiers evaluate the importance of variables when deciding on categories, selecting the most important species categories for classification. The impact of random forest on classification primarily focuses on enhancing classification accuracy, addressing imbalanced datasets, assessing feature importance, managing missing values and outliers, and visualising classification outcomes. Its basic unit is a decision tree, using random splitting with replacement, and each decision tree is a classifier. Assuming that the current focus is on a classification problem, for an input sample, there will be N classification results from N trees. Random forest aggregates all classification voting results and designates the class with the most votes as the final output. We selected the top five variables with a high influence on classification by random forest: *Acinetobacter johnsonii*, *Apophysomyces variabilis*, *Propionibacterium*

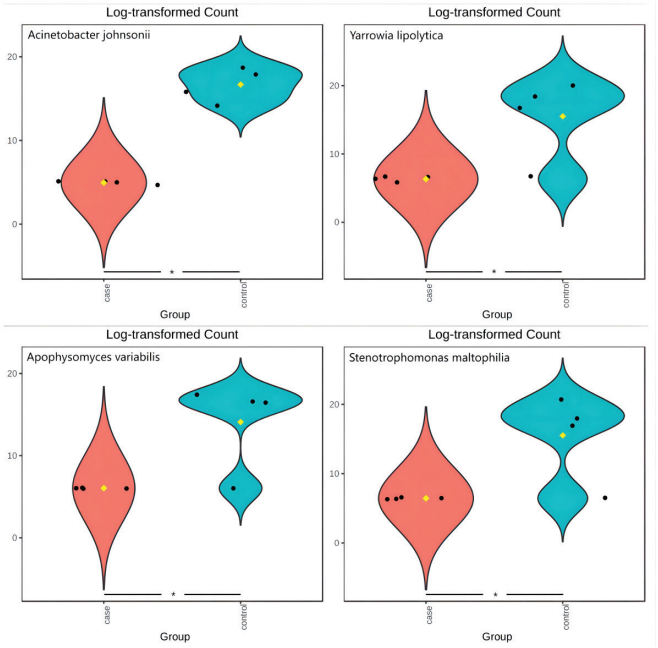

Figure 13. Violin plots of LEfSe analysis of differences at the species level. *$p < 0.05$.

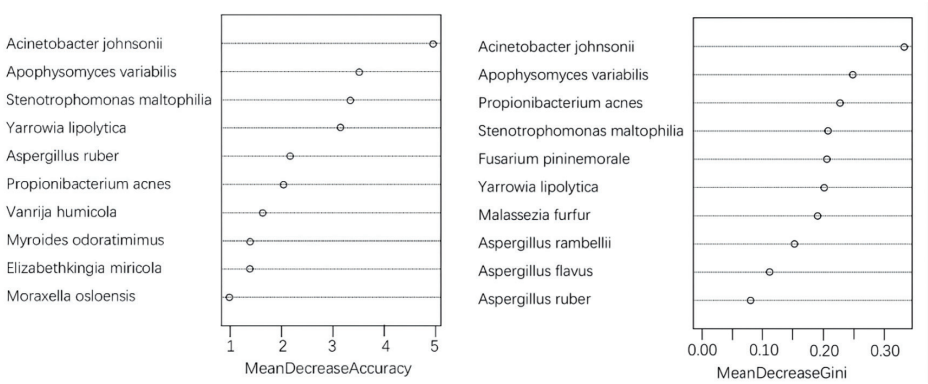

Figure 14. Group discrimination is based on ranking variable importance.

acnes, *Stenotrophomonas maltophilia*, and *Yarrowia lipolytica* (Figure 14). The results are essentially consistent with those generated from the LEfSe analysis.

Ranking variable importance scores presented in the results of random forest analyses are parameters key to predictive accuracy. Larger ranking variable importance scores in the random forest analyses indicate greater importance of the feature. In addition, ranking variable importance reflects observed value heterogeneity.

Discussion

AOSD shares similar symptoms with infectious fever, tumours and other rheumatic diseases. Exclusion of other inflammatory diseases, especially sepsis (Hu et al, 2019b), is required while making the diagnosis of AOSD, a process that could delay the optimal treatment, increase the risk for recurrence, or promote the chronic course of the disease, thereby affecting the treatment, prognosis, and quality of life of patients. In addition, the acute phase of AOSD may occur concurrently with life-threatening complications, such as macrophage activation syndrome, which is considered to be the most severe complication of the disease characterised by a high mortality rate (Giacomelli et al, 2018). Therefore, performing differential diagnosis to differentiate between AOSD and sepsis is of great clinical importance.

The aetiology of AOSD remains largely unknown. Previous studies speculated that its development may be related to certain infections, genetic factors and immune factors. Infectious agents such as Epstein–Barr virus, coxsackie virus, parvovirus B19, human immunodeficiency virus, cytomegalovirus, hepatitis virus, *Enterobacter jejuni*, mycoplasma and chlamydia can trigger the immune system and subsequently activate a cascade of amplified inflammatory responses, thus inducing AOSD (Gerfaud-Valentin et al, 2014). Recently, the growing body of evidence on the relationship between intestinal flora and autoimmune diseases has shed light on the involvement of gut microbes in autoimmune homeostasis, inducing the initiation and regulation of inflammation, and playing an important role in the pathogenesis of autoimmune diseases (Maeda et al, 2016). In previous studies on microorganisms and autoimmune diseases, Rodriguez-Calvo (2019) found enteroviruses in the Langerhans cells of type 1 diabetic patients, which may be involved in the pathogenesis of type 1 diabetes. He et al (2022) found intestinal butyrate-metabolizing species contribute to bone destruction in rheumatoid arthritis. Shen et al (2022) suggested that IAA altered the composition of the intestinal microbiota composition by increasing the abundances of *Bacteroides*, *Bifidobacterium pseudolongum* and *Mucispirillum schaedleri*, and decreasing the abundances of *Proteobacteria* and *Firmicutes*. It has been reported that *Propionibacterium acnes* may be involved in the pathogenesis of synovitis-acne-pustulosis-hyperostosis-osteitis syndrome, mediating the non-specific activation of cellular immunity, and causing osteolytic damage (Amital et al, 2008; Zimmermann and Curtis, 2016). Singh et al (2021) suggested that *Veillonella parvula* may be a biomarker for dry syndrome after conducting a comparison of the differences in oral microbiota between dry syndrome patients and healthy individuals.

However, the gut flora is influenced by a wide range of factors, including genetics, age, geography, ethnicity, antibiotic use, psychological stress, ageing, diet and health status (Scher et al, 2013; Maeda et al, 2016; Martel et al, 2022), and presents highly diverse patterns among individuals, profoundly affecting the identification of microbiome features associated with disease (He et al, 2022). Compared to the gut microbiota, the blood microbiota is less influenced. However, research on blood microbiota and its association with autoimmune diseases remains relatively scarce. It is the scarcity of literature in this field that prompts us to compare the blood microbiota profiles between patients with AOSD and sepsis in the hope of identifying distinguishing markers for diagnostic purposes.

In this study, at the genus level, the relative abundance of species, including the top three bacteria with the highest abundances, showed high individual differences among the eight samples. The beta-diversity analysis revealed relatively minor differences capable of distinguishing between the two groups. Differential bacteria were identified by comparing the blood microbiome profiles between the AOSD group and the sepsis group. We found that *Acinetobacter johnsonii*, *Yarrowia lipolytica*, *Apophysomyces*

variabilis, and *Stenotrophomonas maltophilia* were enriched in the sepsis group. These results align well with random forest analysis. One study has reported the facilitative role of *Stenotrophomonas maltophilia* in the pathogenesis of systemic lupus erythematosus (Chen et al, 2021). Unfortunately, other differential species have not been reported to be associated with autoimmune system diseases. On a separate note, chronic disruption of the intestinal barrier is known to pave the way for the entry of microbial components into the body, producing systemic hypo-inflammation. On this ground, we speculate that dysbiosis of the gut flora is inherent in patients with AOSD, but no studies have yet to confirm such an association.

There are some limitations in this study. Firstly, information bias existed in the process of sample collection and database analysis. Secondly, this study is a small-sample study. The reliability of the results needs to be further validated in the future by expanding the sample size and adopting a multicenter study design. Thirdly, we were unable to perform deeper functional annotation and enrichment analysis as no raw sequence information of nucleic acids or proteins was obtained. In addition, when interpreting mNGS data, we need to be more rational, as the microorganisms detected by mNGS are background microorganisms coexisting with humans, which does not necessarily mean they are pathogenic. Therefore, the study design and analytic approach need to be further improved and optimised in future studies by excluding confounding factors and utilising a larger sample in order to amplify the clinical significance and relevance of the study findings.

Conclusion

In conclusion, this study demonstrated that there are differences in blood microbial composition between AOSD and sepsis, although microbial pathogens potentially contributing to AOSD development were not identified given the limitations of this study. Nevertheless, the utilisation of mNGS holds promise as a candidate technique delivering an effective avenue for distinguishing between AOSD and other similar infectious diseases.

Key points

- Species taxonomic structure and beta-diversity analyses revealed differences in blood microbiota between patients with AOSD and sepsis.
- This is the first study innovatively profiling and comparing blood microbes between AOSD and sepsis using mNGS.
- mNGS has the potential to differentiate between AOSD and other similar infectious diseases, such as sepsis.

Author details

[1]Department of Rheumatology and Immunology, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, Jiangsu, China

Availability of data and materials

The datasets analysed during the current study are available from the corresponding author upon reasonable request.

Author contributions

QQF: writing, original draft, visualisation, software, formal analysis, data curation. JZX: writing, original draft, investigation, formal analysis, data curation. SLY and HQY: writing, editing, methodology, data curation. DMZ: supervision, project administration, conceptualisation. All authors contributed to important editorial changes in the manuscript.

All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics approval and consent to participate

This study was approved by the Institutional Ethics Committee of the Affiliated Hospital of Xuzhou Medical University (XYFY2022-KL275). All participants provided written informed consent in accordance with the principles of the Helsinki Declaration.

Acknowledgement

Not applicable.

Funding

No funding was received for this article.

Conflict of interest

The authors declare that they have no conflicts of interest for this work.

References

Aletaha D, Neogi T, Silman AJ et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010;62(9):2569–2581. https://doi.org/10.1002/art.27584

Amital H, Govoni M, Maya R et al. Role of infectious agents in systemic rheumatic diseases. Clin Exp Rheumatol. 2008;26(1 Suppl 48):S27–S32. https://pubmed.ncbi.nlm.nih.gov/18570751/

Bao Y, Dong C, Ji J, Gu Z. Dysregulation of gut microbiome is linked to disease activity of rheumatic diseases. Clin Rheumatol. 2020;39(9):2523–2528. https://doi.org/10.1007/s10067-020-05170-9

Chen YL, Yang HT, Liu LX et al. Dorsolateral medullary infarction during skin infection by Stenotrophomonas maltophilia in a patient with triple antiphospholipid antibody positivity: a case-based review. Clin Rheumatol. 2021;40(7):2965–2971. https://doi.org/10.1007/s10067-020-05489-3

Dunger-Baldauf C, Nakasato P, Noviello S, Whelan S, Bieth B. Similar clinical outcomes in patients with systemic juvenile idiopathic arthritis and adult-onset Still's disease treated with canakinumab: Bayesian and population model-based analyses. Rheumatol Ther. 2022;9(2):753–762. https://doi.org/10.1007/s40744-021-00422-9

Gerfaud-Valentin M, Jamilloux Y, Iwaz J, Sève P. Adult-onset Still's disease. Autoimmun Rev. 2014;13(7):708–722. https://doi.org/10.1016/j.autrev.2014.01.058

Giacomelli R, Ruscitti P, Shoenfeld Y. A comprehensive review on adult onset Still's disease. J Autoimmun. 2018;93:24–36. https://doi.org/10.1016/j.jaut.2018.07.018

He J, Chu Y, Li J et al. Intestinal butyrate-metabolizing species contribute to autoantibody production and bone erosion in rheumatoid arthritis. Sci Adv. 2022;8(6):eabm1511. https://doi.org/10.1126/sciadv.abm1511

Hu QY, Zeng T, Sun CY et al. Clinical features and current treatments of adult-onset Still's disease: a multicentre survey of 517 patients in China. Clin Exp Rheumatol. 2019a;37 Suppl 121(6):52–57. https://pubmed.ncbi.nlm.nih.gov/31573475/

Hu Q, Gong W, Gu J et al. Plasma microRNA profiles as a potential biomarker in differentiating adult-onset Still's disease from sepsis. Frontiers in Immunology. 2019b;9:3099. https://doi.org/10.3389/fimmu.2018.03099

Jia J, Shi H, Liu M et al. Cytomegalovirus infection may trigger adult-onset Still's disease onset or relapses. Front Immunol. 2019;10:898. https://doi.org/10.3389/fimmu.2019.00898

Lu Y, Zhou G, Ewald J et al. MicrobiomeAnalyst 2.0: comprehensive statistical, functional and integrative analysis of microbiome data. Nucleic Acids Res. 2023;51(W1):W310–W318. https://doi.org/10.1093/nar/gkad407

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(11):2646–2661. https://doi.org/10.1002/art.39783

Martel J, Chang SH, Ko YF et al. Gut barrier disruption and chronic disease. Trends Endocrinol Metab. 2022;33(4):247–265. https://doi.org/10.1016/j.tem.2022.01.002

Nahm CH, Choi JW, Lee J. Delta neutrophil index in automated immature granulocyte counts for assessing disease severity of patients with sepsis. Ann Clin Lab Sci. 2008;38(3):241–246. https://pubmed.ncbi.nlm.nih.gov/18715852/

Rello J, Valenzuela-Sánchez F, Ruiz-Rodriguez M, Moyano S. Sepsis: a review of advances in management. Adv Ther. 2017;34(11):2393–2411. https://doi.org/10.1007/s12325-017-0622-8

Rodriguez-Calvo T. Enterovirus infection and type 1 diabetes: unraveling the crime scene. Clin Exp Immunol. 2019;195(1):15–24. https://doi.org/10.1111/cei.13223

Scher JU, Sczesnak A, Longman RS et al. Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. Elife. 2013;2:e01202. https://doi.org/10.7554/eLife.01202

Shen J, Yang L, You K et al. Indole-3-acetic acid alters intestinal microbiota and alleviates ankylosing spondylitis in mice. Front Immunol. 2022;13:762580. https://doi.org/10.3389/fimmu.2022.762580

Singh M, Teles F, Uzel NG, Papas A. Characterizing Microbiota from Sjögren's Syndrome Patients. JDR Clin Trans Res. 2021;6(3):324–332. https://doi.org/10.1177/2380084420940623

Srzić I, Nesek Adam V, Tunjić Pejak D. Sepsis definition: what's new in the treatment guidelines. Acta Clin Croat. 2022;61(Suppl 1):67–72. https://doi.org/10.20471/acc.2022.61.s1.11

Sun L, Zhang S, Yang Z et al. Clinical application and influencing factor analysis of metagenomic next-generation sequencing (mNGS) in ICU patients with sepsis. Front Cell Infect Microbiol. 2022;12:905132. https://doi.org/10.3389/fcimb.2022.905132

Tian R, Chen X, Yang C et al. Serum heparin-binding protein as a potential biomarker to distinguish adult-onset Still's disease from sepsis. Frontiers Immunol. 2021;12:654811. https://doi.org/10.3389/fimmu.2021.654811

Vétizou M, Pitt JM, Daillère R et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science. 2015;350(6264):1079–1084. https://doi.org/10.1126/science.aad1329

Wang MY, Jia JC, Yang CD, Hu QY. Pathogenesis, disease course, and prognosis of adult-onset Still's disease: an update and review. Chin Med J (Engl). 2019;132(23):2856–2864. https://doi.org/10.1097/CM9.0000000000000538

Wang Y, Sun F, Lin W, Zhang S. AC-PCoA: Adjustment for confounding factors using principal coordinate analysis. PLoS Comput Biol. 2022;18(7):e1010184. https://doi.org/10.1371/journal.pcbi.1010184

Yamaguchi M, Ohta A, Tsunematsu T et al. Preliminary criteria for classification of adult Still's disease. J Rheumatol. 1992;19(3):424–430. https://pubmed.ncbi.nlm.nih.gov/1578458/

Zimmermann P, Curtis N. Synovitis, acne, pustulosis, hyperostosis, and osteitis (SAPHO) syndrome - a challenging diagnosis not to be missed. J Infect. 2016;72 Suppl:S106–S114. https://doi.org/10.1016/j.jinf.2016.04.030