## 2. Materials and methods

### 2.1. Participants

Participants were recruited from Shanghai Mental Health Center in Shanghai, China. Patients with anorexia nervosa who met the diagnostic criteria for anorexia nervosa outlined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; Arlington, VA, 2013), with first-episode presentation and no prior medication history, were included in the study. Participants (both patients and controls) were excluded if they exhibited genetic metabolic diseases, brain injury, acute physical illness, recent antibiotic use (within 2 weeks prior to enrollment), or current medication for gastrointestinal disorders. Healthy controls were recruited via social media platforms, matched to patients by age and sex, and confirmed to be unrelated to the patient cohort. Dietary and lifestyle patterns were assessed using a structured questionnaire adapted from our established methodology, comprising over 50 items addressing body mass index (BMI), nutritional habits, lifestyle behaviors, and mental health parameters. The study protocol adhered to the ethical principles of the Declaration of Helsinki and received approval from the Institutional Review Board of Shanghai Mental Health Center (Approval 伦理批号--------).

### 2.2. Sample collection

Prior to sample collection, researchers provided participants with disposable gloves and sterile fecal collection tubes. Participants were instructed to defecate into a designated commode, followed by handwashing and donning of gloves to aseptically collect the mid-portion fecal sample using standardized protocols. Immediately after collection, samples were placed in a pre-cooled dry ice container (−80 °C) for temporary preservation. All specimens were transported within 24 hours to the laboratory for long-term storage in a −80 °C cryogenic freezer. Genomic DNA extraction was performed using the StoolGen DNA Isolation Kit (CWBiotech Co., Ltd., Beijing, China) following manufacturer specifications.

### 2.3. Shotgun metagenome sequencing

Shotgun metagenome sequencing was performed as described and detailed in the Supplementary Appendix. Brieffy, reads with poor quality or containing human sequences were ffltered out, and MEGAN5 (Huson et al., 2007) was used for metagenomic proffling at the taxonomic and functional levels.

2.4. Gut microbiota-associated epitope (ME) analysis

2.5. Diversity analysis

Gut microbiota diversity and ME alpha diversity were analyzed using the vegan package diversity function. Chao 1 diversity was computed using the vegan package estimateR function; the diversity index was demonstrated using the boxplot function in the R package. The Wilcoxon rank-sum test in R’s wilcox.test function was used to determine the difference in diversity indexes between groups, and signiffcance was deffned as: \*p < 0.05, \*\*p < 0.01, or \*\*\*p < 0.001. To analyze the gut microbiota diversity and ME beta diversity, a nonmetric multidimensional scaling (NMDS) analysis was performed using the metaMDS function in the vegan package.

2.7. Metagenome-wide association study

A metagenome-wide association study (MWAS) was performed as previously described (Qin et al., 2012; Wang et al., 2019). The criteria for signiffcant differences in gut microbial compositions at the taxonomic level were as follows: average relative abundance >0.01%, coverage >0.5, false discovery rate (FDR) corrected p value ≤0.05 for both the Wilcoxon rank-sum test and the Deseq2 test, and absolute value of log2 (fold change) ≥0.2. The criteria for signiffcant differences in gut microbial composition at the functional level was as follows: average relative abundance >0.05%, coverage >0.5, FDR corrected p value ≤0.05 for both the Wilcoxon rank-sum test and Deseq2 test, and absolute value of log2 (fold change) ≥0.2. The criteria for signiffcant differences in MEs was as follows: average relative abundance >0.03%, coverage >0.5, FDR corrected p value ≤0.05 for Fisher's exact test.

2.8. Correlation and regression analyses

To determine the correlation between gut microbiota taxonomies and functional pathways or MEs, the Spearman rank correlation coefffcient was calculated using the cor and cor.test functions in R. The correlation coefffcient and p values were calculated using the cor.test function in R. An FDR-corrected p < 0.05 was considered signiffcant. To determine the linear relationship between diversity index and clinical indices, such as IgA, linear regression analysis was performed using the lm and cor test functions in R, and P values

2.9. Microbial dysbiosis analysis

To determine the degree of microbial dysbiosis, the microbial dysbiosis index (MD index) was computed per Gevers et al. (2014). Here, the MD index was determined as the log of [total abundance of organisms increased in schizophrenia patients]/[total abundance of organisms decreased in controls] for all samples.

2.10. Receiver operating characteristic analysis

To determine whether metagenomic compositions, such as the MD index, IgA and glutamate synthase (GOGAT) can be used as biomarkers for schizophrenia, scikit-learn, a Python-based machine learning method with an L1-regularized logistic regression model was subjected to regression fft analysis, and the receiver operating characteristic (ROC) with 6-fold cross validation was used, and the area under the curve (AUC) was computed.

2.11. Enzyme-linked immunosorbent assay

Enzyme-linked immunosorbent assay (ELISA) was performed as previously described (Zhou et al., 2018a). Brieffy, stool IgA contents were measured using the Salimetrics® SIgA Indirect Enzyme Immunoassay Kit (Salimetrics, Carlsbad, CA, USA) per the manufacturer's instructions. Glutamate synthase (GOGAT) activity was measured using a glutamate synthase activity assay kit (Cat. No. BC0075, Solarbio Science & Technology Co., Ltd, Beijing, China) per the manufacturer’s instructions.

* **Group\_AN（紫色）**：
  + **高丰度功能类别**：在Group\_AN中，以下功能类别的LDA得分较高，表明它们在该组中的相对丰度显著高于Group\_HC：
    - **Membrane\_transport**：与膜运输相关的功能，可能涉及物质进出细胞的调控。
    - **Signal\_transduction**：信号转导功能，可能涉及细胞间或细胞内的信号传递。
    - **Cellular\_community\_prokaryotes**：原核生物的细胞群落相关功能，可能涉及微生物间的相互作用。
    - **Energy\_metabolism**：能量代谢功能，可能涉及能量的产生和利用。
    - **Cellular\_community\_eukaryotes**：真核生物的细胞群落相关功能。
    - **Sensory\_system**：感觉系统功能，可能涉及对环境刺激的感知。
    - **Endocrine\_system**：内分泌系统功能，可能涉及激素等信号分子的调控。
    - **Environmental\_adaptation**：环境适应功能，可能涉及对特定环境条件的响应。
* **Group\_HC（绿色）**：
  + **高丰度功能类别**：在Group\_HC中，以下功能类别的LDA得分较高，表明它们在该组中的相对丰度显著高于Group\_AN：
    - **Transport\_and\_catabolism**：运输和分解代谢功能，可能涉及物质的摄取和分解。
    - **Cell\_growth\_and\_death**：细胞生长和死亡相关功能，可能涉及细胞周期调控等。
    - **Metabolism\_of\_terpenoids\_and\_polyketides**：萜类和聚酮类化合物的代谢，可能涉及次级代谢产物的合成。
    - **Nucleotide\_metabolism**：核苷酸代谢功能，可能涉及DNA和RNA的合成与修复。
    - **Replication\_and\_repair**：复制和修复功能，可能涉及DNA的复制和损伤修复。
    - **Glycan\_biosynthesis\_and\_metabolism**：糖链的生物合成和代谢，可能涉及糖链的合成与降解。
    - **Translation**：翻译功能，可能涉及蛋白质的合成。
* Group\_AN在膜运输、信号转导、能量代谢等功能类别上表现出更高的相对丰度，可能表明该组在这些生物学过程中更为活跃，而在在运输和分解代谢、细胞生长与死亡、核苷酸代谢等功能类别上表现出更低的相对丰度，可能反映其在物质代谢和细胞维持方面的特点。

**分析结果：**

* **Group\_AN（紫色）**：
  + **高丰度功能类别**：在Group\_AN中，以下功能类别的LDA得分较高，表明它们在该组中的相对丰度显著高于Group\_HC：
    - **ABC\_transporters**：ABC转运蛋白相关功能，可能涉及物质的主动运输。
    - **Two\_component\_system**：双组分系统，通常涉及细菌的环境感应和信号传导。
    - **Quorum\_sensing**：群体感应功能，可能涉及细菌密度依赖的基因表达调控。
    - **Pyruvate\_metabolism**：丙酮酸代谢，可能涉及能量代谢和发酵过程。
    - **Sulfur\_relay\_system**：硫传递系统，可能涉及蛋白质的修饰和折叠。
    - **Base\_excision\_repair**：碱基切除修复，可能涉及DNA损伤修复。
    - **Butanoate\_metabolism**：丁酸代谢，可能涉及短链脂肪酸的产生。
    - **Lysine\_degradation**：赖氨酸降解，可能涉及氨基酸代谢。
    - **Nitrogen\_metabolism**：氮代谢，可能涉及氮源的利用。
    - **Galactose\_metabolism**：半乳糖代谢，可能涉及碳水化合物的利用。
    - **Ascorbate\_and\_aldarate\_metabolism**：抗坏血酸和醛酸代谢，可能涉及抗氧化过程。
    - **Biofilm\_formation\_Pseudomonas\_aeruginosa**：铜绿假单胞菌生物膜形成，可能涉及细菌的群体行为。
    - **Lipoic\_acid\_metabolism**：脂酰辅酶A代谢，可能涉及能量代谢。
    - **Photosynthesis**：光合作用相关功能，可能涉及光能利用。
    - **beta\_Lactam\_resistance**：β-内酰胺酶抗性，可能涉及抗生素抗性。
    - **Methane\_metabolism**：甲烷代谢，可能涉及甲烷的产生或利用。
    - **Apoptosis\_fly**：果蝇凋亡相关功能，可能涉及细胞死亡调控。
    - **Biosynthesis\_of\_siderophore\_group\_nonribosomal\_peptides**：铁载体类非核糖体肽的生物合成，可能涉及铁的获取。
    - **Aminobenzoate\_degradation**：氨基苯甲酸降解，可能涉及芳香族化合物的代谢。
    - **Tyrosine\_metabolism**：酪氨酸代谢，可能涉及氨基酸的转化。
    - **Steroid\_hormone\_biosynthesis**：类固醇激素生物合成，可能涉及激素的产生。
    - **Tryptophan\_metabolism**：色氨酸代谢，可能涉及氨基酸的代谢。
    - **Insulin\_signaling\_pathway**：胰岛素信号通路，可能涉及细胞信号传导。
    - **Toluene\_degradation**：甲苯降解，可能涉及有机溶剂的代谢。
    - **Phenazine\_biosynthesis**：吩嗪生物合成，可能涉及抗生素或色素的产生。
    - **Inflammatory\_mediator\_regulation\_of\_TRP\_channels**：炎症介质对TRP通道的调节，可能涉及感觉传导。
* **Group\_HC（绿色）**：
  + **高丰度功能类别**：在Group\_HC中，以下功能类别的LDA得分较高，表明它们在该组中的相对丰度显著高于Group\_AN：
    - **Penicillin\_and\_cephalosporin\_biosynthesis**：青霉素和头孢菌素生物合成，可能涉及抗生素的产生。
    - **Tetracycline\_biosynthesis**：四环素生物合成，可能涉及抗生素的产生。
    - **Lysosome**：溶酶体相关功能，可能涉及细胞内的降解过程。
    - **D\_Glutamine\_and\_D\_glutamate\_metabolism**：D-谷氨酰胺和D-谷氨酸代谢，可能涉及特殊的氨基酸代谢。
    - **Biofilm\_formation\_Vibrio\_cholerae**：霍乱弧菌生物膜形成，可能涉及细菌的群体行为。
    - **Zeatin\_biosynthesis**：玉米素生物合成，可能涉及植物激素的产生。
    - **Epithelial\_cell\_signaling\_in\_Helicobacter\_pylori\_infection**：幽门螺杆菌感染中的上皮细胞信号传导，可能涉及病原体与宿主的相互作用。
    - **Apoptosis**：凋亡相关功能，可能涉及程序性细胞死亡。
    - **Arginine\_biosynthesis**：精氨酸生物合成，可能涉及氨基酸的产生。
    - **Vitamin\_B6\_metabolism**：维生素B6代谢，可能涉及辅酶的产生。
    - **Protein\_processing\_in\_endoplasmic\_reticulum**：内质网中的蛋白质加工，可能涉及蛋白质的折叠和修饰。
    - **DNA\_replication**：DNA复制相关功能，可能涉及细胞分裂。
    - **Purine\_metabolism**：嘌呤代谢，可能涉及核酸的合成和降解。
    - **Glycerophospholipid\_metabolism**：甘油磷脂代谢，可能涉及细胞膜的组成。
    - **Terpenoid\_backbone\_biosynthesis**：萜类化合物骨架生物合成，可能涉及次级代谢产物的产生。
    - **Phenylalanine\_tyrosine\_and\_tryptophan\_biosynthesis**：苯丙氨酸、酪氨酸和色氨酸生物合成，可能涉及氨基酸的产生。
    - **Thiamine\_metabolism**：硫胺素代谢，可能涉及维生素B1的利用。
    - **Cell\_cycle\_Caulobacter**： Caulobacter的细胞周期，可能涉及细菌的分裂调控。
    - **Lipopolysaccharide\_biosynthesis**：脂多糖生物合成，可能涉及细菌细胞壁的组成。
    - **Homologous\_recombination**：同源重组，可能涉及DNA修复和重组。
    - **Ribosome**：核糖体相关功能，可能涉及蛋白质的合成。

**总结：**

* Group\_AN在ABC转运蛋白、双组分系统、光合作用、抗生素抗性等功能类别上表现出更高的相对丰度，可能表明该组在物质运输、环境感应和能量代谢方面更为活跃。在抗生素生物合成、DNA复制、蛋白质加工等功能类别上表现出更低的相对丰度，可能反映其在次级代谢产物产生和细胞维持方面的特点。

结合两张图的分析，可以看出Group\_AN和Group\_HC在功能组成上有显著的差异，这些差异可能与它们的生物学特性、代谢需求或环境适应策略有关。具体的功能差异需要结合实验背景和研究目的进行更深入的探讨。