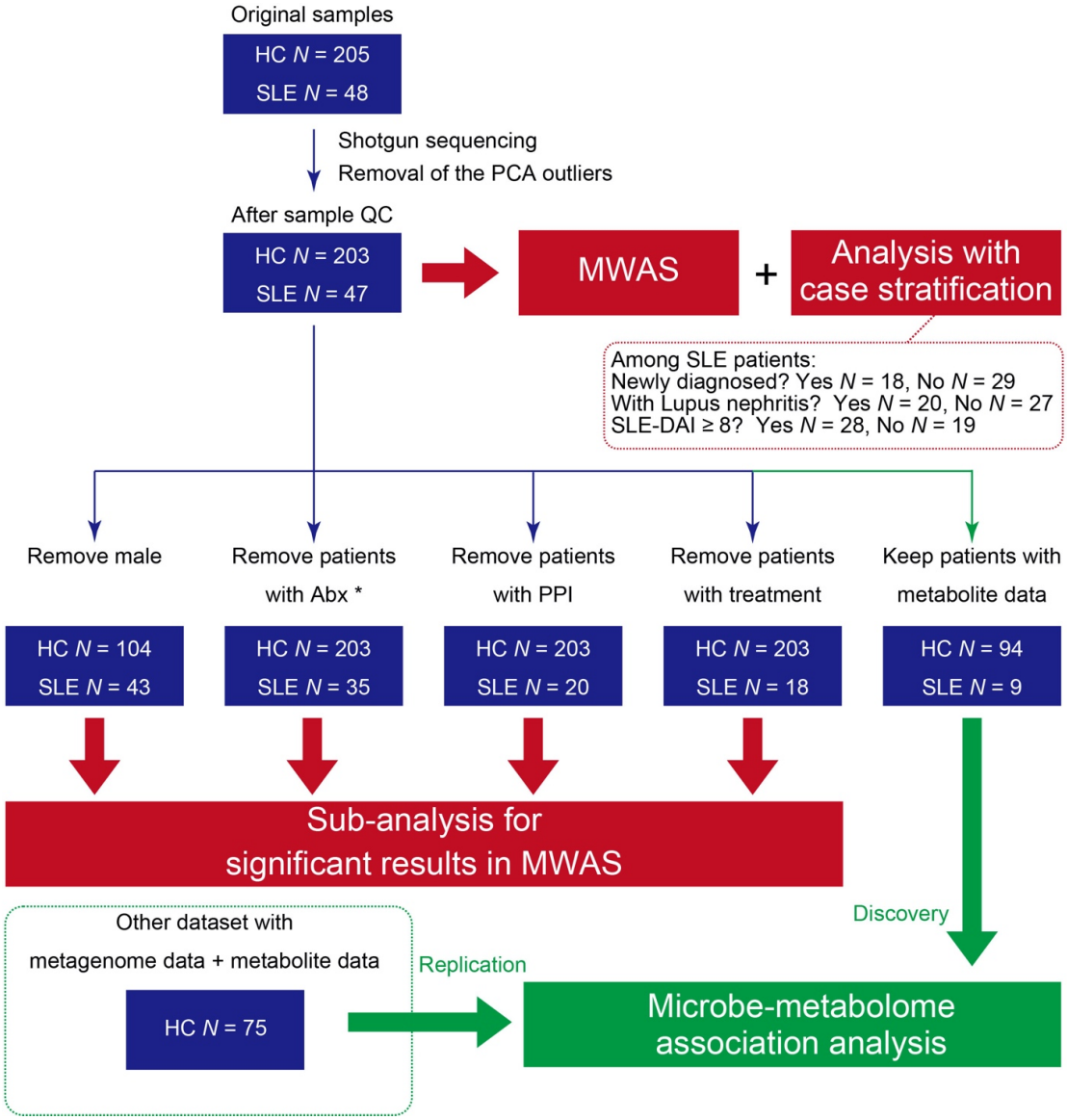
Supplementary information

A metagenome-wide association study revealed disease-specific landscape of the gut microbiome of systemic lupus erythematosus in Japanese

Tomofuji Y et al.

Corresponding to Yukinori Okada (yokada@sg.med.osaka-u.ac.jp)

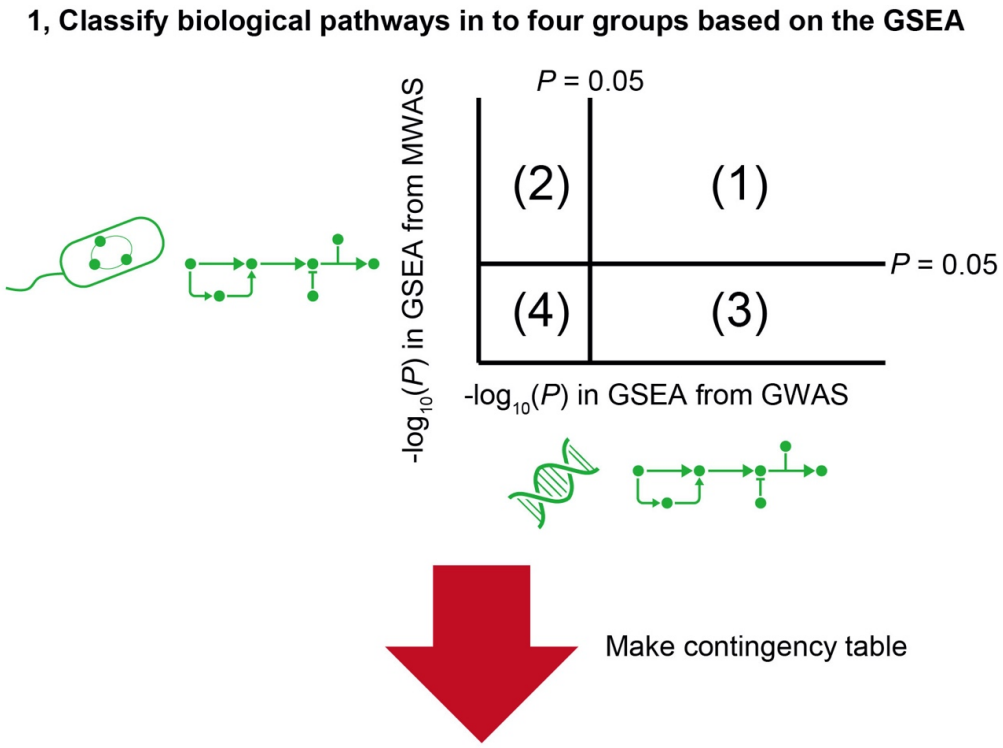
Table of contents



Supplementary Figure 1. A flowchart of the study.

A schematic illustration of the study design. A total of 250 samples (N_{HC} = 203, N_{SLE} = 47) passed the sample QC and were used for the MWAS and analysis with the stratification of the cases. We further extracted subsets of samples for sub-analysis and microbe-metabolite association analysis as described in the figure. Abx, antibiotics; HC, healthy control; MWAS, metagenome-wide association study. PPI, proton pump inhibitor; QC, quality control; SLE, systemic lupus erythematosus. *An SLE patient whose information about antibiotics usage could not be obtained was removed in the sub-analysis.

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687



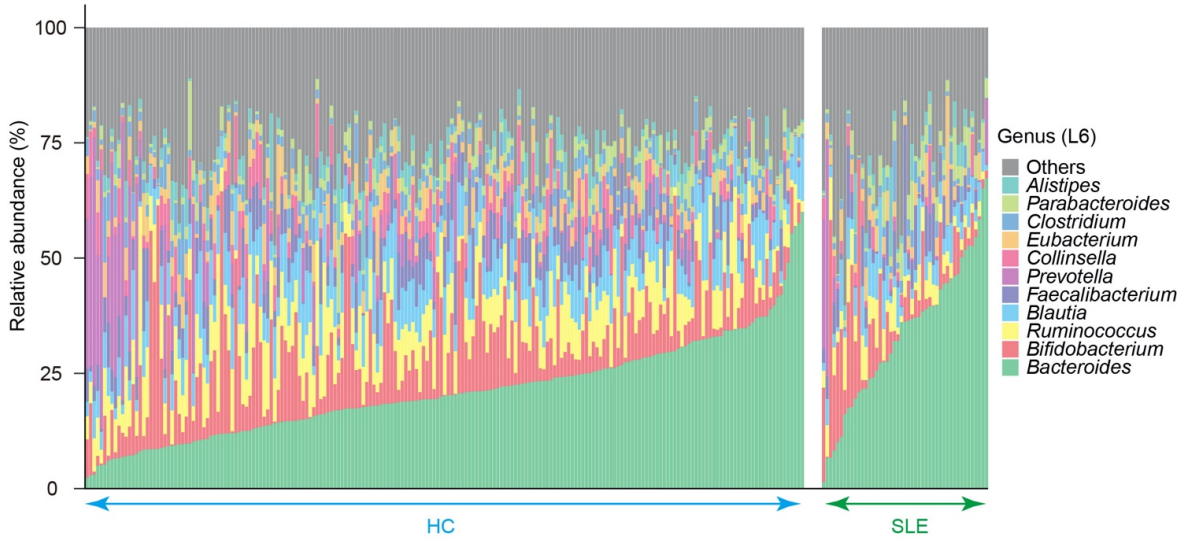
Supplementary Figure 2. A workflow of the MWAS-GWAS interaction analysis.

Firstly, based on the biological pathway enrichment analysis from MWAS and GWAS, we classified biological pathways into the following four groups; (1) Biological pathways enriched in MWAS and GWAS results ($P_{pathway}$ < 0.05 in MWAS and $P_{pathway}$ < 0.05 in GWAS) (2) Biological pathways enriched only in MWAS result ($P_{pathway}$ < 0.05 in MWAS and $P_{pathway}$ ≥ 0.05 in GWAS). (3) Biological pathways enriched only in GWAS result ($P_{pathway}$ ≥ 0.05 in MWAS and $P_{pathway}$ < 0.05 in GWAS). (4) Biological pathways not enriched in MWAS and GWAS results ($P_{pathway}$ ≥ 0.05 in MWAS and $P_{pathway}$ ≥ 0.05 in GWAS). Then,

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

we evaluated the enrichment of the number of biological pathways classified as the group (1) compared to the group (2~4) by one-tailed Fisher's exact test. GSEA, Gene set enrichment analysis; GWAS, Genome-wide association study; MWAS, Metagenome-wide association study.

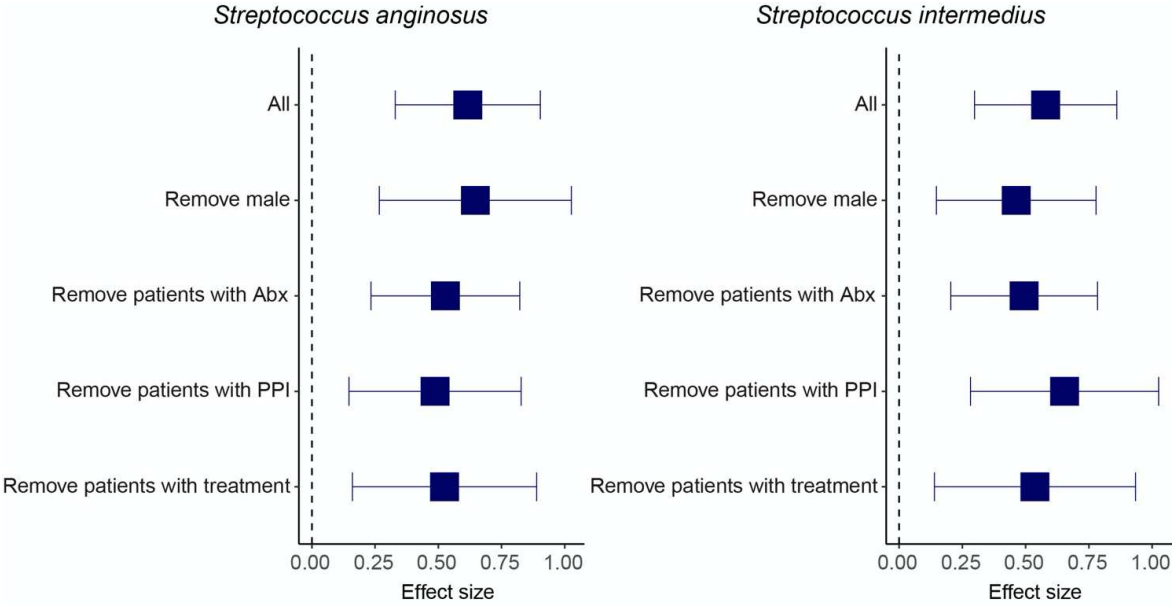
Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687



Supplementary Figure 3. Phylogenetic relative abundance at the genus level (L6).

The relative abundance profiles were constructed utilizing whole genome shotgun sequencing (N_{HC} = 203, N_{SLE} = 47). HC, healthy control; SLE, systemic lupus erythematosus.

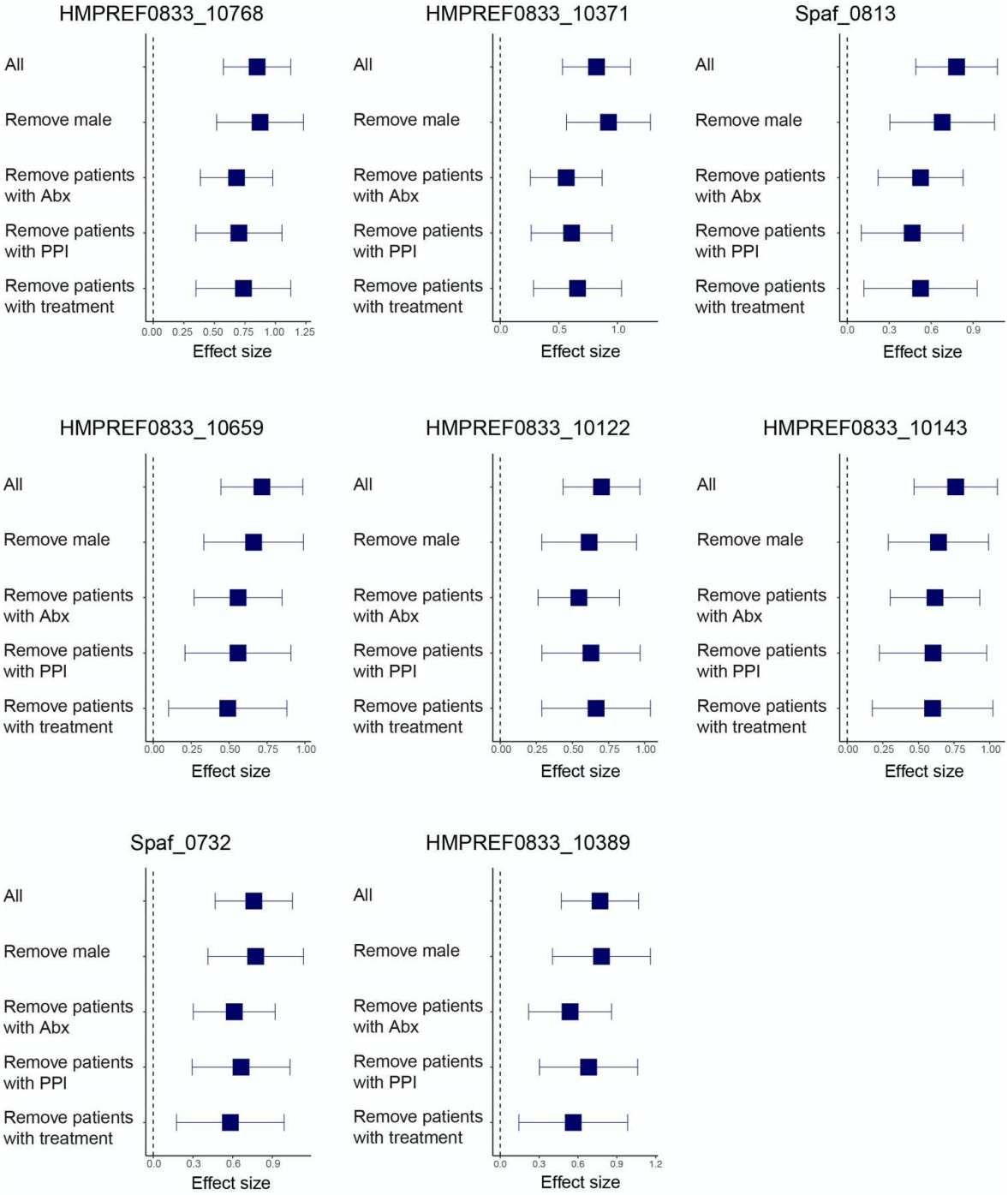
Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687



Supplementary Figure 4. Forest plots from the result of the sub-analysis for the clades with significant SLE case-control discrepancy.

The effect size of the sub-analysis for the clades with significant SLE case-control discrepancy. The boxes indicate the point estimates, and the error bars indicate the 95% confidence interval. Abx, antibiotics; PPI, proton pump inhibitor.

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

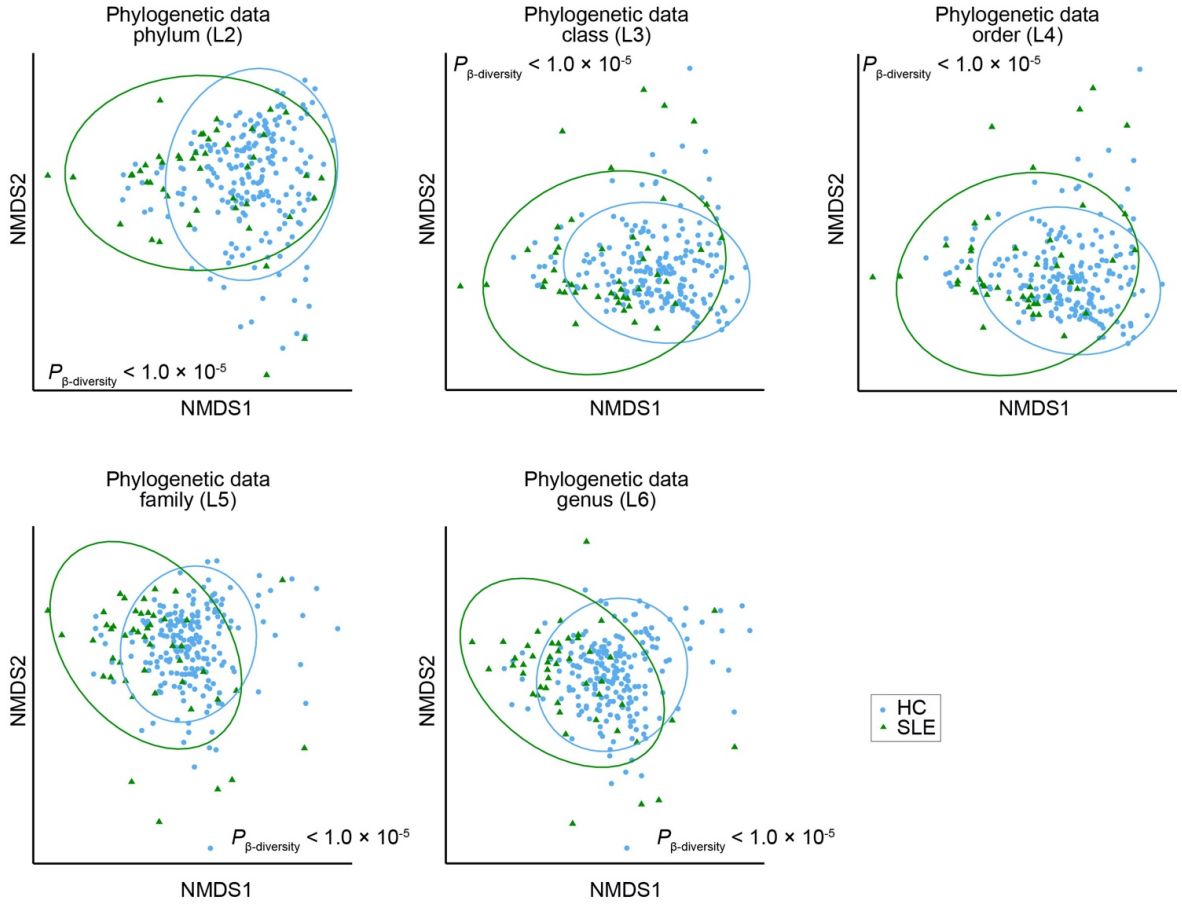


Supplementary Figure 5. Forest plots from the result of the sub-analysis for the genes with significant SLE case-control discrepancy.

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

The effect size of the sub-analysis for the genes with significant SLE case-control discrepancy. The boxes indicate the point estimates, and the error bars indicate the 95% confidence interval. Abx, antibiotics; PPI, proton pump inhibitor.

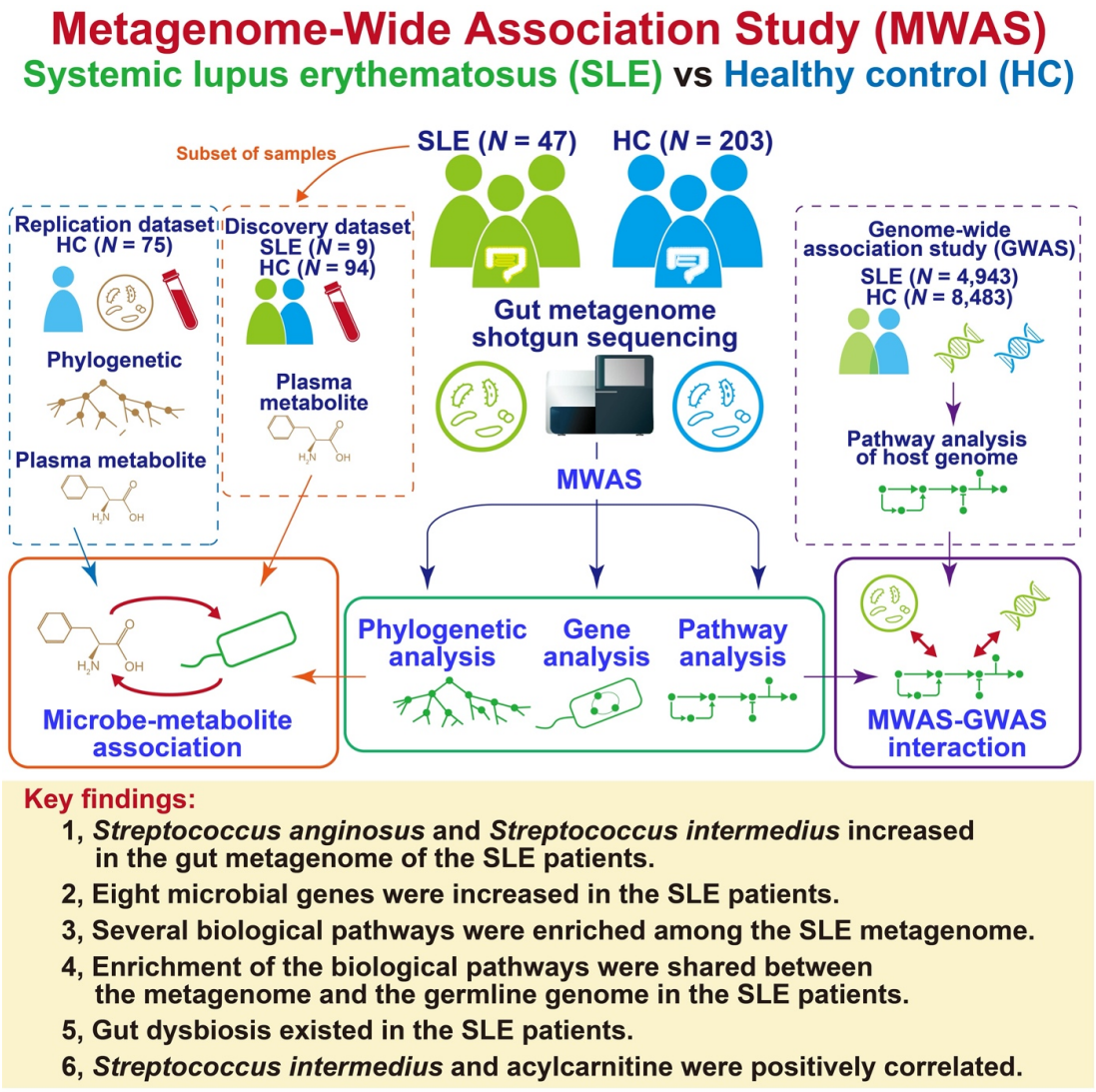
Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687



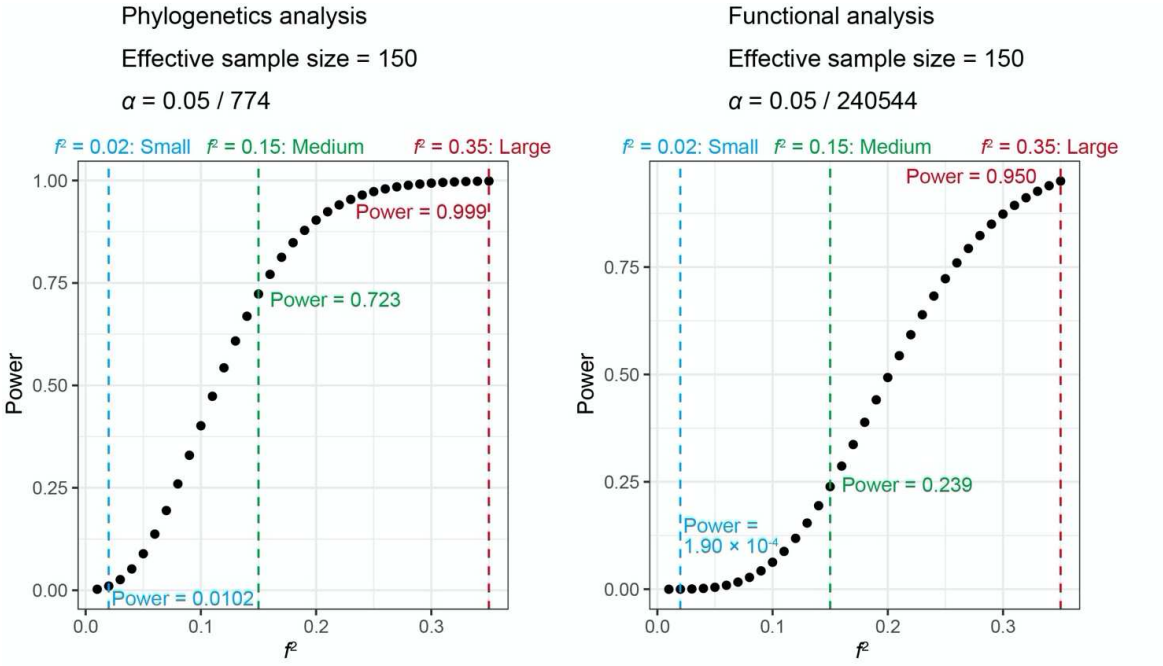
Supplementary Figure 6. β-diversities of the phylogenetic relative abundance data at L2-L6 levels.

The result of NMDS based on Bray-Curtis distance is represented. Blue and green dots represent HC and SLE subjects. HC, healthy control; NMDS, non-metric multidimensional scaling; SLE, systemic lupus erythematosus.

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687



Supplementary Figure 7. Graphical summary of this study.



Supplementary Figure 8. Power calculation of the phylogenetic and functional analysis. Power calculation were performed for phylogenetic and functional analysis. $f^2 \geq 0.02$, $f^2 \geq 0.15$, $f^2 \geq 0.35$ represent small, medium, and large effect sizes, respectively. f^2, Cohen's f^2.

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

Supplementary Table 1. Characteristics of the study population.

	SLE (*N* = 47)	Control (*N* = 203)
Mean age (sd)	42.9 (15.9)	35.4 (12.0)
Sequencing group1	5 (10.6%)	27 (13.3%)
Sequencing group2	10 (21.3%)	71 (35.0%)
Female	43 (91.5%)	104 (51.2%)
Newly diagnosed	18 (38.3%)	-
Not newly diagnosed	29 (61.7%)	-
Treatment		-
Non-treatment	14 (29.8%)	-
Steroid	32 (68.1%)	-
Hydroxychloroquine	7 (14.9%)	-
FK506	5 (10.6%)	-
Ciclosporin	1 (2.1%)	-
Mycophenolate mofetil	8 (17%)	-
Cyclophosphamide	1 (2.1%)	-
Antibiotics use [*1]	11 (23.4%)	-
Proton pump inhibitor use	27 (57.4%)	-
Lupus nephritis	20 (42.6 %)	-
Mean SLE-DAI (sd)	11.1 (8.7)	-

sd, standard deviation; SLE, systemic lupus erythematosus; SLE-DAI, SLE Disease Activity Index.

[*1] Information about antibiotics usage could not be obtained for an SLE patient.

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

Supplementary Table 2. Result of the sub-analysis for the clades with significant SLE case-control discrepancy.

Streptococcus anginosus

Case-control comparison			
Sample set	Effect size	SE	$P_{microbe}$
All	0.617	0.146	3.7×10^{-5}
Remove male	0.647	0.194	0.0011
Remove patients with antibiotics	0.528	0.150	5.4×10^{-4}
Remove patients with proton pump inhibitor	0.487	0.174	0.0056
Remove patients with treatment for SLE	0.525	0.186	0.0053
comparison within the case			
Variable	Effect size	SE	$P_{microbe}$
Not newly diagnosed	0.373	0.705	0.60
Lupus nephritis	0.046	0.421	0.92
SLE-DAI ≥ 8	0.029	0.037	0.45

Streptococcus intermedius

Case-control comparison			
Sample set	Effect size	SE	$P_{microbe}$
All	0.579	0.143	7.5×10^{-5}
Remove male	0.463	0.161	0.0048
Remove patients with antibiotics	0.494	0.148	0.0010
Remove patients with proton pump inhibitor	0.654	0.190	7.0×10^{-4}
Remove patients with treatment for SLE	0.537	0.203	0.0088
comparison within the case			
Variable	Effect size	SE	$P_{microbe}$
Not newly diagnosed	-0.522	0.578	0.38
Lupus nephritis	0.415	0.335	0.23
SLE-DAI ≥ 8	0.0475	0.029	0.12

SE, standard error; SLE, systemic lupus erythematosus.

Supplementary Table 3. Result of the sub-analysis for the genes with significant SLE case-control discrepancy.

Spaf_0732

Case-control comparison			
Sample set	Effect size	SE	P_{KEGG}
All	0.760	0.149	7.1×10^{-7}
Remove male	0.774	0.184	5.2×10^{-5}
Remove patients with antibiotics	0.612	0.158	1.4×10^{-4}
Remove patients with proton pump inhibitor	0.664	0.189	5.4×10^{-4}
Remove patients with treatment for SLE	0.583	0.208	0.0056
comparison within the case			
Variable	Effect size	SE	P_{KEGG}
Not newly diagnosed	0.499	0.449	0.28
Lupus nephritis	0.092	0.265	0.73
SLE-DAI ≥ 8	-0.006	0.021	0.77

Spaf_0813

Case-control comparison			
Sample set	Effect size	SE	P_{KEGG}
All	0.781	0.149	3.6×10^{-7}
Remove male	0.679	0.191	5.5×10^{-4}
Remove patients with antibiotics	0.524	0.155	8.6×10^{-4}
Remove patients with proton pump inhibitor	0.464	0.185	0.013
Remove patients with treatment for SLE	0.524	0.206	0.012
comparison within the case			
Variable	Effect size	SE	P_{KEGG}
Not newly diagnosed	-0.121	0.552	0.83
Lupus nephritis	-0.621	0.275	0.038
SLE-DAI ≥ 8	0.004	0.025	0.87

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

HMPREF0833_10122

Case-control comparison			
Sample set	Effect size	SE	P_{KEGG}
All	0.701	0.135	5.0×10^{-7}
Remove male	0.616	0.167	3.5×10^{-4}
Remove patients with antibiotics	0.544	0.144	2.0×10^{-4}
Remove patients with proton pump inhibitor	0.628	0.173	3.7×10^{-4}
Remove patients with treatment for SLE	0.664	0.192	6.7×10^{-4}
comparison within the case			
Variable	Effect size	SE	P_{KEGG}
Not newly diagnosed	-0.027	0.429	0.95
Lupus nephritis	-0.074	0.244	0.76
SLE-DAI ≥ 8	-0.004	0.020	0.84

HMPREF0833_10143

Case-control comparison			
Sample set	Effect size	SE	P_{KEGG}
All	0.762	0.149	7.0×10^{-7}
Remove male	0.641	0.179	5.1×10^{-4}
Remove patients with antibiotics	0.616	0.160	1.6×10^{-4}
Remove patients with proton pump inhibitor	0.603	0.192	0.0020
Remove patients with treatment for SLE	0.600	0.216	0.0060
comparison within the case			
Variable	Effect size	SE	P_{KEGG}
Not newly diagnosed	-0.147	0.417	0.73
Lupus nephritis	-0.169	0.235	0.48
SLE-DAI ≥ 8	0.005	0.019	0.78

HMPREF0833_10371

Case-control comparison			
Sample set	Effect size	SE	P_{KEGG}
All	0.821	0.148	8.1×10^{-8}
Remove male	0.923	0.182	1.6×10^{-6}
Remove patients with antibiotics	0.563	0.156	3.9×10^{-4}
Remove patients with proton pump inhibitor	0.608	0.176	6.7×10^{-4}
Remove patients with treatment for SLE	0.659	0.192	7.1×10^{-4}
comparison within the case			
Variable	Effect size	SE	P_{KEGG}
Not newly diagnosed	0.113	0.556	0.84
Lupus nephritis	0.135	0.316	0.67
SLE-DAI ≥ 8	0.004	0.026	0.89

HMPREF0833_10389

Case-control comparison			
Sample set	Effect size	SE	P_{KEGG}
All	0.770	0.152	9.1×10^{-7}
Remove male	0.782	0.193	9.1×10^{-5}
Remove patients with antibiotics	0.539	0.163	0.0011
Remove patients with proton pump inhibitor	0.682	0.194	5.3×10^{-4}
Remove patients with treatment for SLE	0.564	0.214	0.0091
comparison within the case			
Variable	Effect size	SE	P_{KEGG}
Not newly diagnosed	-0.091	0.573	0.88
Lupus nephritis	-0.297	0.319	0.37
SLE-DAI ≥ 8	0.000	0.026	1.00

HMPREF0833_10659

Case-control comparison			
Sample set	Effect size	SE	P_{KEGG}
All	0.716	0.138	4.7×10^{-7}
Remove male	0.662	0.168	1.4×10^{-4}
Remove patients with antibiotics	0.560	0.148	2.1×10^{-4}
Remove patients with proton pump inhibitor	0.559	0.178	0.0020
Remove patients with treatment for SLE	0.491	0.199	0.014
comparison within the case			
Variable	Effect size	SE	P_{KEGG}
Not newly diagnosed	0.568	0.409	0.18
Lupus nephritis	-0.223	0.241	0.37
SLE-DAI ≥ 8	-0.019	0.019	0.34

HMPREF0833_10768

Case-control comparison			
Sample set	Effect size	SE	P_{KEGG}
All	0.850	0.141	6.7×10^{-9}
Remove male	0.873	0.181	4.5×10^{-6}
Remove patients with antibiotics	0.681	0.151	1.1×10^{-5}
Remove patients with proton pump inhibitor	0.700	0.180	1.3×10^{-4}
Remove patients with treatment for SLE	0.738	0.198	2.6×10^{-4}
comparison within the case			
Variable	Effect size	SE	P_{KEGG}
Not newly diagnosed	0.266	0.433	0.55
Lupus nephritis	-0.151	0.247	0.55
SLE-DAI ≥ 8	-0.010	0.020	0.61

SE, standard error; SLE, systemic lupus erythematosus; SLE-DAI, SLE Disease Activity Index.

Supplementary Table 4. Pathways with significant SLE case-control discrepancy.

KEGG pathway	Set size	$P_{pathway}$	*q*	Definition
ko00500	2,611	1.5×10^{-5}	0.0015	Starch and sucrose metabolism
B3	4,913	2.3×10^{-5}	0.0015	Energy metabolism
ko00920	805	1.4×10^{-4}	0.0048	Sulfur metabolism
ko02040	320	1.5×10^{-4}	0.0048	Flagellar assembly
ko02010	6,502	0.0011	0.029	ABC transporter
B25	659	0.0017	0.035	Cell motility
ko00460	626	0.0020	0.036	Cyanoamino acid metabolism

KEGG, Kyoto Encyclopedia of Genes and Genomes.

Supplementary Table 5. SLE case-control and within-case comparisons for phylogenetic α-diversity.

Case-control comparison

	Phylum (L2)			Class (L3)		
Sample set	Effect size	SE	$P_{\alpha\text{-diversity}}$	Effect size	SE	$P_{\alpha\text{-diversity}}$
All	-0.032	0.022	0.14	-0.075	0.032	0.019
Remove male	-0.040	0.025	0.11	-0.083	0.037	0.027
Remove patients with antibiotics	-0.021	0.024	0.39	-0.059	0.035	0.095
Remove patients with proton pump inhibitor	-0.043	0.030	0.15	-0.130	0.043	0.0028
Remove patients with treatment for SLE	-0.031	0.034	0.37	-0.140	0.048	0.0040
	Order (L4)			Family (L5)		
Sample set	Effect size	SE	$P_{\alpha\text{-diversity}}$	Effect size	SE	$P_{\alpha\text{-diversity}}$
All	-0.072	0.034	0.038	-0.181	0.038	3.6×10^{-6}
Remove male	-0.083	0.040	0.042	-0.191	0.042	1.4×10^{-5}
Remove patients with antibiotics	-0.056	0.039	0.15	-0.142	0.042	7.9×10^{-4}
Remove patients with proton pump inhibitor	-0.135	0.047	0.0043	-0.265	0.052	7.9×10^{-7}
Remove patients with treatment for SLE	-0.138	0.052	0.0090	-0.223	0.060	2.4×10^{-4}
	Genus (L6)			Species (L7)		
Sample set	Effect size	SE	$P_{\alpha\text{-diversity}}$	Effect size	SE	$P_{\alpha\text{-diversity}}$
All	-0.332	0.054	4.2×10^{-9}	-0.164	0.057	0.0042
Remove male	-0.337	0.060	8.7×10^{-8}	-0.157	0.063	0.014
Remove patients with antibiotics	-0.278	0.060	5.5×10^{-6}	-0.122	0.063	0.052
Remove patients with proton pump inhibitor	-0.396	0.076	5.1×10^{-7}	-0.270	0.078	6.1×10^{-4}
Remove patients with treatment for SLE	-0.335	0.088	1.8×10^{-4}	-0.206	0.090	0.022

comparison within the case

	Phylum (L2)			Class (L3)		
Variable	Effect size	SE	$P_{\alpha\text{-diversity}}$	Effect size	SE	$P_{\alpha\text{-diversity}}$
Not newly diagnosed	0.048	0.060	0.43	0.168	0.092	0.074
Lupus nephritis	-0.007	0.050	0.89	0.015	0.080	0.85
SLE-DAI ≥ 8	-0.048	0.058	0.42	-0.034	0.093	0.71
	Order (L4)			Family (L5)		
Variable	Effect size	SE	$P_{\alpha\text{-diversity}}$	Effect size	SE	$P_{\alpha\text{-diversity}}$
Not newly diagnosed	0.173	0.101	0.093	0.087	0.098	0.38
Lupus nephritis	0.028	0.087	0.75	0.022	0.082	0.79
SLE-DAI ≥ 8	-0.046	0.102	0.66	0.035	0.096	0.72
	Genus (L6)			Species (L7)		
Variable	Effect size	SE	$P_{\alpha\text{-diversity}}$	Effect size	SE	$P_{\alpha\text{-diversity}}$
Not newly diagnosed	0.039	0.138	0.78	0.012	0.146	0.93
Lupus nephritis	-0.027	0.115	0.82	0.064	0.122	0.60
SLE-DAI ≥ 8	0.093	0.134	0.49	0.205	0.140	0.15

SE, standard error; SLE, systemic lupus erythematosus.

Supplementary Table 6. SLE case-control and within-case comparisons for genetic α-diversity.

Case-control comparison			
Sample set	Effect size	SE	$P_{\alpha\text{-diversity}}$
All	-0.228	0.057	7.9×10^{-5}
Remove male	-0.210	0.063	0.0011
Remove patients with antibiotics	-0.181	0.062	0.0042
Remove patients with proton pump inhibitor	-0.344	0.078	1.7×10^{-5}
Remove patients with treatment for SLE	-0.295	0.092	0.0015
comparison within the case			
Variable	Effect size	SE	$P_{\alpha\text{-diversity}}$
Not newly diagnosed	0.038	0.138	0.78
Lupus nephritis	0.064	0.115	0.58
SLE-DAI ≥ 8	0.153	0.133	0.26

SE, standard error; SLE, systemic lupus erythematosus.

Supplementary Table 7. SLE case-control and within-case comparisons for phylogenetic β-diversity.

Group A	Group B	*P* (PERMANOVA)					
		Phylum (L2)	Class (L3)	Order (L4)	Family (L5)	Genus (L6)	Species (L7)
HC	SLE	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}
HC (Only female)	SLE (Only female)	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}
HC	SLE (Remove patients with antibiotics)	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}
HC	SLE (Remove patients with proton pump inhibitor)	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}	1.0×10^{-5}
HC	SLE (Remove patients with treatment for SLE)	0.0013	0.0013	8.4×10^{-4}	2.0×10^{-5}	4.0×10^{-5}	3.0×10^{-5}
SLE (Newly diagnosed)	SLE (Not newly diagnosed)	0.32	0.33	0.33	0.11	0.084	0.12
SLE (without LN)	SLE (with LN)	0.94	0.72	0.67	0.70	0.69	0.57
SLE (SLE-DAI < 8)	SLE (SLE-DAI ≥ 8)	0.12	0.10	0.10	0.027	0.032	0.11

HC, healthy control; SE, standard error; SLE, systemic lupus erythematosus; LN, lupus nephritis.

Supplementary Table 8. SLE case-control and within-case comparisons for genetic β-diversity.

Group A	Group B	*P* (PERMANOVA)
HC	SLE	1.0×10^{-5}
HC (Only female)	SLE (Only female)	1.0×10^{-5}
HC	SLE (Remove patients with antibiotics)	1.0×10^{-5}
HC	SLE (Remove patients with proton pump inhibitor)	1.0×10^{-5}
HC	SLE (Remove patients with treatment for SLE)	3.0×10^{-5}
SLE (Newly diagnosed)	SLE (Not newly diagnosed)	0.23
SLE (without LN)	SLE (with LN)	0.34
SLE (SLE-DAI < 8)	SLE (SLE-DAI ≥ 8)	0.14

HC, healthy control; SE, standard error; SLE, systemic lupus erythematosus; LN, lupus nephritis.

METHODS

Patient participation

We examined the 48 SLE patients at Osaka University Hospital and the National Hospital Organization Osaka Minami Medical Center. The SLE patients were diagnosed according to the systemic lupus international collaborating clinics classification criteria (SLICC),[1]. The 205 healthy controls were enrolled at Osaka University Graduate School of Medicine, Osaka University Hospital and National Hospital Organization Osaka Minami Medical Center. The healthy controls had no personal history of immune-related diseases. Among the healthy control samples, 27 samples which belonged to the sequencing group1 were derived from a previous study,[2].

Participants with extreme diets (e.g., strict vegetarians) and the healthy controls who were treated with antibiotics for at least a month prior to sampling were removed from analysis. Characteristics of the subjects are described in **online supplementary table 1**. All subjects provided written informed consent before participation. The study protocol was approved by the ethical committees of Osaka University and related medical institutions. In microbe-metabolome association analysis, we re-analyzed phylogenetic abundance data from the 75 healthy control subjects in a previously published study,[3].

Sample collection and DNA extraction

Fecal samples were collected in tubes containing RNAlater (Ambion). After the weights of the samples were measured, RNAlater was added to make 10-fold dilutions of homogenates. Fecal samples were stored at -80°C within 24 hours after collection. Bacterial DNA was extracted according to a previously described method,[2]. Briefly, 0.3 g glass beads (diameter 0.1 mm) (BioSpec), and 500 µl EDTA-Tris-saturated phenol were added to the suspension, and the mixture was vortexed vigorously using a FastPrep-24 (MP Biomedicals) at 5.0 power level for 30 seconds. After centrifugation at 20,000 g for 5 minutes at 4°C, 400 µl of

supernatant was collected. Subsequently, phenol-chloroform extraction was performed, and 250 µl of supernatant was subjected to isopropanol precipitation. Finally, DNAs were suspended in 100 µl EDTA-Tris buffer and stored at -20°C.

Whole-genome shotgun sequencing

A shotgun sequencing library was constructed using the KAPA Hyper Prep Kit (KAPA Biosystems) and 150 bp paired-end reads were generated on HiSeq 3000 for the sequencing group1 (average 25.8 million paired end reads per sample) and group2 (average 24.4 million paired end reads per sample), on Novaseq 6000 for the sequencing group3 (average 29.6 million paired end reads per sample). Healthy control data from the sequencing group1 was also used in the previous study,[2]. The sequence reads were converted to FASTQ format using bcl2fastq (version 2.19).

Quality control of sequencing reads

We followed a series of the QC steps to maximize the quality of the datasets. The main steps in the QC process were: (i) trimming of low-quality bases, (ii) identification and masking of human reads, and (iii) removal of duplicated reads. We marked duplicate reads using PRINSEQ-lite (version 0.20.4; parameters: -derep 1). We trimmed the raw reads to clip Illumina adapters and cut off low-quality bases at both ends using the Trimmomatic (version 0.39; parameters: ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:8:true LEADING:20 TRAILING:20 SLIDINGWINDOW:3:15 MINLEN:60). We discarded reads less than 60 bp in length after trimming. Next, we performed duplicate removal by retaining only the longest read among the duplicates with same sequences. As a final QC step, we aligned the quality-filtered reads to the human reference genome (hg38) using bowtie2 (version 2.3.5) with default parameters and BMTagger (version 3.101). We kept only reads of which both paired ends failed to align in either tool.

Taxonomic annotation of metagenome and abundance quantification

To improve both the efficiency and the accuracy of taxonomic assignment, we used curated reference metagenomes as previously described,[2]. The reference metagenomes of the Japanese population constructed by Nishijima et al.,[4] were combined with the metagenomes identified from the cultivated human gut bacteria projects,[5–7]. After filtration to the genomes annotated to the species with more than 50 reference genomes, the taxonomic reference genome dataset consisted of 7,881 genomes. The filtered paired-end reads were aligned to the reference genome dataset using bowtie2 with default parameters. As for multiple-mapped reads, only the best possible alignment was selected by the alignment scores. The number of reads that mapped to each genome was divided by the length of the genome. The value of each genome was summed up by each sample, and the relative abundance of each clade was calculated at six levels (L2: phylum, L3: class, L4: order, L5: family, L6: genus, L7: species). Then, we detected outlier samples by PCA.

Functional annotation and abundance calculation

De novo assembly of the filtered paired-end reads into contigs was conducted using MEGAHIT (version 1.2.9; parameters: --min-contig-len 135). We predicted open reading frames (ORFs) on the contigs with the *ab initio* gene finder MetageneMark (version 3.38; parameters: -a -k -f G). Next, we annotated the ORF catalog with Kyoto Encyclopedia of Genes and Genomes (KEGG) protein database (https://www.kegg.jp),[8]. We utilized a database of prokaryote KEGG genes and MGENES, a database of KEGG genes from metagenome samples annotated based on orthology, with a bit score >60. We aligned putative amino acid sequences translated from the ORF catalog against KEGG protein database with DIAMOND using BLASTP (version v0.9.32.133; parameters: f 6 -b 15.0-k 1 -e 1e-6 --subject-cover 50). We identified 1,814,229 KEGG genes. For quantification of the

ORF abundance, we mapped the filtered paired-end reads to the assembled contigs using bowtie2 with default parameters. To avoid the bias of the gene size, the ORF abundance was defined as the depth of each ORF's region of the ORF catalog according to the mapping result. Then, we detected outlier samples by PCA. Two healthy control samples and an SLE samples were removed from analysis because they were outliers in PCA analysis on the phylogenetic data or the KEGG gene abundance data (**online supplementary figure 1**).

Case-control association test for phylogenetic data

We normalized the relative abundance profiles using log transformation. We removed clades detected (i) in less than 20% of the samples, (ii) in no sample in either cases or controls, (iii) in no sample in any of the three sequencing groups, or (iv) with an average relative abundance of less than 0.001% of total abundance. After selection, we assessed 774 clades (12 phyla, 25 classes, 36 orders, 72 families, 178 genera, and 451 species). Case-control association tests were performed separately for each standardized clade abundance using the lm() function in the R (version 4.0.1) and the effect size of disease state was evaluated. Abundances of the taxa were set as objective variables and distributions of the residuals from linear regression models were checked manually for the significantly changed taxa to confirm the satisfaction of the assumption of normality. We adopted sex, age, sequencing group and the top 25 principal components as covariates. To evaluate the effect of potential confounding factors, we performed sub-analysis with dataset from which (i) males were removed, (ii) those treated with antibiotics were removed, (iii) those treated with proton pump inhibitors were removed or (iv) those treated for the SLE were removed. In addition, the SLE patients were stratified according to (i) newly diagnosed or not, (ii) presence of lupus nephritis, and (iii) SLE-DAI ≥ 8 or not. Then, comparison within case was performed.

Case-control association test for gene abundance data

We converted each ORF abundance to annotated gene abundance for KEGG gene databases. We performed two steps of normalization. First, we adjusted the gene abundance by the sum of ORF abundance for each sample in order to correct the bias of the amount of sequence reads for each sample. Next, we applied a rank-based inverse normal transformation in order to correct the heterogeneity of each gene's abundance and distribution. We removed genes detected (i) in less than 20% of the samples, (ii) in no sample in either cases or controls, or (iii) in no sample in any of the three sequencing groups. After gene selection, we assessed 240,544 genes annotated by KEGG gene database. Case-control association tests were performed separately for each clade using the lm() function in the R and the effect size of the disease state was evaluated. Abundances of the genes were set as objective variables and distributions of the residuals from linear regression models were checked manually for the significantly changed taxa to confirm the satisfaction of the assumption of normality. We adopted sex, age, sequencing group and the top 25 principal components as covariates. To evaluate the effect of potential confounding factors, we performed sub-analysis with the dataset from which (i) males were removed, (ii) those treated with antibiotics were removed, (iii) those treated with proton pump inhibitors were removed, or (iv) those treated for SLE were removed. In addition, the SLE patients were stratified according to (i) newly diagnosed or not, or (ii) presence of lupus nephritis, and (iii) SLE-DAI ≥ 8 or not. Then, comparison within case was performed.

Biological pathway enrichment analysis of the gut microbiome

We performed a gene set enrichment analysis using the R package fgsea (version 1.16.0). Gene sets which contained over 30,000 genes or under 50 genes were excluded from the enrichment analysis. For case-control pathway association tests, genes annotated by the KEGG database were ranked based on their z-value in case-control gene association tests.

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

The KEGG gene sets were defined according to the KEGG pathway. False discovery ratio (FDR) was calculated by Benjamini-Hochberg procedure.

Comparison of the pathway analysis results between the MWAS and the GWAS

We assessed whether there existed shared biological pathways between the gut metagenome and the human germline genome. We compared the pathway enrichment data of the metagenome in the SLE patients with the GWAS of SLE and RA. For GWAS, we used PASCAL,[9] with summary statistics from the SLE GWAS,[10] (4,943 cases and 8,483 controls) and the RA GWAS,[11] (14,361 cases and 43,923 controls) in the European population in order to determine KEGG pathway enrichment of the human genome. Based on the biological pathway enrichment analysis from MWAS and GWAS, we classified biological pathways into the following four groups; (1) Biological pathways enriched in MWAS and GWAS results ($P_{pathway} < 0.05$ in MWAS and $P_{pathway} < 0.05$ in GWAS) (2) Biological pathways enriched only in MWAS result ($P_{pathway} < 0.05$ in MWAS and $P_{pathway} \geq 0.05$ in GWAS). (3) Biological pathways enriched only in GWAS result ($P_{pathway} \geq 0.05$ in MWAS and $P_{pathway} < 0.05$ in GWAS). (4) Biological pathways not enriched in MWAS and GWAS results ($P_{pathway} \geq 0.05$ in MWAS and $P_{pathway} \geq 0.05$ in GWAS). Then, we evaluated the enrichment of the number of biological pathways in the group (1) compared to the group (2~4) by a one-tailed Fisher's exact test (**online supplementary figure 2**).

Empirical estimation of the metagenome-wide significance threshold

We empirically estimated the statistical significance threshold separately for the phylogenetic and the gene case-control analyses, performing a phenotype permutation procedure,[12]. We randomly simulated case-control phenotypes (× 10,000 iterations) and calculated empirical null distributions of the minimum p-values (= P_{min}) in each iteration. We defined an empirical Bonferroni significance threshold at a significance level of 0.05, as the 95th

percentile of P_{min} (= P_{sig}). We calculated the P_{sig} using the Harrell-Davis distribution-free quantile estimator and calculated a 95% confidence interval for P_{sig} by a bootstrapping method in the R package Hmisc (version 4.4_0). To estimate the null distribution of the test statistics, we applied the same process used for minimum p-values to all the p-values above each rank. We defined an empirical FDR threshold of 0.05 as the 95th percentile of all the $-\log_{10}(P)$ above each rank.

Case-control difference between α-diversity and β-diversity of the metagenome

α-diversity (within-sample diversity) was calculated as a Shannon index based on the gene abundance and the six levels of phylogenetic relative abundance (L2-L7) for each sample. Case-control comparison were performed with the lm() function in the R and the effect size of disease state was evaluated. We adopted sex, age and sequencing group as covariates. To evaluate β-diversity, multidimensional scaling (MDS) on the Bray-Curtis dissimilarity was performed. For evaluating case-control differences in the dissimilarity, we performed permutational multivariate analysis of variance,[13] (PERMANOVA) with 99,999 permutations using the adonis() function in R package vegan (version 2.5_6). Sex, binary transformed age (<50 or ≥50) and sequencing group were adjusted as covariate before the variance explained by disease state was evaluated.

Participants of metabolome analysis

We used the previously obtained plasma metabolite data for microbe-metabolite association analysis,[14]. Among the participants, 31 subjects (27 healthy controls, 4 cases) from sequencing group1 and 72 subjects (67 healthy controls, 5 cases) from the sequencing group 2 were profiled their plasma metabolite by CE-TOFMS and LC-TOFMS in the previous study,[14]. In addition, 75 healthy control subjects who belong to replication dataset were also profiled their plasma metabolite in the same study.

Sample collection and metabolome profiling

Details were described in Kishikawa et al,[15]. Briefly, plasma samples were collected in BDTM P100 Blood Collection Tubes (Franklin Lakes, NJ, USA), which contains spray-dried K2 EDTA anticoagulant, proprietary proteinase inhibitors and a mechanical separator to obtain plasma. Within 1h of collection, the samples were centrifuged for 15 min at 2500 × g and stored at -80°C until analysis. Metabolite extraction and metabolome analysis were conducted at Human Metabolome Technologies (HMT), Japan. For CE-TOFMS measurement, 50 µl of plasma was added to 450 µl of methanol containing internal standards (H3304-1002, HMT) at 0°C to inactivate enzymes. The extract solution was thoroughly mixed with 500 µl of chloroform and 200 µl of Milli-Q water and centrifuged at 2300 × g and 4°C for 5 min. The 350 µl of the upper aqueous layer was centrifugally filtered through a Millipore 5-kDa cutoff filter to remove proteins. The filtrate was centrifugally concentrated and re-suspended in 50 µl of Milli-Q water for CE-MS analysis. For LC-TOFMS measurement, 500 µl of plasma was added to 1500 µl of 1% formic acid/acetonitrile containing internal standard solution (Solution ID: H3304-1002, Human Metabolome Technologies, Inc., Tsuruoka, Japan) at 0°C to inactivate enzymes. The solution was thoroughly mixed and centrifuged at 2300 × g and 4°C for 5 min. The supernatant was filtrated by using Hybrid SPE phospholipid (55261-U, Supelco, Bellefonte, PA, USA) to remove phospholipids. The filtrate was desiccated and then dissolved with 100 µl of iso- propanol/Milli-Q for LC-MS analysis.

Metabolome analysis was conducted by the Dual Scan package of HMT using CE-TOFMS and LC-TOFMS for ionic and nonionic metabolites, respectively. CE-TOFMS analysis was carried out using an Agilent CE system equipped with an Agilent 6210 TOFMS, Agilent 1100 isocratic HPLC pump, Agilent G1603A CE-MS adapter kit and Agilent G1607A CE-ESI-MS sprayer kit (Agilent Technologies, Santa Clara, CA, USA). The systems were controlled by Agilent G2201AA ChemStation software version B.03.01 for CE (Agilent

Technologies) and connected by a fused silica capillary (50 μm i.d. × 80 cm total length) with commercial electrophoresis buffer (H3301-1001 and I3302-1023 for cation and anion analyses, respectively, HMT) as the electrolyte. The spectrometer was scanned from m/z 50 to 1000. LC-TOFMS analysis was carried out using an Agilent LC System (Agilent 1200 series RRLC system SL) equipped with an Agilent 6230 TOFMS (Agilent Technologies). The systems were controlled by Agilent G2201AA ChemStation software version B.03.01 (Agilent Technologies) equipped with ODS column (2 × 50 mm, 2 μm). Peaks were extracted using MasterHands, automatic integration software (Keio University, Tsuruoka, Yamagata, Japan) to obtain peak information including m/z, peak area and migration time for CE-TOFMS measurement (MT) or retention time for LC-TOFMS measurement (RT). Signal peaks corresponding to isotopomers, adduct ions and other product ions of known metabolites were excluded, and remaining peaks were annotated according to the HMT metabolite database based on their m/z values with the MTs and RTs determined by TOFMS. Areas of the annotated peaks were normalized based on internal standard levels and sample amounts to obtain relative levels of each metabolite.

Microbe-metabolome association analysis

541 and 512 metabolites were detected in the discovery and the replication dataset. 398 and 385 metabolites which were detected in ≥20% of the samples were retained for analysis. The abundance of metabolite was normalized using log transformation. Microbe-metabolite association tests were performed separately for each microbe-metabolite pair using the lm() function in the R and the effect size of the microbe abundance was evaluated. In the discovery dataset, we adopted sex, age, sequencing group and the top principal component of microbe and the top five principal components of metabolite as covariates. In the replication dataset, we adopted sex, age and the top principal component of microbe and the top five principal

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

components of metabolite as covariates. FDR was calculated by Benjamini-Hochberg procedure.

Power calculation

Power calculation was performed for phylogenetic and functional analysis using the wp.regression() function in R package WebPower (version 0.6) considering of pre-estimated effect sizes ranged from 0.01 to 0.35 (represented as Cohen's f^2). To account for the case-control balance in our dataset, we calculated effective sample size as $N_{eff} = 4 / (1 / N_{case} + 1 / N_{control}) = 152.66 \simeq 150$, according to the previous report,[16]. According to Cohen's (1988) guidelines,[17], $f^2 \geq 0.02$, $f^2 \geq 0.15$, $f^2 \geq 0.35$ represent small, medium, and large effect sizes, respectively.

Patient and public involvement

This research was done without patient and public involvement. Patients and public were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results.

Tomofuji Y, *et al. Ann Rheum Dis* 2021; 80:1575–1583. doi: 10.1136/annrheumdis-2021-220687

URLs

bcl2fastq, https://support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software/downloads.html

Trimmomatic, http://www.usadellab.org/cms/?page=trimmomatic

bowtie2, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml

BMTagger, ftp://ftp.ncbi.nlm.nih.gov/pub/agarwala/bmtagger/

PRINSEQ, http://prinseq.sourceforge.net/

MEGAHIT, https://github.com/voutcn/megahit

Metagenemark, http://exon.gatech.edu/Genemark/meta_gmhmmp.cgi

DIAMOND, https://github.com/bbuchfink/diamond

Samtools, http://www.htslib.org/download/

bedtools, https://github.com/arq5x/bedtools2

R, https://www.r-project.org

Python, https://www.python.org/downloads/release/python-376/

Seqkit, https://bioinf.shenwei.me/seqkit/download/

Hmisc, https://cran.r-project.org/web/packages/Hmisc/index.html

vegan, https://cran.r-project.org/web/packages/vegan/index.html

fgsea, https://github.com/ctlab/fgsea

WebPower, https://CRAN.R-project.org/package=WebPower

Pascal, https://www2.unil.ch/cbg/index.php?title=Pascal

KEGG, https://www.kegg.jp

REFERENCES

1 Petri M, Orbai A-M, Alarcón GS, *et al.* Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis & Rheumatism* 2012;64:2677–86. doi:10.1002/art.34473

2 Kishikawa T, Maeda Y, Nii T, *et al.* Metagenome-wide association study of gut microbiome revealed novel aetiology of rheumatoid arthritis in the Japanese population. *Ann Rheum Dis* 2020;79:103–11. doi:10.1136/annrheumdis-2019-215743

3 Kishikawa T, Ogawa K, Motooka D, *et al.* A Metagenome-Wide Association Study of Gut Microbiome in Patients With Multiple Sclerosis Revealed Novel Disease Pathology. *Front Cell Infect Microbiol* 2020;10:585973. doi:10.3389/fcimb.2020.585973

4 Nishijima S, Suda W, Oshima K, *et al.* The gut microbiome of healthy Japanese and its microbial and functional uniqueness. *DNA Res* 2016;23:125–33. doi:10.1093/dnares/dsw002

5 Almeida A, Mitchell AL, Boland M, *et al.* A new genomic blueprint of the human gut microbiota. *Nature* 2019;568:499–504. doi:10.1038/s41586-019-0965-1

6 Forster SC, Kumar N, Anonye BO, *et al.* A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol* 2019;37:186–92. doi:10.1038/s41587-018-0009-7

7 Zou Y, Xue W, Luo G, *et al.* 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat Biotechnol* 2019;37:179–85. doi:10.1038/s41587-018-0008-8

8 Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. ;:4.

9 Lamparter D, Marbach D, Rueedi R, *et al.* Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. *PLoS Comput Biol* 2016;12:e1004714. doi:10.1371/journal.pcbi.1004714

10 Julià A, López-Longo FJ, Pérez Venegas JJ, *et al.* Genome-wide association study meta-analysis identifies five new loci for systemic lupus erythematosus. *Arthritis Res Ther* 2018;20:100. doi:10.1186/s13075-018-1604-1

11 the RACI consortium, the GARNET consortium, Okada Y, *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;506:376–81. doi:10.1038/nature12873

12 Kanai M, Tanaka T, Okada Y. Empirical estimation of genome-wide significance thresholds based on the 1000 Genomes Project data set. *J Hum Genet* 2016;61:861–6. doi:10.1038/jhg.2016.72

13 Anderson MJ. A new method for non-parametric multivariate analysis of variance: NON-PARAMETRIC MANOVA FOR ECOLOGY. *Austral Ecology* 2001;26:32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x

14 Kishikawa T, Maeda Y, Nii T, *et al.* Increased levels of plasma nucleotides in patients with rheumatoid arthritis. *International Immunology* 2021;33:119–24. doi:10.1093/intimm/dxaa059

15 Kishikawa T, Arase N, Tsuji S, *et al.* Large-scale plasma-metabolome analysis identifies potential biomarkers of psoriasis and its clinical subtypes. *Journal of Dermatological Science*

16 Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26:2190–1. doi:10.1093/bioinformatics/btq340

17 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, N.J: : L. Erlbaum Associates 1988.