Supplemental materials

Materials and Methods

Studies were performed under the supervision of the NYU Institutional Human Subjects Committee, and all subjects provided written informed consent, with samples obtained during visits for routine care. Exclusion and inclusion criteria for patients and healthy controls have previously been described [1]. Ethnicity and race were self-declared based on standard questionnaire. All subjects were females, and age-range for patients and controls were not significantly different. These specific criteria, other aspects of the clinical studies, and the collection and characterization of the gut microbiota are described in Tables S1, S2 and S3. Control subjects were sampled quarterly.

Microbiota analyses

As previously described [1], fecal sampling was performed following a standardized and validated collection protocol using a special media for later recovery of viable bacteria. Microbial DNA from fecal samples was isolated following a validated standard protocol, with extraction directly or after frozen and stored at -80°C (as per Human Microbiome Project website).

For phylogenetic assignments, we analyzed diagnostic 16S ribosomal RNA (16S rRNA) gene sequences in libraries for each fecal sample[2-4]. Briefly, to determine the distribution of operational taxonomic units (OTUs)[5, 6] the diagnostic V4 region of 16S rRNA gene is amplified with flanking oligo primers with embedded 16 bp barcodes, producing a 254 bp read length. From each sample, three replicate libraries were generated with the same bar-coded oligonucleotide primer pair, which were then pooled, then purified[2], and stored until sequencing. With the MiSeq instrument (Illumina) in a 96-well format, these amplimers were generated and later characterized based on ~150 bp reads in both directions.

Upstream informatics analysis. The quality-filtered pre-processed sequences of the community sequence data were analyzed using QIIME pipeline [7]. The pipeline consists of the following steps: (i) clustering of the sequences into operational taxonomical units (OTUs) using UCLUST program at 97% similarity level [8]; (ii) taxonomical assignment of each OTU by running RDP Classifier [9] at 80% bootstrap confidence on a representative sequence from each OTU; (iii) alignment of representative sequences using PyNAST [7] with the Greengenes core-set alignment template; (iv) building a phylogenetic tree for the OTUs using FastTree program[10]; and (v) calculating Unifrac distances between each sample[11]. The data was then exported into R phyloseq[12] data structures and analyzed using custom reproducible RMarkdown scripts. Alpha diversity analyses, association and correlation analyses, as well as most visualizations were performed in R.

Corroboration of results obtained with the QIIME pipeline was done by inference of amplicon sequence variants (ASV) as with the DADA2 pipeline[13] with Silva reference database[14]. To determine the relative abundances with respect to renal disease, univariate testing of ASVs was performed using the PESAME (Predictive Effect Size Analysis in Multivariate Ensembles) protocol[15] and our 16S rRNA amplicon data and the Greengenes database, which assessed the significance of pairwise differences using the Mann-Whitney test. This study was not powered adequately for multiple comparison

correction; however, we examined the significance of our results by false discovery rate[16]. Predictive effect sizes were estimated by converting the U statistics to area under receiver operator characteristic curve. Confidence intervals on the area under receiver operator characteristic curve metrics were estimated by normal approximation [17].

Isolation, characterization, genome sequence determination of RG colonies

To identify and recover *RG* colonies, fecal samples were streaked onto TSBA or BHI plates and grown under anaerobic conditions. Individual colonies, identified based on morphology and growth characteristics of the strains ATCC 21492 (termed RG1) and CC55_001C (termed RG2), were sub-streaked, then sub-cultured. From each colony genomic DNA was recovered with the power soil (Qiagen™) kit, according to the manufacturer's instructions, quantified on a Nanodrop 1000 (Thermo Scientific).

To assess total and RG specific 16S rRNA gene representation [18], PCR assays were performed with the T100 thermocycler (Bio-Rad™) using the primers:

UniF340(5'-ACTCCTACGGGAGGCAGCAGT-3')

UniR514(5'-ATTACCGCGGCTGCTGGC-3').

For the following cycles: an initial 94°C for 3 min, followed by 35 cycles of 94°C for 45 seconds, 50°C for 1 min; and followed by extension at 72°C for 10 min and 4°C hold.

The RG species-specific 16S rRNA was determined with the previously reported oligonucleotide primers [18]:

Fwd 5'-GGACTGCATTTGGAACTGTCAG-3'

Rev 5'-AACGTCAGTCATCGTCCAGAAAG-3'

for the following cycles: an initial 94°C for 3 min, followed by 35 cycles of 94°C for 45 seconds, 58°C for 1 min; and followed by extension at 72°C for 10 min and 4°C hold.

Colonies of interest were named based on the Lupus patient donor that provided the faecal sample of origin, S47- and S107-, followed by the number of the bacterial isolate. DNA from each isolate was then subjected to whole genome sequencing, using both NextSeq 550 Illumina and PacBio technologies.

Bacterial whole-genome sequence analysis

Raw short-read sequence reads were preprocessed using fastp [19] version 0.22 using default settings. Preprocessed reads were screened for within-species contamination using ConFindr [20] as well as cross-species contamination using MetaPhlAn [21]. Preprocessed short reads were assembled using Unicycler [22] using its conservative mode.

Processing and assembly of PacBio long reads was performed using the PacBio SMRT Link software suite. Assemblies were further assessed by BLAST search [23] against the NCBI nt database. Contigs matching *Cutibacterium acnes*, a common skin bacterium, were removed from long-read assemblies, and it was verified that the remaining contigs mapped to corresponding uncontaminated short-read assemblies genome assemblies

were annotated using Prokka [24]. The BEDTools software suite [25] was used to divide genome assemblies into 1kbp windows and to analyze GC content per window. Each window was compared to RG1 (NCBI accession GCF_009831375.1) and RG2 (GCF_000507805.1) using BLAST [23], keeping only alignments with E values below 10^{-20}. Visualizations of each genome assembly were produced using the circlize R package [26].

For classical multi-dimensional analyses of whole genome sequences (WGS) of a dissimilarity matrix paired end sequencing reads were *de novo* assembled into contigs with SPAdes [27]. We then computed whole genome-wise nucleotide similarity with FastANI [28] where collected draft genomes from health control (n=35), inflammatory bowel disease (n=11), SLE strains (n=26) and reference genomes of *R. gnavus* (n=4) were utilized. The distance of each pair of draft genome was defined as 1-(ANIi,j+ANIj,i)/2 where i is the first draft genome, and j is the second draft genome. Iterating over all pairs of draft genomes, a distance matrix of 76*76 was then constructed and subject to classical multidimensional scaling decomposition. The samples were then projected on the first two eigenvectors with health conditions as different colors and reference patterns as different shapes.

Assessment of gene content across assemblies was performed using Orthofinder [29]. The resulting presence/absence matrix of orthogroups was used to generate pairwise Jaccard dissimilarities between isolates using the vegan R package [30]. Snippy was used to produce a core genome alignment [31], using the downloaded RefSeq genome assembly for strain RG1 as reference (NCBI accession number GCF_009831375.1). A phylogenetic tree was inferred from the core alignment using RAxML [32] with a GTRGAMMA model and 100 bootstrap replicates.

Lipoglycan purification and mass spectrometric analysis. RG strains were individually expanded in an anerobic biofermeter to 300 ml of a chopped meat media culture, grown to stationary phase and then pelleted, and used in an established extraction procedure designed to isolate lipoteichoic acids from Gram-positive bacteria [33], which were performed as previously described [1] from which earlier analysis of the RG2 purified moiety was shown to contain a diacylglycerol- (DAG-) containing lipid anchor characteristic of a lipoglycan (LG). Briefly, to purify the cell wall lipoconjugates, bacterial cells were disrupted with a French press, then the precipitate was removed by ultracentrifugation. The supernatant was subjected to butanol-water extraction and the water-soluble component then passaged over a Hydrophobic Interaction Chromatography (HIC) to isolate lipoglycan-containing fractions.

The mass spectrometric analyses of LG preparations were performed on a Q Exactive Plus (ThermoFisher Scientific, Bremen, Germany) using a Triversa Nanomate (Advion, Ithaca, NY) as nano-ESI source. LG extracts were initially dissolved in a concentration of 1 µg µl−1 in water and 10 µl of this solution were mixed with 150 µl of water/propan-2-ol/7 M triethylamine/acetic acid (50:50:0.06:0.02, [v/v/v/v]). Mass spectra were recorded for 0.50 min in the negative mode in an *m/z*-range of 400–2000 or 500-3000 applying a spray voltage of −1.1 kV. Depicted MS^1 spectra were charge deconvoluted (Xtract module of Xcalibur 3.1 software; ThermoFisher Scientific, Bremen, Germany) and all provided values refer to mono-isotopic masses of neutral molecules. Single scan *.mzmL files were generated with MSconvert [34] and used as import for LipidXplorer 1.2.8 [35] to compute

an aligned data set. From this data set the top hundred intense peaks were used to prepare the heatmap shown in Fig. 5D and to compute the similarity score depicted in Fig. 5E based on Spearman rank correlation. For MS^2 experiments aiming to analyze the glycolipid linker composition, the de-*O*-acyl LG3 preparation described earlier [1] was used. Doubly charged ions of interest were selected and spectra were recorded in the negative ion mode at different normalized collision energies (NCE). With the derived data, a structural model (Fig. 5F) was generated with Biorender software (Biorender.com).

Immunoblotting. Electrophoretic separation used Bis-Tris mini gels (Novex, Thermo Fisher) with bacterial extracts loaded at the same concentration, then transferred to membranes, which were incubated with sera diluted at 1:100, and incubated overnight at 4°C. For detection, anti-human IgG biotin conjugated (Jackson ImmunoResearch Labs, USA) was added and developed by IRDye® 800CW Streptavidin (LI-COR®).

Generation of LG-specific murine monoclonal antibodies

A commercial vendor (Envigo Bioproducts Inc., Indianapolis) immunized 10 BALB/c mice with extract of the RG2 strain emulsified into complete Freund's adjuvant and later boosted with lipoglycan purified from the Lupus S47-18 strain, which was emulsified in incomplete Freund's adjuvant. All LG were purified from an RG strain by a method that included fractionation by hydrophobic interaction chromatography, as previously described [1]. The spleen from the mouse with the strongest post-immunization was fused with Ig-deficient NS-1 myeloma cells. The spent supernatants subclones were evaluated for IgG-reactivity, which demonstrated highly correlated reactivity with whole extracts of the immunizing RG strain and purified RG lipoglycan, with the subcloned hybridoma cell lines, termed mAb 33.2.2 and mAb 34.2.2. Antibody gene sequences were determined by Abterra bio, San Diego) (see Supplementary Figure S12).

Direct binding ELISA

To detect the reactivity of the murine monoclonal antibodies 33.2.2 and mAb 34.2.2 with the different RG strains, the ELISA plates were coated with the bacterial extracts from RG2, S47-18, S107-48, S107-86, RG1 as well as with the purified lipoglycan from the strains RG2 and S47-18. Next, the murine monoclonal antibodies were added at 100ng/ml and 25ng/ml in duplicate, after incubation for 2hrs at RT. Binding was detected with goat anti mouse IgG HRP conjugated at 1:10,000 (Jackson ImmunoResearch), then TMB substrate was added to develop the plate.

To detect human serum IgG antibody responses to RG lipoglycan, microtiter wells were coated with purified lipoglycan at 0.5 ug/ml in BBS overnight at 4ºC. After blocking with 1% BSA in PBS, for 1hr at RT, serum samples were added to at 1:3,200 for total IgG, and at 1:800 for IgG subclass-specific detection and incubated for 2hrs at RT. Binding was detected with goat anti-human IgG F'ab'2 biotin-conjugated (Cat. 109-066-006, Jackson ImmunoResearch) at 1:20,000, murine anti-human IgG1 biotin-conjugated (Cat. A10650, Invitrogen) at 1:1000, murine anti-human IgG2 biotin-conjugated (Cat. 05-3540, Invitrogen) at 1:1000, murine anti-human IgG3 biotin-conjugated (Cat. 05-3640, Invitrogen) at 1:1000, murine anti-human IgG4 biotin-conjugated (Cat. A10663, Invitrogen) at 1:500, incubated for 1hr at RT. Then, washed with 0.05% Tween20-PBS, then HRP-conjugated Streptavidin (Thermo-Fisher) at 1:20,000 was added, and

incubated for 1hr at RT, washed then followed by addition of TMB substrate to develop wells.

Multiplex Bead based immunoassay

The assays performed as previously described in [1]. Briefly, serum samples from patients and healthy controls underwent to 4-fold serial dilutions starting at 1:200 to 1:12,800 against a panel of antigens, including extracts of the Lupus strains RG2, S47-18 and RG1, and the purified lipoglycan from the strains RG2 and S47-18, then detecting using goat anti-human IgG PE conjugate (eBioscience).

Generating a recombinant monoclonal chimeric antibody to the RG lipoglycan

To generate a recombinant monoclonal antibody for use as a standard in a human IgG anti-LG assay, we designed a chimeric antibody with the variable regions, for the 33.2.2 hybridoma cell line (see Figure S14), fused to human IgG2-kappa constant regions. For antibody gene synthesis, DNA sequences were generated in which the VH region of the parental 33.2.2 B-cell hybridoma cell line were placed upstream (i.e., 5') to the gene encoding the human g2 subclass constant region. In parallel, the light chain variable region for this cell line was placed upstream of the human kappa constant region gene. These target genes were amplified by PCR, with oligonucleotide primers that facilitated cloning into a compatible mammalian expression vector, which was transfected into HEK293 cells (Sino Biological).

For transfection, the plasmids were mixed with transfection reagents at an optimal ratio and then added into the culture of HEK293 cells, which were grown in a serum-free medium and maintained in Erlenmeyer flasks in a bioreactor with stirring at 37°C for 6 days. Cells were removed by centrifugation, and the cell culture supernatant was loaded onto a protein-A affinity purification column, then IgG was eluted with a mild acid buffer with rapid neutralization and dialysis into a physiologic pH buffer.

The purified protein was analyzed by, SDS-PAGE and size exclusion chromatography (SEC). We characterized the capacity of the chimeric antibody to bind the purified lipoglycan antigen in a validated immunoassay, with a method using the Magpix instrument (Luminex) as previously described [1]. Taken together, a functional recombinant antibody was produced and purified, which was demonstrated to retain high level binding activity for the lipoglycan produced by a strain of R. gnavus.

Statistical analysis

Data are presented as mean ± SD. The student unpaired *t* test with Welch correction was used in 2-group comparisons of normally distributed data, whereas the Mann-Whitney nonparametric test was used when the normality assumption was not met. Fisher's exact test was performed to evaluate bivariate associations between categorical variables, or as described. To test for correlations between two variables Spearman test was used. *p*-values were considered significant at <0.05 for two-tailed tests. Prism software Version 9 (GraphPad) was used for all analyses.

Supplemental figures

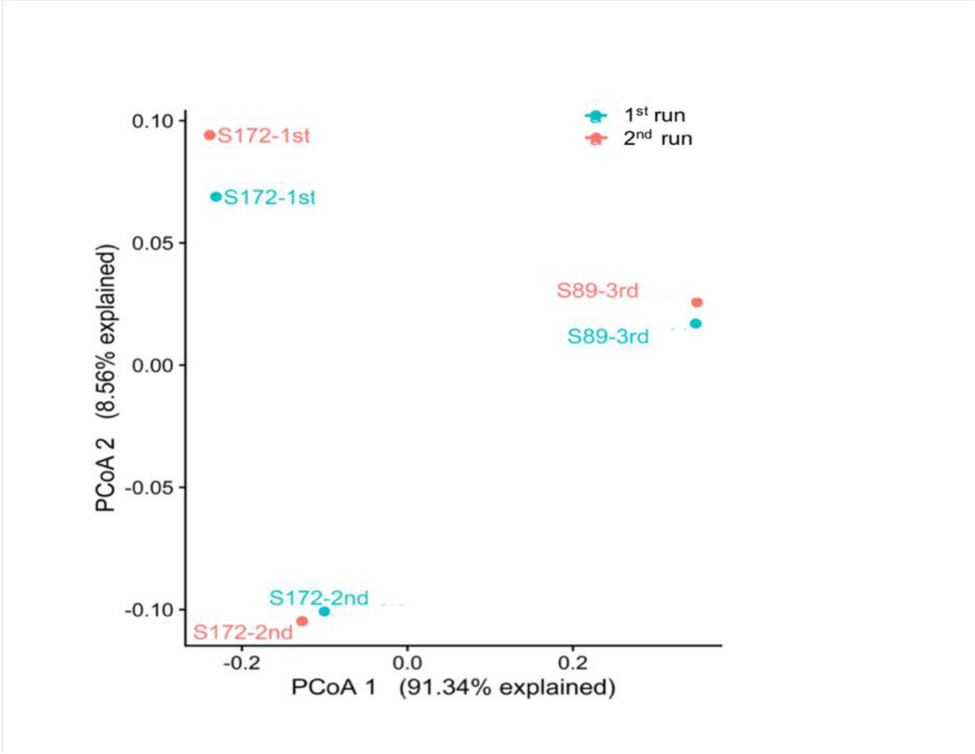


Figure S1. Consistent libraries composition between the two 16S rRNA gene amplimer Miseq batches. Principal Coordinates Analysis (PCoA) on libraries generated from samples of same individuals/samples sequenced in the two different 16S rRNA sequencing runs (1st and 2nd) as an inter-run technical control, shows no consistent bias or batch effect.

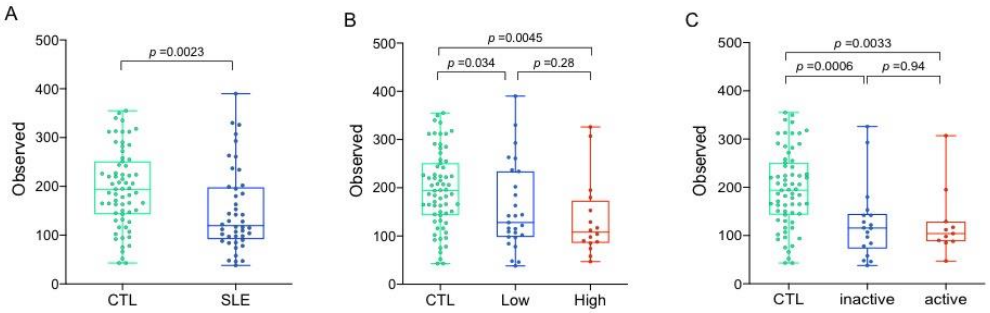


Figure S2. Dysbiosis in **in SLE microbiota communities**. (**A**) The number of distinct taxa were estimated based on observed ASV, and alpha diversity richness was reduced in samples from Lupus patients (SLE) compared to healthy controls (CTL) (Wilcoxon, *p*=0.0023). (**B**) Compared to CTL, alpha diversity was reduced in Lupus patients with low disease activity (based on SLEDAI score), with even greater contractions in the group with high disease activity, (Wilcoxon, *p*=0.034, *p*=0.0045, respectively). (**C**) Compared to CTL, alpha diversity was reduced in those with inactive renal disease, and further contracted in those with active renal disease (Wilcoxon, *p*=0.0006 and *p*=0.0033, respectively). The purpose was to assess correlations in this data set, and hence these analyses did not consider statistical adjustments for multiple samples from the same patient.

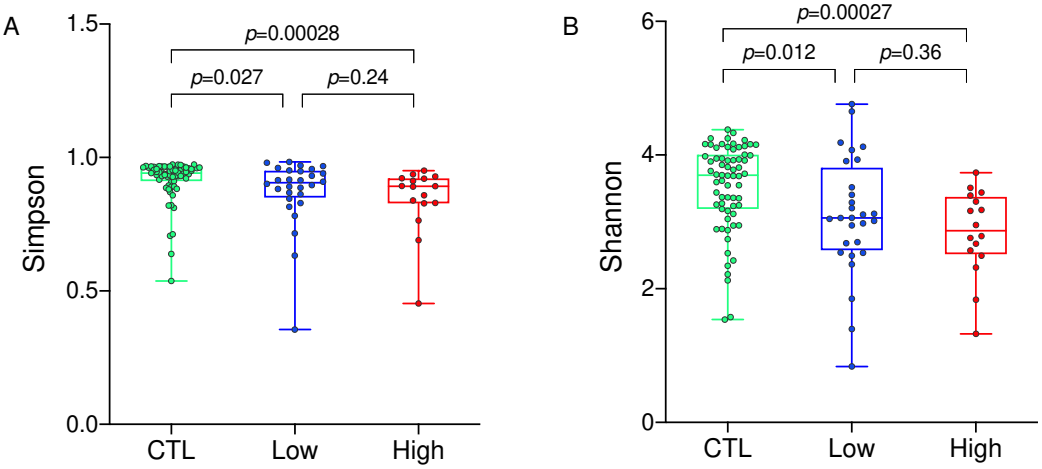


Figure S3. Alpha diversity is reduced in libraries from patients with high Lupus disease activity. Analyses are as shown in Fig. 1. High disease activity is defined as a composite SLEDAI score of ≥8.

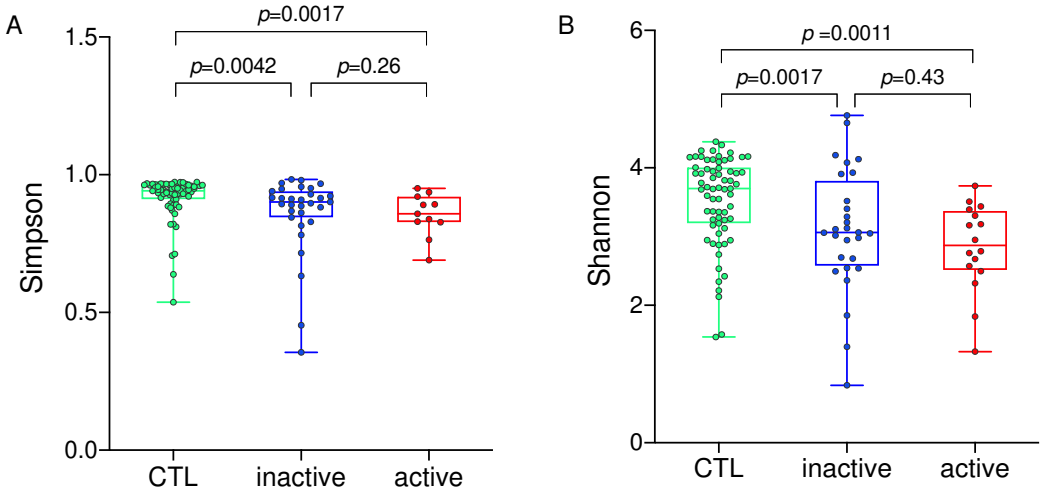


Figure S4. Alpha diversity is reduced in libraries from patients with active renal disease compared to inactive renal disease. Analyses performed as shown in Fig. 2. Active renal disease was defined by standard clinical laboratory criteria.

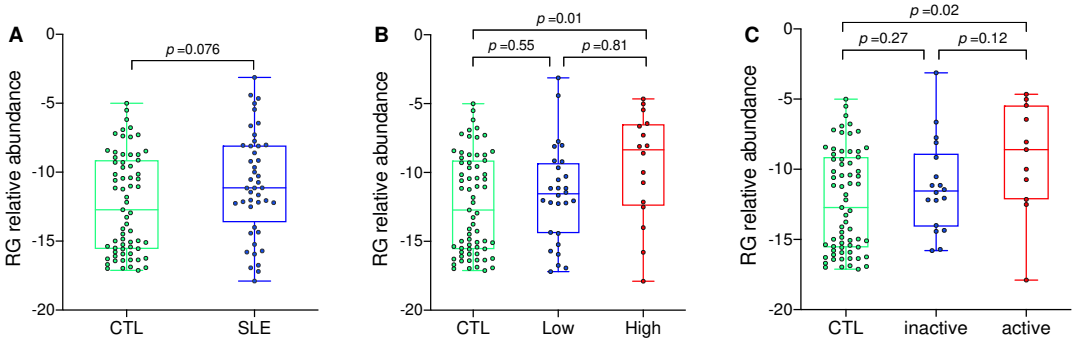


Figure S5. RG expansions occur at the time of high Lupus disease activity and active LN. (**A**) All samples from SLE patients have a numerical but non-significant trend toward increased RG abundance compared to healthy CTL (Wilcoxon, *p*=0.0760). (**B**) Samples from patients with high disease activity (based on SLEDAI) showed a greater RG abundance, compared to low disease activity and CTL (Wilcoxon, *p* =0.81, *p* =0.01, respectively). (**C**) RG expansions were common in the active LN group compared to healthy CTL (Wilcoxon *p*=0.02), but not significantly different in the inactive LN group. Wilcoxon, *p*=0.27, NS. RG relative abundance is shown in log 2 values.

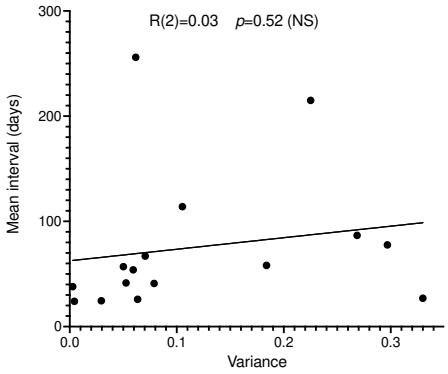


Figure S6. Variance within the microbiota communities of an SLE affected individual did not shift in a time-dependent manner. As an independent means to evaluate whether there is evidence of a time-dependence of community variance in these SLE patients sampled over time (Figure 2), we have performed a Spearman correlation analysis of time interval vs. variance. These values did not show significant correlation.

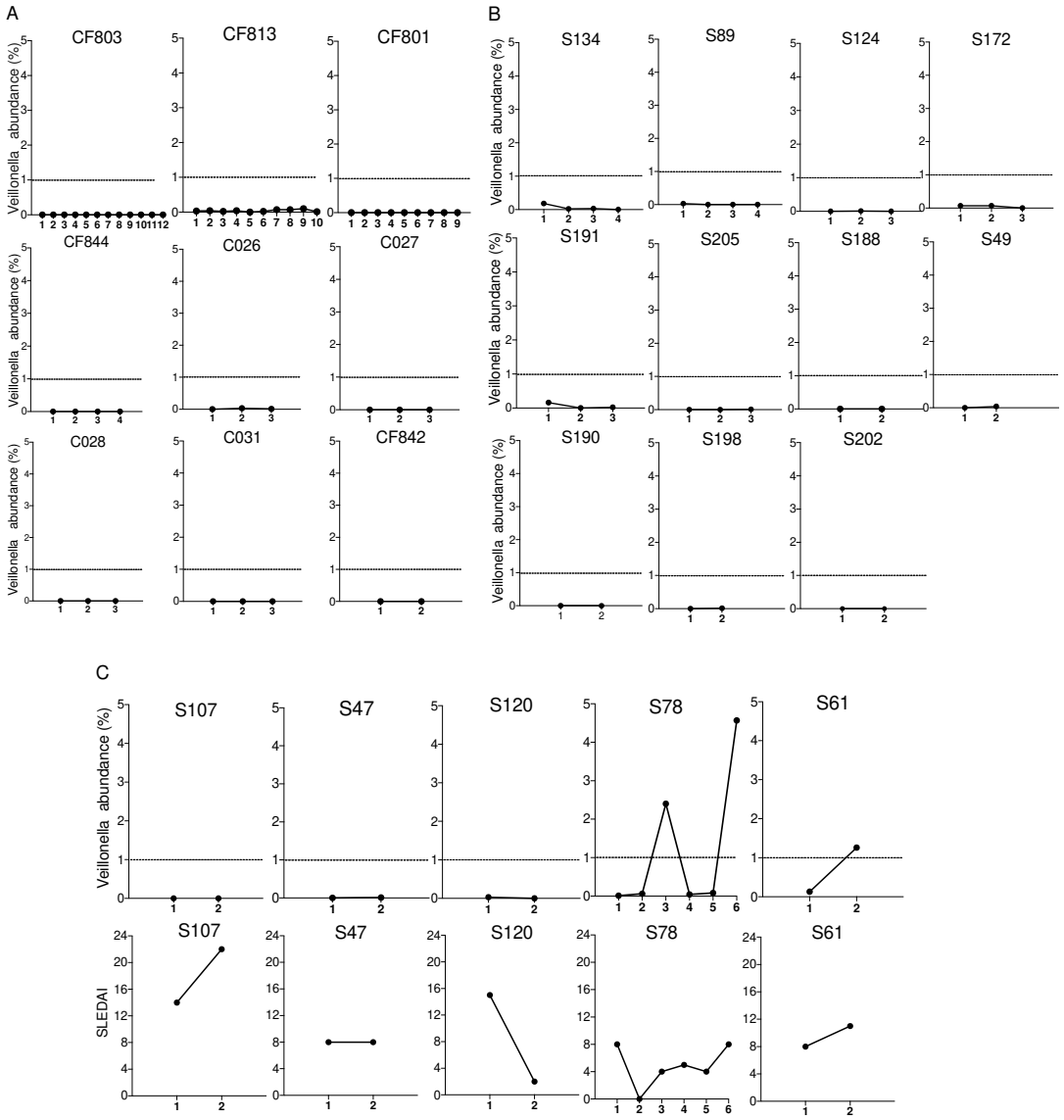


Figure S7. **Blooms of *Veillonella* do not occur concurrent with Lupus disease activity**. (**A**) *Veillonella* abundance in healthy individuals. (**B**) *Veillonella* abundance in SLE patients. (**C**) *Veillonella* abundance in SLE patients with above-described RG blooms concordant with Lupus disease activity flares. Abundance based on ASV representation in total amplicon libraries.

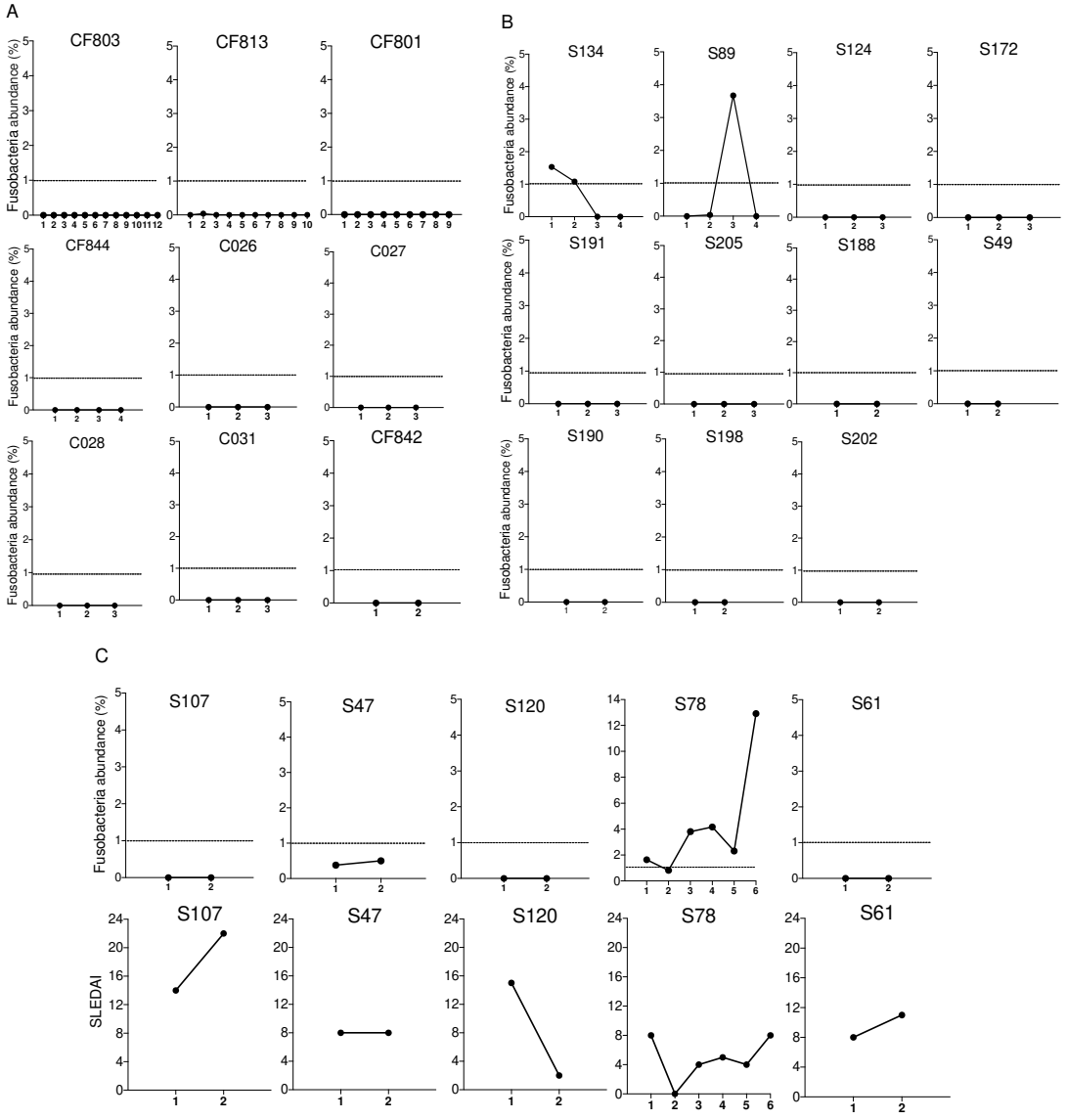


Figure S8. Blooms of *Fusobacterium* do not occur concurrent with peak flares of Lupus disease activity. (**A**) *Fusobacterium* abundance in healthy individuals. (**B**) *Fusobacterium* abundance in SLE patients. (**C**) *Fusobacterium* abundance in SLE patients with above-described RG blooms concordant with Lupus disease activity flares. Abundance based on ASV representation in total amplicon libraries.

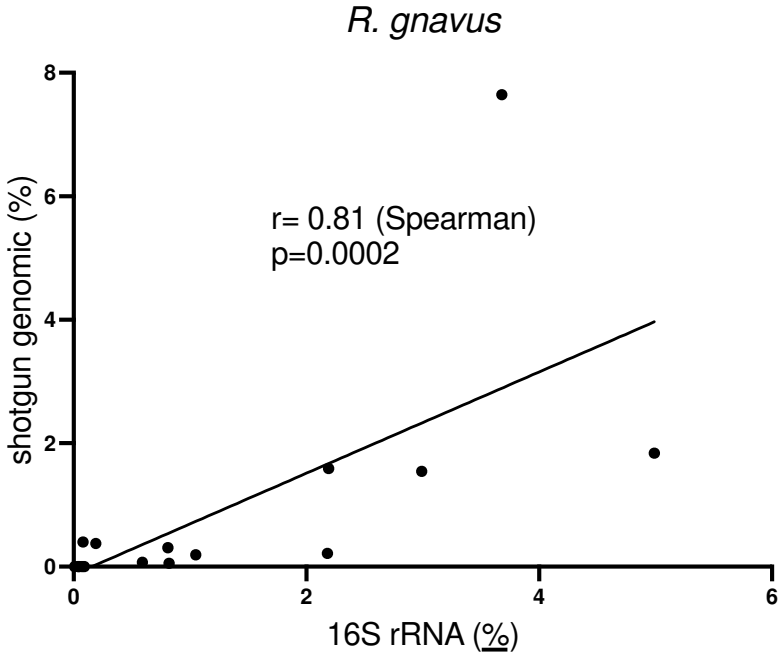


Figure S9. Abundance of R. gnavus in SLE fecal genomic samples as determined by 16srRNA and shotgun metagenomic analysis is highly correlated. Here, the abundance of R. gnavus species was determined by these two approaches for 16 SLE fecal samples. Values for Spearman analysis with p value for two-tailed analysis are shown.

Methods for shotgun metagenomic analysis. Fecal genomic DNA was used from 16 SLE patients was used to generate individual metagenomic sequencing runs. After quality filtering using fastp v. 0.23.2 [36], the number of reads in each run ranged from 62 million to 79 million. MetaPhlAn v. 4.0.6 [37], which were used to quantify the relative abundance of taxa from each sequencing run. The relative abundance of Ruminococcus gnavus in these 16 patient samples, as determined by metagenomic shotgun sequencing, were then compared with results as determined by 16S rRNA library analysis. The calculated correlation coefficients of the relative abundances obtained with each of the two methods for R. gnavus, by Pearson correlation was 0.66 and by Spearman correlation was 0.81.

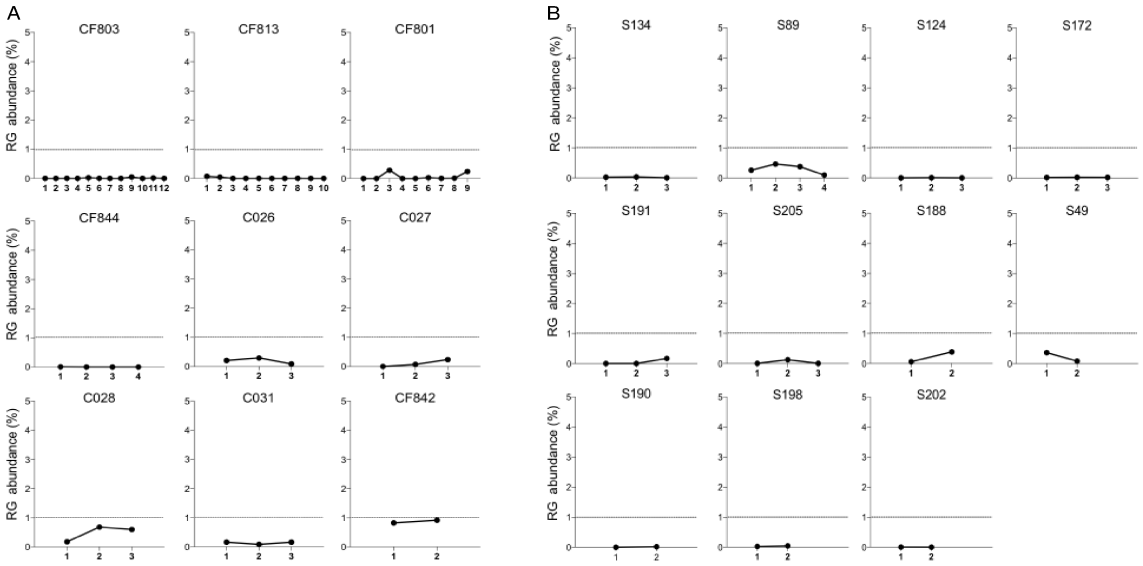


Figure S10. Lack of **dynamic changes in RG abundance documented at sequential time points in healthy and many Lupus-affected individuals.** (**A**) All healthy control subjects displayed a stable low abundance in RG representation. (**B**) For 11 of the 16 SLE patients under investigation, a stable low abundance in RG representation was detected. By comparison, in Figure 2, data for 5/16 (31%) of the SLE patients evaluated overtime, which documented that the abundance of RG fluctuated greatly overtime. In those cases, RG abundance at much higher levels were present in fecal samples obtained proximal to visits in which disease flares were documented. Clinical and demographic data are shown in table S1&S3. Dotted line depicts an arbitrary 1% threshold of 16S rRNA amplicon representing RG abundance that is highly above the mean 0.15% level in these healthy controls.

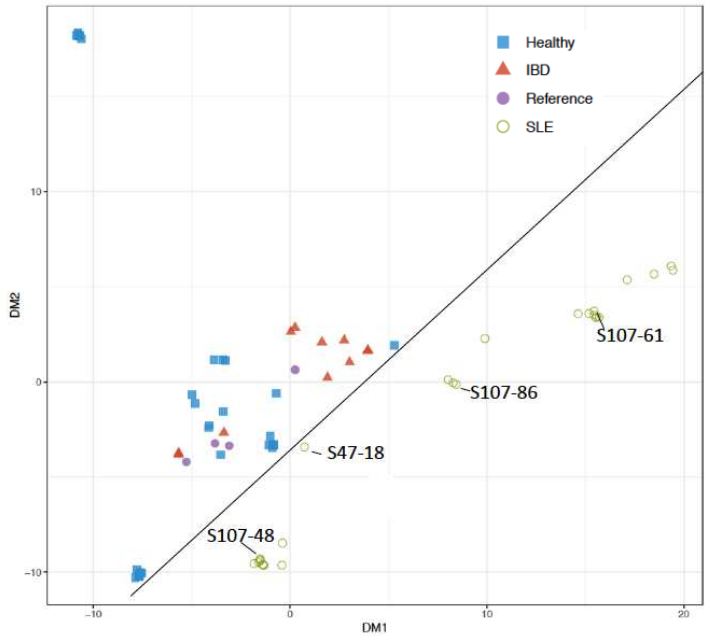


Figure S11. Classical multi-dimensional analyses of whole genome sequences (WGS) of a dissimilarity matrix reveals inter-person RG variability and SLE-unique clades**.** Within these comparisons are strains from healthy adults [38], IBD-related strains [39] and 27 strains from SLE patients with LN flares. The LN strains are shown to distribute into four subgroupings for which a representative strain is identified. The LN strains show differences from all other strains. Reference refers to genome sequences for RG1 (ATCC29149) and RG2 (CC_001C).

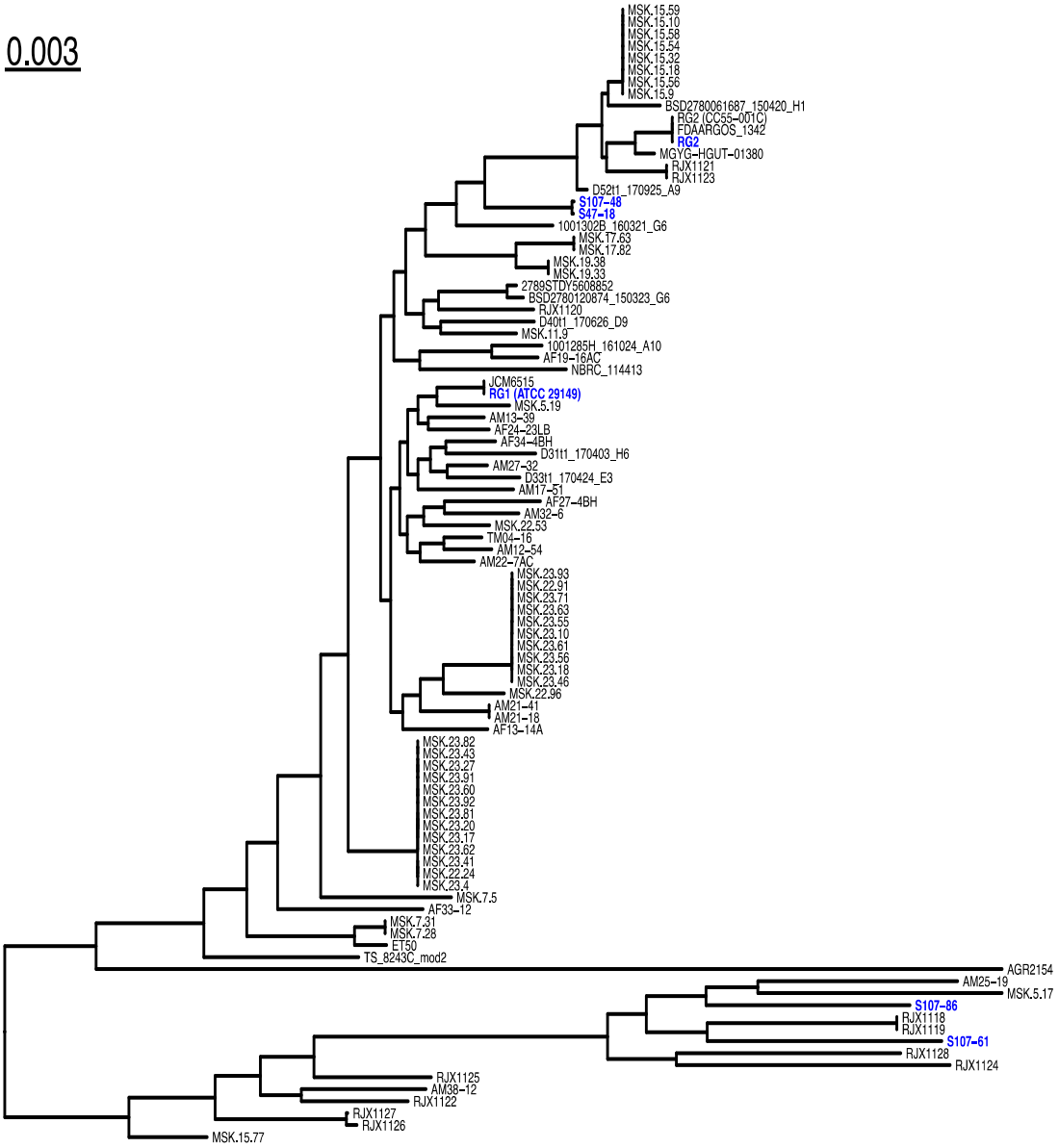


Figure S12. Phylogenetic tree based on core alignment of *Blautia (Ruminococcus) gnavus* (RG) genome assemblies downloaded from NCBI RefSeq, together with five newly generated genome assemblies. Newly sequenced isolates are shown in blue, along with downloaded RG1 (ATCC 29149) strain for reference.

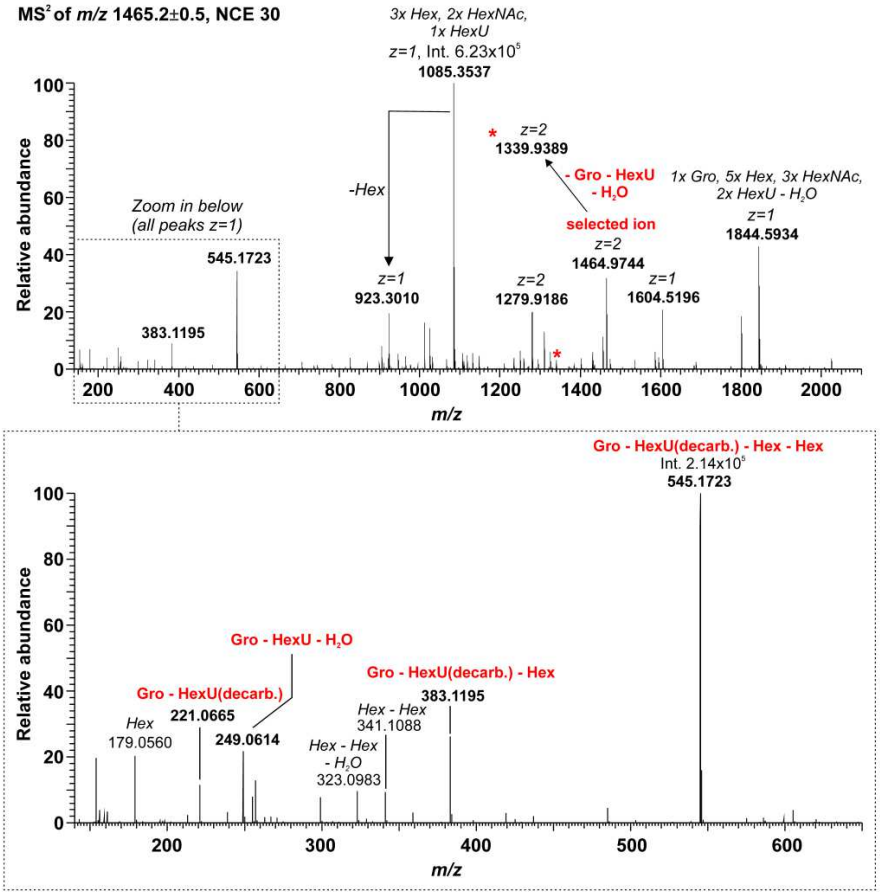


Figure S13. MS^2 analysis of the basic de-*O*-acyl LG without hexose extension (composed of 1 Gro, 8 Hex, 5 HexNAc, 3 HexU; calc. mono-isotopic mass: 2931.963 Da) of strain RG2 obtained after hydrazine-treatment [1]. In the experiment shown here, we isolated the double charged peak for this molecule and applied an NCE of 30 to induce fragmentation. Full spectrum is depicted in the upper panel. A zoom into the region of *m/z* 140 to 650 comprising the small molecular fragments generated under these conditions is depicted below. Fragments indicative for the presence of a glycerol–hexuronic acid unit (221.0665 Da: glycerol–hexuronic acid (decarboxylated); 249.0614 Da: glycerol–hexuronic acid, with loss of water) which is extended by at least two hexoses (545.1723 Da: glycerol–hexuronic acid (decarboxylated)–hexose–hexose) can be observed. The loss of the glycerol–hexuronic acid unit including loss of water can also be observed from the selected ion. (z=1, single charged ions $[M-H]^-$; z=2, double charged ions $[M-2H]^{2-}$).

Murine monoclonal antibodies recognize structurally related LGs from Lupus RG strains

To independently investigate the antigenic diversity expressed by different RG strains, we generated murine monoclonal antibodies (mAbs) by RG bacterial immunization and boosting with a purified RG LG (see methods). By ELISA, both the mAb 33.2.2 and mAb 34.2.2 strongly react with the purified LGs from the S47-18 strain used for the immunization boost as well as the purified LG from the RG2 strain (Fig. S12A). Both mAbs were reactive with the oligo bands of the same apparent MW in extracts from the RG2 strain and Lupus strains, S47-18, S107-48 and S107-86, that were isolated from two different LN patients (Fig. S12). In immunoblotting studies, these two mAbs recognized the same non-protein oligo band antigen of the same MW recognized in extracts of the Lupus RG strains from two patients, S107-48, S107-86 and S47-18, as well as the index RG2 strain, as well as purified LG from both the S47-18 strain and RG2. Notably, the mAb 33.2.2 is of the murine IgG2a subclass while mAb 34.2.2 is of the IgG1 subclass, and antibody gene analysis documented the same antibody germline gene usage in both hybridomas in independent rearrangements with only limited somatic hypermutations (fig. S13), which strongly suggests these are products of B-cell clones that are convergent in encoding for binding with the same RG LG antigen-specificity. In an attempt to further characterize their antigenic targets, these same mAbs were tested for reactivity with a collection of 313 purified bacterial glycans, predominantly capsular polysaccharides (see table S6), Yet, none displayed binding reactivity by the LG-specific mAbs, which supports the notion that in the RG LGs express a novel type of cross-reactive antigen.

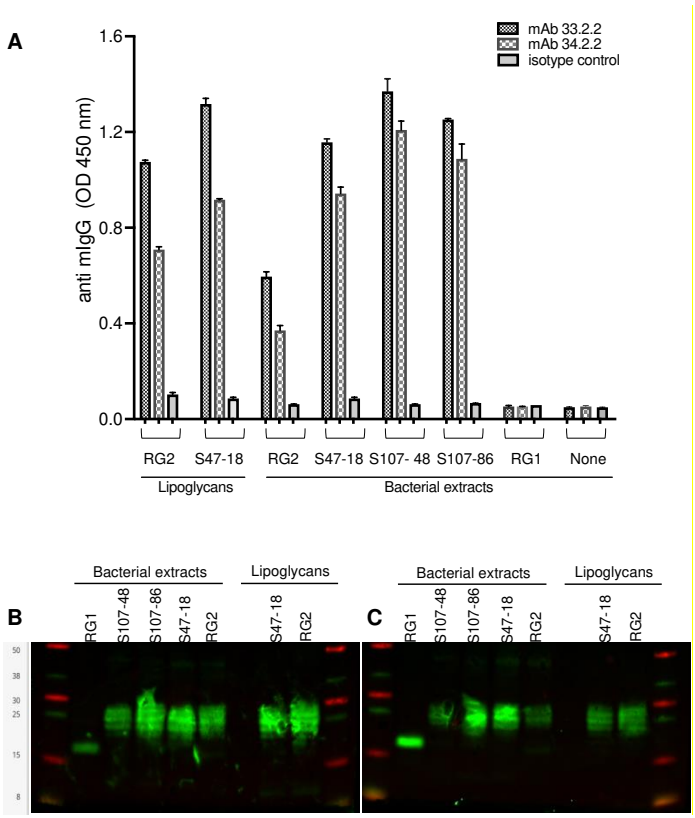
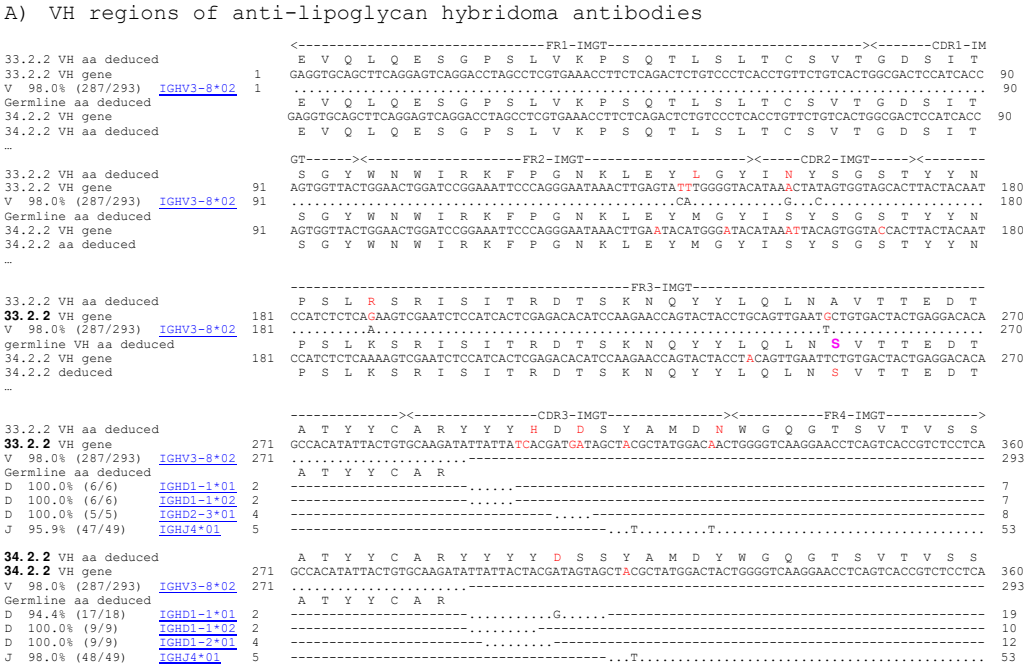


Figure S14. Reactivity of post-immunization murine monoclonal antibodies is restricted to conserved cross-reactive determinants on the oligobands of protease-resistant lipoglycans from RG strains derived from clinically active LN patients. (**A**) Direct binding ELISA demonstrates reactivity of both mAbs, 33.2.3 and 34.2.2, with the purified RG2 and S47-18 LGs, and bacterial extracts from the RG2 strain and Lupus-derived strains; S47-18, S107-48 and S107-86. In this assay, LG or nuclease-treated RG strain extracts were precoated directly onto microtiter wells, then after blocking mAbs or isotype control were incubated, then washed and developed (see methods). (**B**) Immunoblots with 33.2.3 mAb or (**C**) with the 34.2.2 mAb detect antigenically related LG oligo bands in extracts of *RG* strains isolated from active LN patients. In each panel are shown samples of purified lipoglycan (LG) from the Lupus S47-18 strain, and the RG2 strain. At left, extracts of whole bacteria are shown for the RG1 strain (from a healthy donor), and the Lupus-derived RG strains; S107-48, S107-86, S47-18, as well as the RG2 strain.

A) VH regions of anti-lipoglycan hybridoma antibodies



B) VL regions of anti-lipoglycan hybridoma antibodies

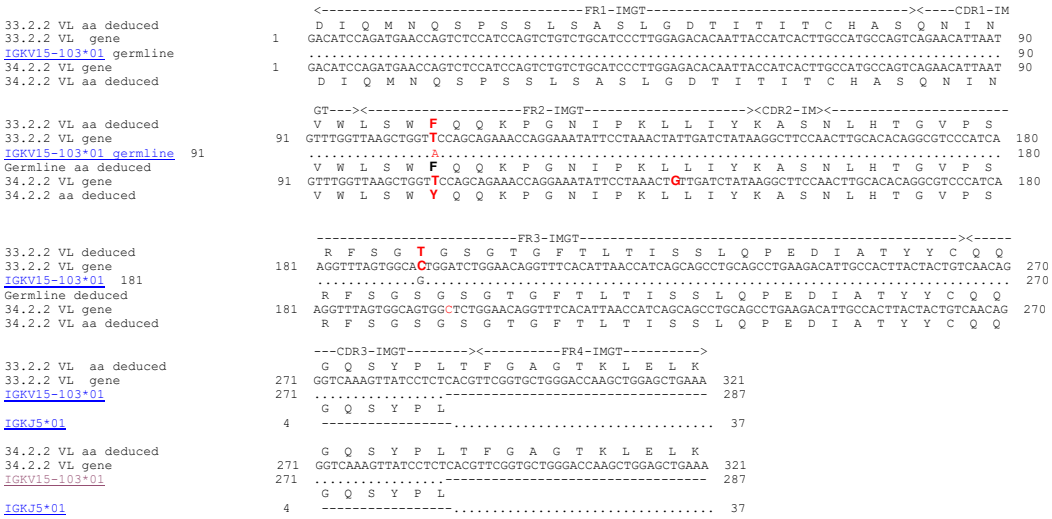


Figure S15. Antibody gene sequences and deduced amino acid sequences of the murine anti-RG LG antibodies. Alignments were based on Immunogenetics (IMGT) information systems web-based analysis.

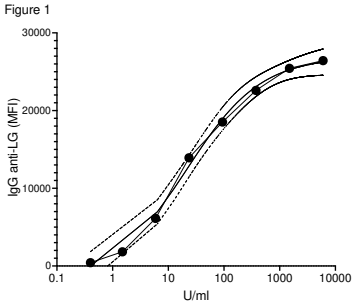


Figure S16. Binding curve for serial concentration of the 33.2.2 chimeric antibody binding (see Fig S14 for variable region genes) to custom beads coated with purified R. gnavus lipoglycan. Production and purification of the chimeric human anti-lipoglycan is described in the Methods section. On the Y axis is the mean fluorescence intensity detected for antibody binding by the Magpix instrument (Luminex). On the X axis is the concentration of the anti-lipoglycan antibody, wherein 1 ng represents 1 antibody binding activity unit (U). Each point is the mean of duplicate measurements for this sample a 2000-fold dilution. Binding activity was detected with a signal above background for the assay with a IgG concentration of below 400 pg/ml, with activity also documented that was below saturation of the assay using 1500 ng/ml of the chimeric antibody. Curve-fit was performed with Prism 9 for macOS software (Graphpad, San Diego). Lipoglycan, LG.

Table S1A. Demographic, clinical and treatment features of Lupus patients with renal involvement evaluated overtime.

Patient ID	Age	Ethnicity	Sample	Collection week span	SLEDAI	Renal ACR	Renal active	MEDs		R. *gnavus* %
								Prednisone	MMF	
S47	39	Asian	1st	0	8	1	1	0	1	1.1
			2nd	256	8		1	0	1	2.3
S78	38	White Hispanic	1st	0	8	1	0	10	1	0.4
			2nd	176	6		0	7.5	0	0.0
			3rd	205	4		0	5	0	0.2
			4th	246	5		0	5	0	0.0
			5th	268	4		0	5	0	0.1
			6th	291	8		1	5	0	9.5
S89	37	Asian	1st	0	8	1	1	40	0	0.2
			2nd	142	6		0	2.5	0	0.5
			3rd	163	8		1	5	0	0.4
			4th	233	10		1	0	0	0.1
S107	32	White Hispanic	1st	0	14	1	1	20	0	1.0
			2nd	38	22		1	25	0	3.1
S120	37	African American	1st	0	15	1	1	40	0	3.9
			2nd	215	2		0	0	0	0.0
S124	35	White Hispanic	1st	0	16	1	1	20	0	0.0
			2nd	114	8		0	0	0	0.0
			3rd	228	10		0	5	1	0.0
S134	33	White	1st	0	4	1	0	0	0	0.0
			2nd	201	6		0	0	0	0.0
			3rd	236	6		0	0	0	0.0
			4th	260	6		0	0	0	0.0
S172	24	Asian	1st	0	12	1	1	0	0	0.0
			2nd	34	9		0	0	1	0.0
			3rd	49	3		0	0	0	0.0
S202	38	White	1st	0	16	1	1	60	1	0.0
			2nd	57	2		0	0	1	0.0

Control subjects (HC) were resampled on a quarterly basis (see fig.S8).

Table S1B. Demographic, clinical and treatment features of Lupus patients without renal involvement evaluated overtime.

Patient ID	Age	Ethnicity	Sample	Collection week span	SLEDAI	Renal ACR	Renal active	MEDs		R. *gnavus* %
								Prednisone	MMF	
S49	52	African American	1st	0	2	0	0	0	0	0.4
			2nd	41	3		0	0	0	0.1
S61	42	Asian	1st	0	8	0	0	60	0	0.6
			2nd	27	11		0	3.75	0	4.7
S188	57	Asian	1st	0	9	0	un	0	0	0.1
			2nd	24	5		un	2.5	0	0.4
S190	34	White Hispanic	1st	0	4	0	0	0	0	0.0
			2nd	67	4		un	0	0	0.0
S191	43	White Hispanic	1st	0	4	0	0	0	0	0.0
			2nd	54	0		0	0	0	0.0
			3rd	83	4		0	0	1	0.2
S198	35	White Hispanic	1st	0	6	0	0	0	1	0.0
			2nd	54	0		0	0	1	0.0
S205	46	Asian	1st	0	4	0	0	0	0	0.0
			2nd	31	6		0	0	0	0.1
			3rd	52	4		0	0	0	0.0

Table S2A. SLEDAI domain scoring in patients with Lupus nephritis.

Patient ID	Sample	SLEDAI	Involvements in the SLEDAI Score
S47	1st	8	proteinuria, ↓complement, ↑dsDNA
	2nd	8	proteinuria, ↓complement, ↑dsDNA
S78	1st	8	rash, alopecia, ↓complement,↑dsDNA
	2nd	6	rash, pericarditis, ↑dsDNA
	3rd	4	rash, ↑dsDNA
	4th	5	rash, ↓WBC, ↑dsDNA
	5th	4	rash, ↑dsDNA
	6th	8	proteinuria, ↓complement, ↑dsDNA
S89	1st	8	proteinuria, ↓complement, ↑dsDNA
	2nd	6	pleurisy, ↓complement, ↑dsDNA
	3rd	8	proteinuria, ↓complement, ↑dsDNA
	4th	10	proteinuria, pleurisy, ↓complement, ↑dsDNA
S107	1st	15	proteinuria, rash, alopecia, ulcers, ↓complement, ↓WBC, ↑dsDNA
	2nd	23	arthritis, proteinuria, pyuria, rash, alopecia, pleurisy, ↓complement, ↓WBC, ↑dsDNA
S120	1st	15	myositis, proteinuria, pleurisy,↓ complement,↑dsDNA, fever
	2nd	2	↑dsDNA
S124	1st	16	visual disturbances, arthritis, ↓complement,↑dsDNA
	2nd	8	arthritis, ↓complement, ↑dsDNA
	3rd	10	arthritis, pericarditis, ↓complement, ↑dsDNA
S134	1st	4	↓complement, ↑dsDNA
	2nd	6	rash, ↓complement, ↑dsDNA
	3rd	6	rash, ↓complement, ↑dsDNA
	4th	6	rash, alopecia, ↓complement
S172	1st	12	hematuria, proteinuria, ↓complement, ↑dsDNA
	2nd	9	proteinuria, ↓complement, ↓WBC, ↑dsDNA
	3rd	7	proteinuria, ↓WBC, ↑dsDNA
S202	1st	16	hematuria, proteinuria, pyuria, ↓complement, ↑dsDNA
	2nd	2	↑dsDNA

Table S2B. Organ involvements reflected by SLEDAI in non-renal Lupus patients.

Patient ID	Sample	SLEDAI	Involvements in the SLEDAI Score
S49	1st	2	↓complement
	2nd	3	↓complement , ↓WBC
S61	1st	8	rash, alopecia, ↓complement, ↑dsDNA
	2nd	11	arthritis, alopecia, ↓complement, ↓WBC, ↑dsDNA
S188	1st	9	arthritis, ↓complement, ↓WBC, ↑dsDNA
	2nd	5	↓complement, ↓WBC, ↑dsDNA
S190	1st	4	↓complement, ↑dsDNA
	2nd	4	↓complement, ↑dsDNA
S191	1st	4	arthritis
	2nd	0	–
	3rd	4	arthritis
S198	1st	6	arthritis, ↑dsDNA
	2nd	0	–
S205	1st	4	↓complement, ↑dsDNA
	2nd	6	↓complement, ↑dsDNA, alopecia
	3rd	4	↓complement, ↑dsDNA

Table S3. RG abundance in healthy control subjects

No of collection	CF803	CF813	CF801	CF842	C026	C027	C028	C031	CF844
1	0.00119	0.0717	0.000836	0.8264	0.2046	0.00177	0.1787	0.1599	0.00568
2	0.00094	0.044	0.00235	0.9167	0.2914	0.0735	0.682	0.0874	0.00115
3	0.00437	0.00162	0.2827		0.0882	0.2357	0.6005	0.1598	0.00216
4	0.00159	0.000769	0.00119						0.000701
5	0.0215	0.00339	0.00209						
6	0.00307	0.00126	0.0278						
7	0.00256	0.00113	0.00723						
8	0.00205	0.00292	0.0127						
9	0.0473	0.000804	0.2387						
10	0.00143	0.000782							
11	0.0147								
12	0.00227								

Table S4A. Other medications of patients with Lupus Nephritis

Patient ID	Sample	Other medications
S47	1st	hydroxychlorquine 400 mg, azathioprine 150 mg
	2nd	hydroxychlorquine 400 mg
S78	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 400 mg, milatuzumab/placebo trial
	3rd	hydroxychlorquine 400 mg, methotrexate 20 mg
	4th	hydroxychlorquine 400 mg, methotrexate 20 mg
	5th	hydroxychlorquine 400 mg, methotrexate 20 mg
	6th	hydroxychlorquine 400 mg, methotrexate 20 mg
S89	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 400 mg, belimumab 10 mg, azathioprine 100 mg
	3rd	hydroxychlorquine 400 mg, belimumab 10 mg, azathioprine 100 mg
	4th	hydroxychlorquine 400 mg, methylprednisolone 1000 mg
S107	1st	hydroxychlorquine 400 mg, anifrolumab/placebo study
	2nd	hydroxychlorquine 400 mg
S120	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 200 mg, azathioprine 100 mg
S124	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 400 mg
	3rd	hydroxychlorquine 400 mg
S134	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 300 mg, belimumab 10 mg
	3rd	hydroxychlorquine 300 mg, belimumab10 mg
	4th	hydroxychlorquine 300 mg
S172	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 400 mg
	3rd	hydroxychlorquine 300 mg
S202	1st	hydroxychlorquine 400 mg, methotrexate 12.5 mg
	2nd	hydroxychlorquine 300 mg

Table S4B. Medications of patients without Lupus Nephritis.

Patient ID	Sample	Other medications
S49	1st	hydroxychlorquine 400 mg, azathioprine 100mg
	2nd	hydroxychlorquine 400 mg, azathioprine 150mg
S61	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 400 mg
S188	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 400 mg
S190	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 300 mg
S191	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 200 mg
	3rd	hydroxychlorquine 200 mg
S198	1st	hydroxychlorquine 400 mg
	2nd	hydroxychlorquine 400 mg
S205	1st	hydroxychlorquine 300 mg
	2nd	hydroxychlorquine 400 mg
	3rd	hydroxychlorquine 300 mg

Table S5. Orthologues of IBD-associated RG genes found in strains from LN patients.

PanPhlAn identifier	Old annotation	KEGG Ortholog ID	Ortholog Name	S107-61	S107-48	S107-86	S47-18	RG1/ATCC_29149	RG2/CC_001C
xg000037	PTS sugar transporter [[Ruminococcus] gnavus]								
xg000038	fibronectin [[Ruminococcus] gnavus]	K15923	AXY8, FUC95A, afcA; alpha-L-fucosidase 2 [EC:3.2.1.51]						
xg000086									
xg000091	ABC transporter								
xg000096	X-X-X-Leu-X-X-Gly heptad repeats protein	K01421*	putative membrane protein						
xg000183	MFS transporter								
xg000314	GNAT family N-acetyltransferase								
xg000605									
xg000616	plasmid mobilization relaxosome protein MobC								
xg000619									
xg000625	Signal transduction histidine kinase								
xg000685	prepilin peptidase	K02654*	leader peptidase (prepilin peptidase) / N-methyltransferase						
xg000698	peptide ABC transporter permease	K02034	ABC.PE.P1; peptide/nickel transport system permease protein						
xg000699									
xg000713	homoserine dehydrogenase	K00003	hom; homoserine dehydrogenase [EC:1.1.1.3]						
xg000714	ABC transporter	K09817*	zinc transport system ATP-binding protein [EC:7.2.2.20]						
xg000796									
xg000797	diguanylate cyclase								
xg000866									
xg001053	transcriptional regulator								
xg001176									
xg001180									
xg001199	histidine kinase								
xg001226									
xg001227	stage II sporulation protein D	K06381	spoIID; stage II sporulation protein D						
xg001228	peptidase M23								
xg001327	glycosyl transferase family 1	K19002*	1,2-diacylglycerol 3-alpha-glucosyltransferase [EC:2.4.1.337]						
xg001329	glycosyl transferase family 1	K19002*	1,2-diacylglycerol 3-alpha-glucosyltransferase [EC:2.4.1.337]						
xg001333	acylphosphatase	K001512*	acylphosphatase [EC:3.6.1.7]						
xg001345	Methionine--tRNA ligase	K06878	K06878; tRNA-binding protein						
xg001349									
xg001425	peroxiredoxin	K03564*	thioredoxin-dependent peroxiredoxin						
xg003868									
xg009760	serine hydroxymethyltransferase	K00600	glyA, SHMT; glycine hydroxymethyltransferase [EC:2.1.2.1]						

Table S6. Calculated and experimentally determined monoisotopic masses of LG species observed in MS[1] spectra of the LG mixture isolated from RG2 strain (shown in Figure 5A). Only species with a monoisotopic mass peak abundance >2% are listed. Annotation accuracy of chemical composition to mass measurements are stated as Δppm.

Acylation status	Fatty acid sum composition	Glycan composition	Calculated exact mass [Da]	Observed monoisotopic mass [Da]	Error [Δppm]
mono-acyl	16:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3170.193	3170.201	2.5
mono-acyl	16:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 1 Hex	3332.246	3332.254	2.4
mono-acyl	16:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 2 Hex	3494.298	3494.306	2.3
mono-acyl	16:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 3 Hex	3656.351	3656.358	1.9
mono-acyl	17:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3184.208	3184.217	2.8
di-acyl	30:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3380.391	3380.399	2.4
di-acyl	31:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3394.407	3394.415	2.4
di-acyl	31:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 1 Hex	3556.460	3556.466	1.7

di-acyl	31:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 2 Hex	3718.512	3718.519	1.9
di-acyl	31:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 3 Hex	3880.565	3880.570	1.3
di-acyl	31:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 4 Hex	4042.618	4042.621	0.7
di-acyl	32:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3408.422	3408.431	2.6
di-acyl	32:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 1 Hex	3570.475	3570.482	2.0
di-acyl	32:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 2 Hex	3732.528	3732.534	1.6
di-acyl	32:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 3 Hex	3894.581	3894.585	1.0
di-acyl	33:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3422.438	3422.442	1.2
di-acyl	33:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 2 Hex	3746.544	3746.543	-0.3
di-acyl	33:0	1 Gro, 8 Hex, 5 HexNAc, 3	3908.597	3908.591	-1.5

		HexU + 3 Hex			
tri-acyl	45:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3604.605	3604.612	1.9
tri-acyl	46:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3618.621	3618.627	1.7
tri-acyl	46:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 1 Hex	3780.674	3780.679	1.3
tri-acyl	46:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 2 Hex	3942.726	3942.731	1.3
tri-acyl	46:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 3 Hex	4104.779	4104.782	0.7
tri-acyl	47:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3632.636	3632.643	1.9
tri-acyl	47:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 1 Hex	3794.689	3794.695	1.6
tri-acyl	47:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 2 Hex	3956.742	3956.746	1.0
tri-acyl	47:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 3 Hex	4118.795	4118.796	0.2
tri-acyl	47:0	1 Gro, 8 Hex, 5	4280.848	4280.856	1.9

tri-acyl	47:0	HexNAc, 3 HexU + 4 Hex 1 Gro, 8 Hex, 5	4442.901	4442.907	1.4
tri-acyl	47:0	HexNAc, 3 HexU + 5 Hex 1 Gro, 8	4604.954	4604.960	1.3
tri-acyl	48:0	Hex, 5 HexNAc, 3 HexU 1 Gro, 8	3646.652	3646.657	1.4
tri-acyl	48:0	Hex, 5 HexNAc, 3 HexU + 1 Hex 1 Gro, 8	3808.705	3808.709	1.1
tri-acyl	48:0	Hex, 5 HexNAc, 3 HexU + 2 Hex 1 Gro, 8	3970.758	3970.760	0.5
tri-acyl	48:0	Hex, 5 HexNAc, 3 HexU + 3 Hex 1 Gro, 8	4132.811	4132.811	0.0
tri-acyl	48:0	Hex, 5 HexNAc, 3 HexU + 4 Hex 1 Gro, 8	4294.863	4294.871	1.9
tri-acyl	49:0	Hex, 5 HexNAc, 3 HexU 1 Gro, 8	3660.668	3660.669	0.3
tri-acyl	49:0	Hex, 5 HexNAc, 3 HexU + 1 Hex 1 Gro, 8	3822.721	3822.722	0.3
tri-acyl	49:0	Hex, 5 HexNAc, 3 HexU + 2 Hex	3984.773	3984.770	-0.8

tri-acyl	49:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU + 3 Hex	4146.826	4146.820	-1.4
tri-acyl	50:0	1 Gro, 8 Hex, 5 HexNAc, 3 HexU	3674.683	3674.681	-0.5

Table S7. Bacterial polysaccharides not recognized by the mAb 33.2.2 and mAb 34.2.2.

Providencia stuartii O49	PO49 Core-linked
Providencia stuartii O52	PO52 Core-linked
Pseudomonas aeruginosa O4 (Habs serotype 4)	PO4 Core-linked
Pseudomonas aeruginosa O1 (Fisher immunotype 4)	PO1 Core-linked
Pseudomonas aeruginosa O2 (Fisher immunotype 3)	PO2 Core-linked
Pseudomonas aeruginosa O13 (Sandvik serotype II)	PO13 Core-linked
Pseudomonas aeruginosa O9 (9a, 9b, 9d)	PO9 Core-linked
Pseudomonas aeruginosa O6a (Habs serotype6, fraction IIa)	PO6a Core-linked-O-unit
Pseudomonas aeruginosa O6a (Habs serotype6, fraction IIb)	PO6a unsubstituted core
Salmonella typhimurium SL 11881 (Re mut)	LPS-L9516
Salmonella typhimurium TV 119 (Ra mut)	LPS-L6016
Salmonella typhimurium SL 684 (Rc mut)	LPS-L5891
Pseudomonas aeruginosa O10	L8643
Salmonella typhimurium dodeca saccharide	4809
Salmonella enteritidis dodeca saccharide	1262
Salmonella typhimurium LPS	L2262
Serratia marcescens LPS	L6136
Escherichia coli K235 LPS	L2143
Escherichia coli O128-B12 LPS	L2755
Salmonella enterica abortus equi LPS	L5886
Salmonella typhosa LPS	L2387
Salmonella enteritidis LPS	L2012
Shigella bodyii type2	
Shigella bodyii type4	
Shigella bodyii type10	
Shigella dysennteriae type 3	
Shigella dysennteriae type 8 (batch 12)	
Shigella dysennteriae type 11	
Shigella dysennteriae type 13	
Escherichia coli O29	
Escherichia coli O40	
Escherichia coli O106	
Escherichia coli O130	
Escherichia coli O148	
Escherichia coli O150	
Escherichia coli O180	
Proteus mirabilis O3a, 3c (G1)	
Proteus mirabilis O8 (TG326)	

Proteus mirabilis O10 (HJ4320)	
Proteus mirabilis O29a, 29b (2002)	
Proteus mirabilis O50 (TG332)	
Proteus mirabilis O54a, 54b (10704)	
Proteus mirabilis O57 (TG319)	
Proteus penneri O8 (106)	
Proteus penneri O64a, 64b, 64d (39)	
Proteus penneri O66 (2)	
Proteus penneri O69 (25)	
Proteus penneri O71 (42)	
Proteus penneri O72a, 72b (4)	
Pseudomonas aeruginosa O2 (2a),2d,2f	IATS 10 , OPS
Pseudomonas aeruginosa O2 2a,2b	IATS 16 OPS
Pseudomonas aeruginosa O2 2a,2b,2e	IATS NO, OPS
Pseudomonas aeruginosa O2 2a,2d	IATS 5 OPS
Pseudomonas aeruginosa O2 Immuno 7	IATS 18, OPS
Pseudomonas aeruginosa O3 3a,3b	IATS NO, OPS
Pseudomonas aeruginosa O3 3a,3b,3c	IATS 3, OPS
Pseudomonas aeruginosa O3 3a,3d	IATS NO, OPS
Pseudomonas aeruginosa O4 4a,4c	IATS NO, OPS
Pseudomonas aeruginosa O6 6a	IATS 6, OPS
Pseudomonas aeruginosa O6 6a,6c	IATS NO, OPS
Pseudomonas aeruginosa O6 Immuno 1	IATS NO, OPS
Pseudomonas aeruginosa O7 7a,7b,7c	IATS 7,LPS
Pseudomonas aeruginosa O7 7a,7b,7d	IATS 8,LPS
Pseudomonas aeruginosa O7 7a,7d	IATS NO, LPS
Pseudomonas aeruginosa O10 10a,10b	IATS 10, OPS
Pseudomonas aeruginosa O10 10a,10c	IATS 19, OPS
Pseudomonas aeruginosa O11 11a,11b	IATS 11, OPS
Pseudomonas aeruginosa O12 12	IATS 12, OPS Habs 12
Pseudomonas aeruginosa O13 13a,13c	IATS 14, OPS
Pseudomonas aeruginosa O14 14	IATS 17,OPS Meitert X
Pseudomonas aeruginosa O15 15	IATS 15, OPS
Proteus vulgaris O1 (18984)*	LPS
Proteus vulgaris O4 (PrK 9/57)	OPS
Proteus vulgaris O12 (PrK 25/57)	OPS
Proteus vulgaris O13 (8344)	OPS
Proteus vulgaris O15 (PrK 30/57)	OPS
Proteus vulgaris O17 (PrK 33/57)	OPS
Proteus vulgaris O19a (PrK 37/57)	OPS
Proteus vulgaris O21 (PrK 39/57)*	LPS
Proteus vulgaris O22 (PrK 40/57)	OPS
Proteus vulgaris O25 (PrK 48/57)	OPS

Proteus vulgaris O34 (4669)*	LPS
Proteus vulgaris O37a,b (PrK 63/57)	OPS
Proteus vulgaris O37a,c (PrK 72/57)	OPS
Proteus vulgaris O44 (PrK 67/57)	OPS
Proteus vulgaris O45 (4680)	OPS
Proteus vulgaris O53 (TG 276-10)	OPS
Proteus vulgaris O54a,54c (TG 103)	OPS
Proteus vulgaris O55 (TG 155)	OPS
Proteus vulgaris O65 (TG 251)	OPS
Proteus mirabilis O6 (PrK 14/57)	OPS
Proteus mirabilis O11 (PrK 24/57)	OPS
Proteus mirabilis O13 (PrK 26/57)	OPS
Proteus mirabilis O14a,14b (PrK 29/57)	OPS
Proteus mirabilis O16 (4652)	OPS
Proteus mirabilis O17 (PrK 32/57)	OPS
Proteus mirabilis O23a,b,d (PrK 42/57)	OPS
Proteus mirabilis O26 (PrK 49/57)	OPS
Proteus mirabilis O27 (PrK 50/57)	OPS
Proteus mirabilis O28 (PrK 51/57)	OPS
Proteus mirabilis O29a (PrK 52/57)	OPS
Proteus mirabilis O40 (10703)	OPS
Proteus mirabilis O41 (PrK 67/57)	OPS
Proteus mirabilis O51 (19011)*	LPS
Proteus mirabilis O74 (10705, OF)	OPS
Proteus mirabilis O75 (10702, OC)	OPS
Proteus mirabilis O77 (3 B-m)	OPS
Proteus penneri O31a (26)	OPS
Proteus penneri O52 (15)	OPS
Proteus penneri O58 (12)	OPS
Proteus penneri O59 (9)	OPS
Proteus penneri O61 (21)	OPS
Proteus penneri O62 (41)	OPS
Proteus penneri O63 (22)	OPS
Proteus penneri O64a,b,c (27)	OPS
Proteus penneri O65 (34)	OPS
Proteus penneri O67 (8)	OPS
Proteus penneri O68 (63)	OPS
Proteus penneri O70 (60)	OPS
Proteus penneri O73a,b (103)	OPS
Proteus myxofaciens O60	OPS
Proteus O56 (genomospecies 4)	OPS

Providencia stuartii O4	OPS
Providencia stuartii O18	OPS
Providencia stuartii O20*	LPS
Providencia stuartii O43	OPS
Providencia stuartii O44	OPS
Providencia stuartii O47	OPS
Providencia stuartii O47, Core 9	OPS
Providencia stuartii O49, Core 1	OPS
Providencia stuartii O57	OPS
Providencia alcalifaciens O5	OPS
Providencia alcalifaciens O6*	LPS
Providencia alcalifaciens O19	OPS
Providencia alcalifaciens O19	LPS
Providencia alcalifaciens O19	LPS/NaOH
Providencia alcalifaciens O21	OPS
Providencia alcalifaciens O23	OPS
Providencia alcalifaciens O27	OPS
Providencia alcalifaciens O29	OPS
Providencia alcalifaciens O30	OPS
Providencia alcalifaciens O32	OPS
Providencia alcalifaciens O36*	LPS-NH4OH
Providencia alcalifaciens O39	OPS
Providencia rustigianii O14	OPS
Providencia rustigianii O16	OPS
Providencia rustigianii O34	OPS
Yersinia pestis, KM260(11)-Δ0187	LPS
Yersinia pestis, KM260(11)-Δ0187	Core oligo saccharide
Yersinia pestis, KM260(11)-Δrfe	LPS
Yersinia pestis, KM260(11)-Δrfe	Core oligo saccharide
Yersinia pestis, 1146-25	LPS
Yersinia pestis 1146-25	Core oligo saccharide
Yersinia pestis, 1146-37	LPS
Yersinia pestis, 1146-37	Core oligo saccharide
Yersinia pestis, 0KM218-37	LPS
Yersinia pestis, KM218-37	Core oligo saccharide
Yersinia pestis, KM218-25	LPS
Yersinia pestis, KM218-25	Core oligo saccharide
Yersinia pestis, KM260(11)-ΔpmrF	LPS
Yersinia pestis, KM260(11)-ΔpmrF	Core oligo saccharide
Yersinia pestis, KM260(11)-Δ0186	LPS
Yersinia pestis, KM260(11)-Δ0186	Core oligo saccharide

Yersinia pestis, KM260(11)-ΔwaaQ	LPS
Yersinia pestis, KM260(11)-ΔwaaQ	Core oligo saccharide
Yersinia pestis, KM260(11)-ΔwaaL	LPS
Yersinia pestis, KM260(11)-25	LPS
Yersinia pestis, KM260(11)-25	Core oligo saccharide
Yersinia pestis, KM260(11)-37	Core oligo saccharide
Yersinia pestis, KIMD1-37	Core oligo saccharide
Yersinia pestis, KIMD1-25	Core oligo saccharide
Yersinia pestis, 11M-25	LPS
Yersinia pestis, 11M-37	LPS
Proteus mirabilis O23a, 23b, 23c (CCUG 10701)	OPS
Proteus vulgaris O24 (PrK 47/57)	LPSOH
Yersinia pestis KM260(11)-6C	LPS
Yersinia pestis 260(11)-37C-186	LPS
Yersinia pestis 260(11)-37C-187	LPS
Yersinia pestis 260(11)-37C-416	LPS
Yersinia pestis 260(11)-37C-417	LPS
Yersinia pestis P-1680-25C	OS
Yersinia pestis P-1680-37C	LPS
Yersinia pestis I-2377-25C	OS
Yersinia pestis I-2377-37C	LPS
Francisella novicida OPS	OPS
Francisella tularensis OPS	OPS
Klebsiella O1 OPS	OPS
Klebsiella O2a OPS	OPS
Klebsiella O2ac OPS	OPS
Klebsiella O3 OPS	OPS
Klebsiella O4 OPS	OPS
Klebsiella O5 OPS	OPS
Klebsiella O8 OPS	OPS
Klebsiella O12 OPS	OPS
Shigella boydii type 1	LPSOH
Shigella boydii type 3	OPS
Shigella boydii type 5	OPS
Shigella boydii type 9	OPS
Shigella boydii type 11	OPS
Shigella boydii type 12	OPS
Shigella boydii type 15	OPS
Shigella boydii type 16	OPS
Shigella boydii type 17	OPS
Shigella boydii type 18	OPS
Escherichia coli O49	OPS
Escherichia coli O52	OPS

Escherichia coli O58	OPS
Escherichia coli O61	LPSOH
Escherichia coli O73	OPS
Escherichia coli O112ab	OPS
Escherichia coli O118	OPS
Escherichia coli O125	OPS
Escherichia coli O151	OPS
Escherichia coli O168	OPS
Shigella dysenteriae type 2	LPSOH
Shigella dysenteriae type 4	OPS
Shigella dysenteriae type 5	OPS
Shigella dysenteriae type 6 SR-strain	SR-strain
Shigella dysenteriae type 7	OPS
Shigella dysenteriae type 8 (Russian)	OPS
Shigella dysenteriae type 9	OPS
Escherichia coli O111:B4 LPS- solution at 1 mg/mL	L5293-2ML (LPS) (Sigma)
Escherichia coli O26:B6 LPS- solution at 1 mg/mL	L5543-2ML (LPS) (Sigma)
Escherichia coli O55:B5 LPS- solution at 1 mg/mL	L5418-2ML (LPS) (Sigma)
Escherichia coli O127:B8 LPS- solution at 1 mg/mL	L5668-2ML (LPS) (Sigma)
Streptococcus pneumoniae type 1 (Danish type 1)	161-X // Capsular PS
Streptococcus pneumoniae type 2 (Danish type 2)	165-X// Capsular PS
Streptococcus pneumoniae type 3 (Danish type 3)	169-X// Capsular PS
Streptococcus pneumoniae type 4 (Danish type 4)	173-X// Capsular PS
Streptococcus pneumoniae type 5 (Danish type 5)	177-X// Capsular PS
Streptococcus pneumoniae type 8 (Danish type 8)	185-X// Capsular PS
Streptococcus pneumoniae type 9 (Danish type 9N)	189-X// Capsular PS
Streptococcus pneumoniae type 12 (Danish type 12F)	193-X// Capsular PS
Streptococcus pneumoniae type 14 (Danish type 14)	197-X// Capsular PS
Streptococcus pneumoniae type 17 (Danish type 17F)	201-X// Capsular PS
Streptococcus pneumoniae type 19 (Danish type 19F)	205-X// Capsular PS
Streptococcus pneumoniae type 20 (Danish type 20)	209-X// Capsular PS
Streptococcus pneumoniae type 22 (Danish type 22F)	213-X// Capsular PS
Streptococcus pneumoniae type 23 (Danish type 23F)	217-X// Capsular PS
Streptococcus pneumoniae type 26 (Danish type 6B)	225-X// Capsular PS
Streptococcus pneumoniae type 34 (Danish type 10A)	229-X// Capsular PS
Streptococcus pneumoniae type 43 (Danish type 11A)	233-X// Capsular PS
Streptococcus pneumoniae type 51 (Danish type 7F)	237-X// Capsular PS

Streptococcus pneumoniae type 54 (Danish type 15B)	241-X// Capsular PS
Streptococcus pneumoniae type 56 (Danish type 18C)	245-X// Capsular PS
Streptococcus pneumoniae type 57 (Danish type 19A)	249-X// Capsular PS
Streptococcus pneumoniae type 68 (Danish type 9V)	253-X// Capsular PS
Streptococcus pneumoniae type 70 (Danish type 33F)	257-X// Capsular PS
Yersinia pestis KM218-6C	OS
Yersinia pestis KM260(11)-yjhW-6C	OS
Yersinia pestis KM260(11)-wabD/waaL	OS
Yersinia pestis KM260(11)-wabC/waaL	OS
Yersinia pseudotuberculosis 85pCad-37C	OS
Yersinia pseudotuberculosis 85pCad-20C	OS
Yersinia pseudotuberculosis O:2a	PS
Yersinia pseudotuberculosis O:2a-dhmA	PS
Yersinia pseudotuberculosis O:2c	PS
Yersinia pseudotuberculosis O:3	PS
Yersinia pseudotuberculosis O:4b	PS
Proteus vulgaris O2 (OX2)	PS
Proteus mirabilis O3ab (S1959)	PS
Proteus mirabilis O5 (PrK 12/57)	PS
Proteus mirabilis O9 (PrK 18/57)	PS
Proteus mirabilis O11 (9B-m)	PS
Proteus penneri O17 (16)	PS
Proteus mirabilis O18 (PrK 34/57)	LPSOH
Proteus mirabilis O20 (PrK 38/57)	LPSOH
Proteus penneri O31ab (28)	PS
Proteus mirabilis O33 (D52)	PS
Proteus mirabilis O43 (PrK 69/57)	PS
Proteus vulgaris O47 (PrK 73/57)	Not stated
Proteus mirabilis O49 (PrK 75/57)	PS
Proteus mirabilis O54ab (OE)	PS
Proteus penneri O73ac (75)	PS
Proteus vulgaris O76 (HSC438)	PS
Shigella flexneri type 1a	PS
Shigella flexneri type 1b	PS
Shigella flexneri type 2a	PS
Shigella flexneri type 2b	PS
Shigella flexneri type 3a	PS
Shigella flexneri type 3b	PS
Shigella flexneri type 4a	PS
Shigella flexneri type 4b	PS

Shigella flexneri type 5b	PS
Shigella flexneri type 6a	PS
Shigella flexneri type 6	PS
Shigella flexneri type X	PS
Shigella dysenteriae type 1	PS
Shigella boydii type 6	PS
Shigella boydii type 7	PS
Shigella boydii type 8	PS
Shigella boydii type 13	LPSOH
Shigella boydii type 14	LPSOH
Escherichia coli O71	PS
Escherichia coli O85	PS
Escherichia coli O99	PS
Escherichia coli O145	LPSOH
Escherichia coli O107	PS
Salmonella enterica O17	PS
Salmonella enterica O28	PS
Salmonella enterica O47	PS
Salmonella enterica O55	PS
Escherichia coli K92	CPS
Escherichia coli K5	CPS
Escherichia coli K13	CPS
Neisseria meningitidis Group C	CPS
Davanat	
Laminarin	
Yeast Mannan	
Escherichia coli O86	
Galactomannan DAVANT (160102) Pro-Pharmacenti	
Yeast Mannan Sigma M-3640	
1-2 Mannan Acetobacter methanolieus MB135	

References for Supplementary Materials

1. Azzouz D, Omarbekova A, Heguy A, Schwudke D, Gisch N, Rovin BH, et al. Lupus nephritis is linked to disease-activity associated expansions and immunity to a gut commensal. Ann Rheum Dis. 2019 Jul; 78(7):947-956.
2. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012 Aug; 6(8):1621-1624.
3. Determinants of warfarin use and international normalized ratio levels in atrial fibrillation patients in Japan. - Subanalysis of the J-RHYTHM Registry. Circ J. 2011; 75(10):2357-2362.
4. Vaughan EE, Schut F, Heilig HG, Zoetendal EG, de Vos WM, Akkermans AD. A molecular view of the intestinal ecosystem. Curr Issues Intest Microbiol. 2000 Mar; 1(1):1-12.
5. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010 Mar 04; 464(7285):59-65.
6. Kuczynski J, Stombaugh J, Walters WA, Gonzalez A, Caporaso JG, Knight R. Using QIIME to analyze 16S rRNA gene sequences from microbial communities. Curr Protoc Bioinformatics. 2011 Dec; Chapter 10:Unit 10 17.
7. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010 May; 7(5):335-336.
8. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010 Oct 01; 26(19):2460-2461.
9. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007 Aug; 73(16):5261-5267.
10. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One. 2010 Mar 10; 5(3):e9490.
11. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. ISME J. 2011 Feb; 5(2):169-172.
12. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 2013; 8(4):e61217.
13. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016 Jul; 13(7):581-583.
14. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013 Jan; 41(Database issue):D590-596.
15. Gonzalez ME, Schaffer JV, Orlow SJ, Gao Z, Li H, Alekseyenko AV, et al. Cutaneous microbiome effects of fluticasone propionate cream and adjunctive bleach baths in childhood atopic dermatitis. J Am Acad Dermatol. 2016 Sep; 75(3):481-493 e488.
16. Benjamini Y. Controlling the false discovery rate: a proactical and useful approach to multiple testing. J Royal Stat Soc Ser. 1995(57):289-300.
17. Newson R. Confidence intervals for rank statistics: Somers'D ans extensions. Stata J. 2006; 6:309-334.
18. Png CW, Linden SK, Gilshenan KS, Zoetendal EG, McSweeney CS, Sly LI, et al. Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. Am J Gastroenterol. 2010 Nov; 105(11):2420-2428.
19. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018 Sep 1; 34(17):i884-i890.

20. Low AJ, Koziol AG, Manninger PA, Blais B, Carrillo CD. ConFindr: rapid detection of intraspecies and cross-species contamination in bacterial whole-genome sequence data. PeerJ. 2019; 7:e6995.
21. Beghini F, McIver LJ, Blanco-Miguez A, Dubois L, Asnicar F, Maharjan S, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife. 2021 May 4; 10.
22. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017 Jun; 13(6):e1005595.
23. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 Oct 5; 215(3):403-410.
24. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014 Jul 15; 30(14):2068-2069.
25. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010 Mar 15; 26(6):841-842.
26. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. Bioinformatics. 2014 Oct; 30(19):2811-2812.
27. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012 May; 19(5):455-477.
28. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018 Nov 30; 9(1):5114.
29. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019 Nov 14; 20(1):238.
30. Oksanen JGB, Friendly M, Kindt R, and Wagner H. vegan community ecology package version 2.5-7 November 2020. 2020.
31. Seemann T. Snippy: fast bacterial variant calling from NGS reads. 2015 [cited; Available from:
32. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014 May 1; 30(9):1312-1313.
33. Hess N, Waldow F, Kohler TP, Rohde M, Kreikemeyer B, Gomez-Mejia A, et al. Lipoteichoic acid deficiency permits normal growth but impairs virulence of Streptococcus pneumoniae. Nat Commun. 2017 Dec 12; 8(1):2093.
34. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, et al. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol. 2012 Oct; 30(10):918-920.
35. Herzog R, Schwudke D, Schuhmann K, Sampaio JL, Bornstein SR, Schroeder M, et al. A novel informatics concept for high-throughput shotgun lipidomics based on the molecular fragmentation query language. Genome Biol. 2011; 12(1):R8.
36. Chen S Z, Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor Bioinformatics. 2018; 34:i884-i890.
37. Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. Microbial strain-level population structure and genetic diversity from metagenomes. Genome Res. 2017 Apr; 27(4):626-638.
38. Sorbara MT, Littmann ER, Fontana E, Moody TU, Kohout CE, Gjonbalaj M, et al. Functional and Genomic Variation between Human-Derived Isolates of Lachnospiraceae Reveals Inter- and Intra-Species Diversity. Cell Host Microbe. 2020 Jul 8; 28(1):134-146 e134.
39. Hall AB, Yassour M, Sauk J, Garner A, Jiang X, Arthur T, et al. A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. Genome Med. 2017 Nov 28; 9(1):103.