

UCSF

UC San Francisco Previously Published Works

Title

Gut microbiome of multiple sclerosis patients and paired household healthy controls reveal associations with disease risk and course.

Permalink

<https://escholarship.org/uc/item/9rg2j1z4>

Journal

Cell, 185(19)

Publication Date

2022-09-15

DOI

10.1016/j.cell.2022.08.021

Peer reviewed



Published in final edited form as:

Cell. 2022 September 15; 185(19): 3467–3486.e16. doi:10.1016/j.cell.2022.08.021.

Gut microbiome of multiple sclerosis patients and paired household healthy controls reveal associations with disease risk and course

Xiaoyuan Zhou¹, Ryan Baumann¹, Xiaohui Gao¹, Myra Mendoza¹, Sneha Singh¹, Ilana Katz Sand², Zongqi Xia³, Lau M. Cox⁴, Tanuja Chitnis⁴, Hongsup Yoon^{5,6}, Laura Moles⁷, Stacy J. Caillier¹, Adam Santaniello¹, Gail Ackermann⁸, Adil Harroud¹, Robin Lincoln¹, Refujia Gomez,

Antonio González Peña⁸, Elise Digga², Daniel Joseph Hakim⁹, Yoshiki Vazquez-Baeza¹⁰, Karthik Soman¹, Shannon Warty¹, Greg Humphrey⁸, Mauricio Farez¹¹, Lisa Ann Gerdes¹², Jorge R. Oksenberg¹, Scott S. Zamvil¹, Siddharthan Chandran¹², Peter Connick¹², David Otaegui⁷, Tamara Castillo-Triviño^{7,13}, Stephen L. Hauser¹, Jeffrey M. Gelfand¹, Howard L. Weiner⁴, Reinhard Hohlfeld⁵, Hartmut Wekerle⁶, Jennifer Graves¹⁴, Amit Bar-Or¹⁵, Bruce A.C. Cree¹, Jorge Correale¹¹, Rob Knight¹⁰, Sergio E. Baranzini¹, iMSMS Consortium*

¹Weill Institute for Neurosciences. Department of Neurology, University of California, San Francisco, CA, USA

*Lead Contact / Correspondence: Sergio.Baranzini@ucsf.edu (Sergio E. Baranzini).

Author Contributions

Conceptualization, S.E.B., I.K.S., S.C., S.L.H., B.A.C.C., H.L.W., R.H., H.W., J.C., R.K.

Methodology, R.B., Y.V-B, S.J.C., A.G.P, G.H., G.A., R.K.

Software, X.Z., A.S., D.J.H.,

Validation, X.G., R.B.,

Formal Analysis, X.Z., D.J.H., S.E.B.

Investigation, X.Z., I.K.S, Z.X., L.M.C., T.C., H.Y, L.M., A.H., M.F, L.A.G., J.R.O., S.S.Z., D.O., S.L.H., B.A.C.C., H.L.W., R.H.,

H.W., J.G., J.C.,

Resources, R.L., X.G., M.M., S.S., Z.X., S.J.C, A.H., E.D., M.F, L.A.G., J.R.O., S.S.Z., S.C., P.C., T.C-T, S.L.H., J.M.G., B.A.C.C.,

J.G., A.B-O, J.C,

Data Curation, X.Z., M.M., S.S., A.S., R.G., E.D.,

Writing – Original Draft, X.Z., S.E.B.

Writing – Review & Editing, X.Z., I.K.S., Z.X., L.M.C., S.L.H., J.C., S.E.B.

Visualization, X.Z.,

Supervision, S.E.B.

Project Administration, R.L., S.W.,

Funding Acquisition, S.E.B., S.L.H.

Consortia

The members of iMSMS are:

Xiaoyuan Zhou, Ryan Baumann, Xiaohui Gao, Myra Mendoza, Sneha Singh, Ilana Katz Sand, Zongqi Xia, Laura M. Cox, Tanuja Chitnis, Hongsup Yoon, Laura Moles, Stacy J. Caillier, Adam Santaniello, Gail Ackermann, Adil Harroud, Robin Lincoln, Refujia Gomez, Antonio González Peña, Elise Digga, Daniel Joseph Hakim, Yoshiki Vazquez-Baeza, Karthik Soman, Shannon Warty, Greg Humphrey, Mauricio Farez, Lisa Ann Gerdes, Jorge R. Oksenberg, Scott S. Zamvil, Siddharthan Chandran, Peter Connick, David Otaegui, Tamara Castillo-Triviño, Stephen L. Hauser, Jeffrey M. Gelfand, Howard L. Weiner, Reinhard Hohlfeld, Hartmut Wekerle, Jennifer Graves, Amit Bar-Or, Bruce A.C. Cree, Jorge Correale, Rob Knight, Sergio E. Baranzini
(See Document S1 for consortium member affiliations).

Declaration of Interests

The authors declare no competing interests.

Inclusion and Diversity

We worked to ensure that the study questionnaires were prepared in an inclusive way. One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design, analysis, and/or interpretation of the work.

²Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

³Department of Neurology, University of Pittsburgh, Pittsburgh, PA, USA

⁴Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

⁵Institute of Clinical Neuroimmunology, Biomedical Center and University Hospitals, Ludwig-Maximilians-Universität München, and Munich Cluster of Systems Neurology (SyNergy), München, Germany

⁶Department Neuroimmunology, Max Planck Institute (MPI) of Neurobiology, Munich, Germany

⁷Neurosciences Area, Biodonostia Health Research Institute, San Sebastián, Spain

⁸Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA

⁹Department of Bioinformatics and Systems Biology, University of California, San Diego, La Jolla, CA, USA

¹⁰Center for Microbiome Innovation, University of California, San Diego, La Jolla, CA, USA

¹¹Department of Neurology, Institute for Neurological Research Dr. Raul Carrea (FLENI), Buenos Aires, Argentina

¹²Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK

¹³Department of Neurology, Hospital Universitario Donostia and Neurosciences Area, Biodonostia Health Research Institute, San Sebastián, Spain

¹⁴Department of Neurosciences, University of California, San Diego, CA, USA

¹⁵Department of Neurology, University of Pennsylvania, Pennsylvania, PA, USA

Abstract

Changes in gut microbiota have been associated with several diseases. Here the international Multiple Sclerosis Microbiome Study (iMSMS) studied the gut microbiome of 576 MS patients (36% untreated), and genetically unrelated household healthy controls (1,152 total subjects). We observed a significantly increased proportion of *Akkermansia muciniphila*, *Ruthenibacterium lactatiformans*, *Hungatella hathewayi* and *Eisenbergiella tayi* and decreased *Faecalibacterium prausnitzii* and *Blautia* species. The phytate degradation pathway was over-represented in untreated MS, while pyruvate-producing carbohydrate metabolism pathways were significantly reduced. Microbiome composition, function and derived metabolites also differed in response to disease modifying treatments. The therapeutic activity of interferon- β may in part be associated to upregulation of short chain fatty acid transporters. Distinct microbial networks were observed in untreated MS and healthy controls. These results strongly support specific gut microbiome associations with MS risk, course and progression, and functional changes in response to treatment.

Introduction

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system (CNS) characterized by demyelination, axonal damage, and progressive neurologic disability. The etiology and pathogenesis of MS is complex and remain elusive, although both genetic and environmental factors are involved. Gut microbiota, an important modulator of the immune response (Geva-Zatorsky et al., 2017) and brain function, has emerged as a likely environmental contributor to MS. (Esmail Amini et al., 2020; Kadowaki and Quintana, 2020; Probstel and Baranzini, 2018)

Alterations in commensal gut microbiota have been linked to many inflammatory conditions. (Honda and Littman, 2016) Numerous studies including ours have shown both depletion and enrichment of certain bacteria in MS patients compared with healthy controls, (Berer et al., 2017; Chen et al., 2016a; Cox et al., 2021; Jangi et al., 2016) suggesting certain taxa might be associated to either disease pathogenesis or progression. It remains uncertain whether the disease results from microbial alterations, or viceversa. Mouse and human studies indicate that microbiota can potentially affect the onset and progression of diseases mediated by different immune effector cells and soluble metabolic, immune, and neuroendocrine factors modulated by gut microbes. (Camara-Lemarroy et al., 2018; Probstel and Baranzini, 2018)

Although microbial changes in MS have been detected across different studies, most of the alterations were reported in relapsing-remitting MS (RRMS), whereas few studies investigated the microbiome in progressive MS. Furthermore, it is difficult to identify a common pattern since results are rarely concordant (Cox *et al.*, 2021; Kozhieva et al., 2019; Reynders et al., 2020) The gut microbiota can also be altered by drugs with either beneficial or undesirable effects. Many common drugs have antimicrobial effects, or exert a large impact on the composition of gut microbiome suggesting that therapeutic efficacy may be due to the effects of DMTs on gut microbiota. (Castillo-Alvarez et al., 2018; Cox *et al.*, 2021; Jangi *et al.*, 2016; Maier et al., 2018; Sand et al., 2019)

Current microbiome studies in MS are limited by the relatively small size of the cohort analyzed, and inadequate handling of multiple confounding factors, such as genetic heterogeneity of participants, geographic location, disease subtype, treatment and diet. Also, many studies rely on 16S ribosomal RNA sequencing, which offers low resolution to identify MS associated species. To overcome these challenges, the international MS Microbiome Study (iMSMS) is systematically recruiting MS patients and household healthy controls in the US, Europe and South America. The advantages of the household-controlled experimental design, sequencing method and handling of confounding factors (e.g. geographic location and diet) on gut microbiome were recently reported in a pilot cohort of 128 patient:control pairs. (Zhou et al., 2020) Here we present a large microbiome study of MS and healthy controls (n= 576 pairs), and investigate relationships with MS susceptibility, progression, and treatment.

Results

A total 576 pairs of MS patients and genetically unrelated household healthy controls (HHCs) were recruited between September 2015 and January 2019 from seven sites (recruiting centers) located in San Francisco, Boston, New York, Pittsburgh, Buenos Aires, Edinburgh and San Sebastián (Figure 1, Table 1, Table S1). The first 128 pairs were recruited before October 2016 (Cohort1(Zhou *et al.*, 2020)) and the subsequent 448 pairs were recruited before January 2019 (Cohort 2). Among the 576 MS patients, 209 (36%) were untreated and 367 (63%) were treated with a disease modifying therapy (DMT). Treatments included oral agents Fingolimod (n=71), and dimethyl fumarate (DMF, n=86);, injectables glatiramer acetate (GA, n=68) and interferon (IFN, n=87); and infusion agents anti-CD20 monoclonal antibodies (n=28), and natalizumab (n=27). Of the total of 576, 437 (76%) patients had RRMS, 68 (12%) secondary progressive MS (SPMS) and 71 (12%) primary progressive (PPMS). Given the heterogeneity in the assessment of patients with SPMS and PPMS, they were combined into a single category, progressive MS (PMS, n=139, 24%) for subsequent analyses.

All participants completed a clinical survey to report the disease status and treatment, and a high proportion of participants (94%, n = 1,086) completed the subject survey to report the demographics, medication, lifestyle and physiology factors (Figure 1A, and Table S1). Most participants (90%, n= 1,034) also completed the online food frequency questionnaire (FFQ). A summary of dietary questionnaires and the dietary intake is provided in Table S2. The Healthy eating index (HEI2015 with 10 components) was also calculated for all qualifying participants (Table S3).

Altered Gut microbial composition in MS

We first used 16s rRNA data to study the global microbial composition (α - and β -diversity). The V4 region of the bacteria 16S ribosomal RNA gene was amplified on an Illumina MiSeq platform using the Earth Microbiome Project protocol.(Caporaso et al., 2012) Amplicon reads from two cohort samples were analyzed using QIITA(Gonzalez et al., 2018; Hillmann et al., 2018) to combine the forward and reverse reads, trim short reads of less than 150bp and assign filtered reads to amplicon sequencing reads (ASVs) using default Deblur parameters against Greengenes (version 13.8 at 99% identity) as described in QIIME2 documents.(Caporaso et al., 2010)16S rRNA sequencing has been more commonly used in microbiome studies to date, thus several well-established databases (e.g. Greengenes(DeSantis et al., 2006)) are available. The 576 pairs were processed and sequenced in two cohorts (128 pairs in Cohort 1 and 448 pairs in Cohort 2) (STAR Methods, Table S4). For the first cohort, Q-tip samples (i.e. dry) and snap frozen (i.e. wet) samples were processed using the QIAamp PowerFecal DNA Kit (ref 12830–50). The second cohort samples were processed using the MagAttract PowerSoil DNA EP Kit (ref 27100–4-EP).

The ASVs characterized by 16s rRNA sequencing were rarefied to 10,000 sequences per participant sample for microbial diversity analysis. α -diversity was measured by Shannon(Shannon, 1997) and Chao1(Chao, 1984) indices (Table S4). The microbial composition and diversity were highly correlated in duplicate samples sequenced in the two cohorts (Figure S1A–C), and also in duplicated samples processed by different DNA

isolation methods (Figure S1D–F), thus allowing us to merge all sequencing samples for a joint analysis. After removing samples with low coverage (<10,000 reads), 500 pairs of MS and household control participant samples were used for diversity analysis (Table S4).

No significant difference in α -diversity was observed between MS and HHC groups as measured by Shannon and Chao1 metrics ($n = 1000$, ANOVA $p > 0.05$) (Figure 1B, Figure S2A). We also found no significant difference in α -diversity across patient:HHC pairs of RRMS ($n=766$), PMS ($n= 234$, ANOVA $p > 0.05$) (Figure 1B), untreated MS ($n=358$), and treated MS ($n=642$, ANOVA $p > 0.05$) (Figure S2B). Intriguingly, β -diversity-based sample clustering revealed a significant difference in MS regardless of treatment, and also differed in untreated or treated MS group status compared to their HHC (PERMANOVA FDR < 0.05) (Figure 1C). No significant difference was observed between untreated and treated MS patients (PERMANOVA FDR > 0.05) (Figure 1C). Different microbial communities were also observed across patient:HHC pairs of RRMS and PMS patients, and when comparing RRMS versus PMS patients (PERMANOVA FDR < 0.05) (Figure 1D).

We next tested how much of the variance in microbial diversity (weighted uniFrac distances) was explained by the host confounders, including demographics, lifestyle, disease, medication, and physiology factors (Table S4). (Vujkovic-Cvijin et al., 2020; Zhou et al., 2020) Not surprisingly, the recruitment site showed a significant and dominant effect on the microbial composition (PERMANOVA FDR < 0.05) (Figure 1E), as we and other studies have reported. (Gaulke and Sharpton, 2018; Zhou et al., 2020) By checking the gut microbial α -diversity in individuals from each recruitment site, we observed lower microbial diversity in both healthy and MS participants from New York whereas a higher diversity in participants from San Francisco and San Sebastián (ANOVA $p > 0.05$) (Figure S2C). We hypothesize that these differences in microbial diversity might be associated with different dietary habits or the FFQ not completely capturing the nuances of diets different countries (see dietary analysis). Microbial differences associated with geography were also shown by the PCoA of the microbiome β -diversity (Figure S2D). The second and third largest component was explained by disease status (RRMS/HHC, PMS/HHC) and treatment status (treated MS/HHC, untreated MS/HHC), implying an altered gut microbiome in MS patients versus HHC as well as an effect of treatment on changing microbial structure.

Age, sex and BMI also showed significant effects on microbial compositions (Figure 1E). Our household design effectively reduces age-associated variation as the great majority of household participants are spouses of comparable age (Table 1). Smoking and education also exerted significant effects, but these effects were variable across recruitment sites (e.g. participants from San Francisco, Boston and New York are less likely to smoke and reported higher education) (Figure S2E–F). A smaller effect was identified by medication use and MS comorbidities (Figure 1E) as MS patients tend to use more medications and have depression or anxiety (Fisher's exact test $p < 0.001$) (Figure S2G). No significant microbial divergence was related to factors such as household pets, birth method, or asthma in our study (Table S4).

Disease associated microbial changes adjusted for confounders

Shallow shotgun sequencing with as little as 0.5 million sequences per sample has been shown to be a powerful and cost-effective alternative to whole metagenome sequencing. (Hillmann *et al.*, 2018; Shapira *et al.*, 2009) 1 ng of input DNA was used in a 1:10 miniaturized Kapa HyperPlus protocol. The pooled library was sequenced as a paired-end 150-cycle run on an Illumina HiSeq2500 v2 (cohort1) or NovaSeq 6000 (cohort2) at the UCSD IGM Genomics Center with sequencing depth 0.5 million reads per sample. Due to the high correlation between 16S rRNA and shallow sequencing at both phylum and genus levels (Zhou *et al.*, 2020) we used shallow metagenomic data to identify disease-associated taxa and their functions. To achieve this, we performed a mixed linear regression model on metagenomics taxa (Table S5) in untreated MS versus HHCs. For this analysis, microbial composition and pathway were normalized as relative abundance and further transformed with a variance-stabilizing arcsine square-root transformation. (Lloyd-Price *et al.*, 2019; Morgan *et al.*, 2012; Sokal, 1982) The organism-pathway-reaction-compound network was built by Scalable Precision Medicine Oriented Knowledge Engine (SPOKE), a large graph with multiple types of nodes and relationships integrated from more than 30 publicly available databases covering human and bacterial molecular interactions. (Himmelstein *et al.*, 2017; Nelson *et al.*, 2021)

Compared to HHCs, 7 species were significantly reduced in untreated MS, whereas 16 species including were significantly increased in this group (FDR < 0.05) (Figure 2A). We observed a similar trend for these same species in untreated RRMS and progressive MS, although some did not reach significance likely due to the smaller sample size and relatively higher interindividual heterogeneity of these groups. Intriguingly, a larger decrease of *F. saccharivorans* and *F. prausnitzii* and a larger increase of *R. lactatiformans*, *H. hathewayi* and *Eisenbergiella tayi* were found in untreated PMS compared to untreated RRMS (Figure 2B), suggesting the alteration of these species could be associated with disease progression.

We next tested the correlation between microbiota and the Multiple Sclerosis Severity score (MSSS), adjusting for age and BMI. Several species showed correlations with disease severity in untreated RRMS and PMS patients (Spearman's correlation $p < 0.05$) (Figure 2C–D), consistent with a recent study. (Cox *et al.*, 2021) Specifically, some Bacteroides species were correlated with lower MSSS in RRMS and short-chain fatty acid producers like *Butyrivibrio*, *Clostridium* and *Ruminococcus* were correlated with lower MSSS in PMS. Conversely, *Collinsella aerofaciens*, shown to increase disease severity in collagen-induced arthritis mice (Chen *et al.*, 2016b) was associated with a higher MSSS in RRMS patients. Consistent with studies (Larsen, 2017) showing increased inflammatory properties of several *Prevotella* species (including *P. buccalis*, *P. corporis*, *P. disiens*, and *P. copri*) in chronic inflammation, we found these were associated with higher MSSS in progressive MS patients. Finally, *Streptococcus thermophilus*, *Azospirillum sp. 47_25* and *Rhodospirillum sp. UNK.MSG-17* were also correlated with MSSS albeit in different direction for RRMS (positive) and PMS (negative).

Functional alterations in the gut microbiome of untreated MS patients

We next explored the functional potential of the MS gut metagenome using the HUMAnN2 workflow. Generally, all microbes perform four core metabolic pathways, biosynthesis, degradation, energy metabolism and macromolecule modification (Figure S3A). No significant differences in the functional potential of gut microbes were observed. PCA analysis of the abundance of functional pathways also failed to identify significant changes between untreated MS patients (RRMS or PMS) and HHCs (PERMANOVA $p > 0.05$) (Figure S3B). However, when testing individual pathways, we found that phytate degradation I (PWY-4702), was significantly more represented in MS patients (mixed linear regression FDR < 0.05) (Figure 3A). Several species, including *Akkermansia muciniphila*, *Escherichia coli* and *Cronobacter sakazakii*, have the ability to degrade phytate via this pathway. We found *A. muciniphila* (Figure 2A) and two of its encoded proteins in this pathway (4-phytase, Amuc_0145, and Inositol-1-monophosphatase, Amuc_1242) significantly more represented in untreated RRMS and progressive MS patients (Figure 3B). As multiple (and sometime opposing) functional capabilities have been reported for *A. muciniphila* strains, we implemented the Metagenomic Intra-Species Diversity Analysis System (MIDAS) (Nayfach et al., 2016) to estimate strain-level genomic variation. In total, 58 samples had sufficient sequencing coverage allowing us to identify single nucleotide polymorphisms (SNPs) and gene content from *A. muciniphila* (Figure S4). To distinguish possible strains from these samples, we compared 2,913 genes (presence/absence) and identified four gene clusters and two sample clusters. None of the sample clusters was significantly correlated with disease status, sex, treatment status or geographic site (Chi-squared test $p > 0.05$). The majority of *A. muciniphila* genes were shared across samples, (i.e. core genes), but some genes showed a distinct pattern. Functional analysis revealed that “Sulfite reductase [NADPH] hemoprotein beta-component (EC 1.8.1.2)”, encoded by the *cysI* gene, was present in cluster1 but not cluster2. While additional studies are needed to establish their relevance to MS, we were able to identify the presence of at least two *A. muciniphila* strains with differences in sulfur metabolism.

Conversely, 6 pathways were more represented in HHC, mostly explained by the increase of *Faecalibacterium prausnitzii* (mixed linear regression FDR < 0.05) (Figure 3B, Figure S3B). We also found four other carboxylates metabolism pathways which produce pyruvate via protein 2-dehydro-3-deoxy-phosphogluconate aldolase (EC 4.1.2.14), were enriched in healthy controls (Figure 3C). This protein was identified in *Faecalibacterium prausnitzii* (FP2_23290, D4K064) and significantly decreased in both untreated RRMS and PMS patients (Paired T-test $p < 0.05$) (Figure 3D). Finally, by integrating the metabolic pathways into higher-class pathway level, we identified that cyclitols degradation and fermentation to acids were more abundant in untreated MS patients, while carboxylates degradation, lysine synthesis, S-adenosyl-L-methionine biosynthesis and sucrose degradation were enriched in healthy controls (mixed linear regression FDR < 0.05) (Figure 3E).

We found different pathways associated with disease severity in untreated RRMS and PMS (Spearman’s correlation $p < 0.05$) (Figure 3F). “PWY-4981: L-proline biosynthesis II (from arginine)” was positively correlated with higher MSSS, mostly explained by the abundance of *Collinsella aerofaciens* (Figure S3C) which was also correlated with higher MSSS

(Figure 2C). Conversely, lower representation of the “PWY-5097: L-lysine biosynthesis VI” pathway (Figure 3A) was correlated with a lower MSSS PMS mostly explained by *Bacteroides* species and *F. prausnitzii* (Figure S3D).

Specific interacting microbial communities were enriched in MS

We next computed species-species co-abundance networks using SparCC(Friedman and Alm, 2012) method (in R using *SpiecEasi* package(Kurtz et al., 2015)), which is a tool to infer linear relationships with high precision for high diversity compositional data. SparCC correlations were adjusted for age, sex and BMI using *cor2por* function from R package “*corpor*”. In total, 1677 species were used for the analysis, resulting in 116,397 correlations across 1372 species in MS patients and 105,304 correlations across 1375 species in HHC (absolute Sparcc $r > 0.1$, FDR < 0.05 , Table S6). After filtering out correlations with $r < 0.4$ (based on the network centrality distribution) (Figure S5A) and subnetworks with fewer than 2 species, we identified a network of 773 taxa with 5688 correlations in HHC (dominated by 555 Firmicutes and 196 Bacteroidetes species), and a network of 786 taxa with 6742 correlations in MS (dominated by 549 Firmicutes and 197 Bacteroidetes species) (Figure S5B). Notably, the majority of taxa ($n = 702$) and correlations ($n = 4131$) between MS and HHC microbial networks overlapped (Figure S5C), suggesting that the fundamental role of commensal microbes remains stable even under different biological conditions.

Cohort-specific analysis revealed 215 correlations across 119 species in untreated MS patients (mean $r = 0.78$, FDR < 0.05), and 195 correlations across 139 species in HHCs (mean $r = 0.783$, FDR < 0.05) (Figure 4A–B). Cohort-specific species were linked to their MetaCyc pathways. As many species share pathways, we focused on those that are unique to the cohort specific species. Species from the same phylum were clustered together in both MS and HHC networks, suggesting a cooperative role of these species in response to the environment. Remarkably, we observed different Firmicutes/Bacteroidetes (F/B) ratio for the MS-specific network (F/B=2.5) and HHC-specific network (F/B=1.03) (hypergeometric test $p < 0.01$). Interestingly, 45 unique species (largely dominated by *Bacteroides* and *Prevotella*) composed the HHC network (Figure 4C).

Surprisingly, among 21 significantly altered species (untreated MS versus HHC, Figure 2A), seven were identified in both the MS and healthy specific networks, and only one species (*Varibaculum cambriense*), was found in the MS network (Figure 4D). This suggests that co-abundance relationships and differential microbial abundance reflect orthogonal information. While the group-specific species did not show significant differences in abundance, some have unique functions (Figure 4E–F). For example, *[Clostridium] innocuum* and *Salmonella enterica* (with potential roles in drug resistance and pathogenicity) were specific to the MS network, while *Bacteroides vulgatus*, *Bacteroides thetaiotaomicron*, *Prevotella fusca* and *Prevotella denticola* (with potential roles in glycan biosynthesis), were specific to the HHC network. Altogether, these results suggest that microbial co-abundance network analyses can identify highly interacting communities that may contribute to health or disease status.

Impact of treatment on gut microbiota

We next evaluated how DMT may affect gut microbiome composition in RRMS patients receiving any of 6 commonly used treatments in our study. Overall, the microbial composition measured by β -diversity did not differ between treated and untreated RRMS patients (except for the interferon treated group). However, significant differences in β -diversity were observed when patients within each treatment group were compared to their corresponding HHC (PERMANOVA) (Figure 5A).

Due to the heterogeneity of treated and untreated RRMS patients recruited from multiple locations and unequal sample sizes of these groups, we mainly focused our analyses on gut microbiome by comparing untreated or treated RRMS groups to their HHCs (mixed linear regression) (Figure 5B–C). A direct comparison between untreated RRMS and treated RRMS was shown in Figure S6A–B. Intriguingly, the microbial changes observed in untreated RRMS patients (versus HHCs) were not replicated in treated RRMS (versus HHCs). Specifically, several taxa increased in untreated RRMS subjects showed no difference within DMTs groups, including *Parabacteroides merdae* CAG:48, *A. muciniphila* and other *Akkermansia* species. Use of DMTs was also associated with changes in multiple taxa that were not significantly different between MS and HHC. For example, DMF, which is hydrolyzed into monomethyl fumarate (MMF) before exerting its therapeutic effect, specifically reduced *Bacteroides stercoris*, *Clostridium* and *Eubacterium* species, and fingolimod specifically reduced *Bacteroides finegoldii* CAG:203, *Roseburia faecis* and *Blautia* species. Interferon- β treatment, thought to decrease proinflammatory cytokines and prevent the migration of activated T cells across the blood-brain barrier, was associated with dysbiosis of short-chain fatty acid-producing bacteria like *Ruminococcus sp.*, *Clostridium sp.*, *Faecalibacterium prausnitzii*, *Roseburia inulinivorans* and *Roseburia intestinalis* while also increased *Parabacteroides distasonis*, which have been shown to have multiple metabolic benefits in obesity.(Wang et al., 2019) Notably, *Bacteroides uniformis* was significantly increased by interferon treatment but reduced by GA and natalizumab therapy. This bacterium was reported to be associated with MS(Miller et al., 2015) but also with a protective role in obesity.(Lopez-Almela et al., 2021) GA exerted a modest impact on gut microbes compared to other DMTs. Lastly, infusion of natalizumab or anti-CD20 monoclonal antibody altered gut microbes significantly. *Phascolarctobacterium sp. CAG:207* was increased while *Prevotella* species and *Bifidobacterium longum* were decreased in response to natalizumab. A reduction of *Bacteroides finegoldii* CAG:203 and *Blautia sp. CAG:37* were observed in anti-CD20-treated patients.

Based on metagenomic sequencing, numerous metabolic pathways appeared to be altered in response to the different DMTs, and many of them are included in the same high-class pathway (Figure S6C). We found that pathways related to lysine synthesis, sugar nucleotides and unsaturated fatty acids biosynthesis were decreased significantly in untreated RRMS but modulated differently by the various DMTs. Of interest, the increased cyclitols degradation pathways in untreated RRMS remained highly abundant even after treatment (Figure 5C and Figure S6C). We also identified various metabolic pathways that were differentially modulated by specific therapies. For example, DMF use increased heme synthesis and enzyme cofactor biosynthesis pathways. In addition, DMF and interferon use was associated

with an increase in L-ornithine biosynthesis and carrier biosynthesis. Furthermore, GA use was associated with increased peptidoglycan biosynthesis and natalizumab with increased lipid biosynthesis, whereas a decrease of guanosine nucleotides degradation pathway was associated to Fingolimod treatment (Figure 5C and Figure S6C). Altogether, disease modifying therapies showed significant and specific impact on gut microbiome both structurally and functionally, indicating the importance of stratifying microbiome analyses by treatment status.

To further investigate the mechanism of disease modifying therapies (DMT) in MS and their interactions with gut microbiota, we performed metabolomic profiling in untreated RRMS patients (N=79), and in those treated with dimethyl fumarate (n=47), fingolimod (n=39), GA (n=31), and interferon- β (n=49), as well as their corresponding household healthy controls. A panel of global metabolites and 8 targeted short-chain fatty acids (SCFAs) in both feces and serum samples were measured using ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS) by Metabolon Inc. (Durham, North Carolina) (Table S7).

We found 31 metabolites significantly different between untreated patients and controls, or in response to at least one MS drug (mixed linear regression $p < 0.05$) (Figure 6A). Consistent with their expected functions and origin, we found higher variability in fecal metabolites compared their corresponding serum levels (with the notable exception of increased serum fumarate in DMF-treated patients). We also identified significant changes in the levels of 8 SCFAs in either serum or stool for at least one group (mixed linear regression $p < 0.05$) (Figure 6B). Remarkably, the vast majority of changes in microbiota-derived fecal metabolites were towards lower levels among MS patients and even more significantly in response to DMTs (except for GA, Figure 6A). Higher levels of metabolic dysfunction have been reported to be associated with increased disability in MS. (Lazzarino et al., 2017; Villoslada et al., 2017) We found no difference of disease severity (measured by global MS severity score) among RRMS patients (treated or untreated) (Figure 6C). This suggests the altered metabolites reported here are in response to the MS drugs, not the disease process. Interestingly, we found specific signatures of microbe-derived metabolites (stool) in RRMS patients in response to each treatment. The most notable changes in gut metabolites were induced in response to Fingolimod and IFN- β .

While Fingolimod is an oral drug, and changes to the gut microbiota might be expected, the profound metabolic signature of IFN- β (an injectable) was most intriguing. A functional analysis of the 23 IFN- β -associated metabolites, revealed a significant enrichment in pathways involving amino acid metabolism (e.g. "Arginine biosynthesis"), carbohydrate (i.e. "Citrate cycle"), nucleotide (i.e. "Purine"), and energy ("Nitrogen") metabolism (MetaboAnalyst pathway enrichment FDR < 0.05) (Figure 6D). In contrast, GA exerted an almost null impact on stool-derived metabolites. These findings are in agreement with previous studies, in which only modest transcriptional changes were observed in PBMCs after treatment with GA compared to IFN- β . (De Jager et al., 2009; Ottoboni et al., 2012) Also of interest, these distinct metabolomic alterations were consistent with functional predictions derived from shotgun sequencing (Figure 5B–C).

We noted that pyruvate was significantly decreased in both feces and serum samples from RRMS patients treated with Fingolimod. Interestingly, this finding correlates well with the significant depletion of taxa containing the pathway “CARBOXYLATES-DEG” (which produces pyruvate) in Fingolimod treated patients (Figure 5C, Figure S6). We also observed the concentration of fecal SCFA (such as acetate and propionate) was consistently lower in RRMS patients, regardless of treatment (Figure 6B), consistent with our finding of the depletion of *F. prausnitzii* (a major SCFA-producing bacteria) in MS. A decreased amount of fecal SCFA has also been reported in RRMS and PPMS patients in other studies. (Takewaki et al., 2020; Zeng et al., 2019)

Propionate supplementation in MS patients was associated with an increased Treg/Th17 ratio, leading to long-term clinical improvement. (Duscha et al., 2020) Interestingly, we found a significant increase in serum propionic acid (Acetic and Butyric acid also followed this same trend, without reaching statistical significance) in RRMS patients treated with IFN- β (Figure 6E, Figure 6B). Since most SCFAs produced in the colonic lumen are actively transported to the lamina propria and further into the blood stream, (Olsson et al., 2021; Venegas et al., 2019) we hypothesized that IFN- β may increase the intestinal absorption of propionate, as part of its immunomodulatory effect. To address this hypothesis, we searched whether expression of the genes encoding for SCFA transporters MCT1 (*SLC16A1*) and SMCT1 (*SLC5A8*) (Miyachi et al., 2004; Ritzhaupt et al., 1998) were upregulated by IFN- β . The Interferome database (Rusinova et al., 2013) reports an increase of SLC16A1 expression in human bronchial epithelial cells (no data is available for intestinal epithelial cells) treated with IFN- β (Figure 6F), potentially supporting our hypothesis.

Diet and gut microbiome

Diet is thought to explain over 20% of microbial structural variations in humans, implying the potential for dietary strategies in disease management through gut microbiota modulation. (Rothschild et al., 2018) We administered the validated Block 2005 food frequency questionnaire (FFQs) (Block, 2005) to our participants (89.8% completion rate) (Table S2). Recent epidemiologic studies of diet and health outcomes have also focused on the overall diet quality, (Guo et al., 2004) which can be measured by the Healthy Eating Index (HEI-2015), where a higher HEI-2015 score indicates greater diet quality (See Supplemental document).

Significant differences in diet (measured by Jaccard dissimilarity) were associated with BMI, participant household, recruitment site, education and age (PERMANOVA FDR < 0.05) (Figure 7A). Not surprisingly, a higher BMI correlated with a lower HEI-2015 score in both MS patients and healthy individuals (Pearson's correlation $p < 0.05$) (Figure S7A), consistent with evidence that an imbalanced diet exerts a significant influence on weight. (Guo *et al.*, 2004) We also observed that diet quality increased with age (Pearson's correlation $p < 0.05$) (Figure S7A). There is considerable variation in dietary intakes across countries (Figure 7B). In particular we found a significantly lower average HEI-2015 score in participants from Buenos Aires when compared to those in San Francisco, New York, Edinburgh and San Sebastian. While this may indeed indicate a lower health index, it is noteworthy that the FFQ is standardized for the US average participant, and diets in other

parts of the world may not adjust properly to this standard. As gut microbial diversity differed among recruitment sites, we hypothesized that the diversity was influenced by diet. Indeed, we found that higher microbial diversity significantly correlated with a higher HEI-2015 score in both healthy and MS individuals (Pearson's correlation $p < 0.01$) (Figure 7C). However, although participants from Buenos Aires had lower HEI-2015 scores, their microbial diversity remained high compared to other sites, whereas New York had higher HEI-2015 scores but comparatively lower microbial diversity (Figure S2C). This may indicate that standardized questionnaires, even if validated, do not fully capture the wide range of dietary habits from iMSMS participants, but also suggests that the gut microbiota could be influenced by other factors, such as physiological activity, water and air, among others. Also, shifts in oral microbe composition need to be considered as studies have shown oral-derived bacteria can colonize and persist in the intestines.(du Teil Espina et al., 2019; Hatton et al., 2018)

Despite the large variance in dietary habits among participants, we identified a significantly higher diet similarity within household pairs compared to that of random pairs drawn from within the same city (ANOVA FDR < 0.05) (Figure S7B). The lowest diet similarity was found when random pairs of MS and HHC were assembled from different cities, consistent with our previous findings(Zhou *et al.*, 2020) and reflecting distinct dietary habits across cities and countries (Figure 7B). Finally, we observed a significant correlation between education, nonsmoker (or former smoker) status, and female sex with a higher HEI-2015 score (ANOVA FDR < 0.05) (Figure S7C–E), also consistent with findings from previous studies.(Arabshahi et al., 2011; Thorpe et al., 2019)

Although a more similar diet was shared among household participants, the HEI-2015 score of MS patients was significantly higher than those of healthy controls (paired T-test $p < 0.001$) (Figure 7D). However, microbial taxa associated with MS status did not overlap with those associated with diet, thus likely not representing a confounder. Indeed, we specifically assessed which dietary components were consumed differently by MS and healthy participants and whether these differences were associated with species previously shown to be altered in MS. By comparing the ten components from the HEI-2015 (Table S3), we found MS participants consumed more fruit, vegetables and unsaturated fatty acid when compared to HHCs (paired T-test $p < 0.05$) (Figure 7E), which contributed to their scores (Figure 7D). We also found that *Eubacterium eligens* was highly correlated with a higher HEI-2015 score (Pearson's correlation $p < 0.01$) (Figure 7F), and particularly correlated with intake of whole grains, fruit and vegetables (mixed linear regression $p < 0.05$ after adjusting for age, sex, BMI and recruitment site) (Figure S7F), consistent with previous studies showing that *E. eligens* responded significantly to dietary fiber.(Chung et al., 2016) *Faecalibacterium sp.*, *Eubacterium sp.* and *Blautia sp.* were also positively correlated with higher intake of whole grains. Increased *Alistipes obesi* abundance was also correlated with healthier diet (Pearson's correlation $p < 0.05$) (Figure 7F). Interestingly, other studies found low *Alistipes* abundance in individuals with obesity (Thingholm et al., 2019) and associated to higher meat intake,(Garcia-Ribera et al., 2020). Furthermore, *Alistipes* abundance was identified as a predictor of successful weight loss in a two-year intervention (including healthier diet) in adults with obesity,(Louis et al., 2016) suggesting a potential beneficial role of this bacterium in the context of metabolic health. Altogether, although diet does correlate

with changes in the host microbiota, we were able to tease apart the effects of diet and disease, in large part due to the household paired design employed (Figure S7F).

As expected, diet showed a modest effect on MS-associated taxa after controlling for the environmental impact by household design in all three groups (mixed linear regression) (Figure 7G). Still, some disease-associated species were also related to diet. For example, *Ruminococcus torques* was enriched in MS, and showed a negative correlation with sodium intake, whereas no difference in sodium intake was found between MS versus HHCs. *Faecalibacterium prausnitzii* correlated positively with fruit (which MS patients consumed more), but the bacterium remained reduced in MS compared to healthy controls. These examples suggest that these species were more likely related to disease status than diet.

Phytate degradation I (PWY-4702) pathway was found to be overrepresented in MS patients (Figure 2A). Phytate, a plant-based antioxidant compound, is a strong chelator of divalent minerals (e.g. calcium, magnesium, iron and zinc), which bacteria metabolize into myo-inositol, a compound with immunoregulatory properties, (Nerurkar et al., 2020) which was found at lower levels in MS sera and CSF. (Zahoor et al., 2021) Thus, we hypothesized that this bacterial pathway was activated: i) in response to increased dietary intake of divalent minerals by MS patients or; ii) as a compensatory mechanism to produce more myo-inositol. To test this hypothesis, we compared the dietary mineral intake between MS patients and their HHCs but found no significant difference (after adjusting for age, BMI and sex) (Figure S8A).

Finally, we observed MS patients took more vitamin D supplementation than healthy controls (paired T-test $p < 0.001$) (Figure S8B), possibly in response to studies showing an association with reduced risk of developing MS and of disease activity in MS patients. (Munger et al., 2006; Runia et al., 2012) When assessing the impact of vitamin D usage on microbial composition, we were unable to find a correlation. We did find a trend towards a negative correlation with microbial α -diversity for both MS or HHCs samples, but without reaching significance (Figure S8C). Similarly, β -diversity was not significantly influenced by vitamin D intake (Figure S8D).

Discussion

Microbiome composition and function significantly differed across disease subtypes, responded differently to disease modifying treatments, and were modestly associated with diet. We found that the microbial composition was to a lesser extent, associated with factors such as geographic location, age, sex and BMI. The influence of other confounding factors was reduced by our paired household design, thereby potentially enhancing power to identify MS-associated microbial features. In addition to confirming and extending previous reports, (Berer *et al.*, 2017; Cekanaviciute et al., 2017; Chen *et al.*, 2016a; Cox *et al.*, 2021; Jangi *et al.*, 2016) this work provides a large reference dataset that can be used to understand microbial variation across individuals with MS, disease subtypes and in response to different therapeutic interventions.

Consistent with earlier studies, we found no difference of α -diversity between MS patients and healthy individuals (Berer *et al.*, 2017; Cekanaviciute *et al.*, 2017; Jangi *et al.*, 2016). However, in contrast to previous studies, we observed a significant difference of β -diversity in disease status (regardless of treatment status). Interestingly, there was no difference in β -diversity between untreated MS and treated MS, potentially indicating that disease status exerts a stronger effect on gut microbiome than does treatment. (Cox *et al.*, 2021) Overall, our findings revealed a robust alteration of gut microbial composition related to the disease and therapy.

While an increase in *A. muciniphila* has also been reported in previous studies, (Berer *et al.*, 2017; Cekanaviciute *et al.*, 2017; Cox *et al.*, 2021; Probstel *et al.*, 2020) interpretation of its specific role remains controversial. *A. muciniphila* is a mucin-degrading bacteria shown to exert pro-inflammatory effects on T cells *in vitro* (Cekanaviciute *et al.*, 2017) and to exacerbate inflammation during infection. (Ganesh *et al.*, 2013) Interestingly, peptides from *A. muciniphila* have been recently shown to stimulate human myelin autoreactive CD4+ T cell clones, thus suggesting molecular mimicry is a potential mechanism for MS pathogenesis (Wang *et al.*, 2020). However, *A. muciniphila* has also been proposed as a contributor to maintaining gut health, improving glucose homeostasis, increasing gut mucin integrity and enhancing effect of checkpoint inhibitor immunotherapy. (Cani and de Vos, 2017; Liu *et al.*, 2019; Routy *et al.*, 2018) Different functional capabilities across *A. muciniphila* strains that may affect how these bacteria interact with the host. (Becken *et al.*, 2021; Karcher *et al.*, 2021; Kirmiz *et al.*, 2020) At least two *A. muciniphila* strains were identified in our samples with differences in their functions such as sulfur metabolism, but none of them was enriched in MS in our dataset. Through pathway analysis we found “phytate degradation I” (PWY-4702) (a cyclitols degradation pathway), mainly driven by *A. muciniphila*, was significantly increased in untreated MS patients. This pathway converts phytate into *Myo*-inositol. Phytate is a strong chelator of divalent minerals such as calcium, magnesium, iron and zinc. Previous studies suggested that high levels of iron and zinc could play a role in MS activity and progression. (Ferreira *et al.*, 2017; Hametner *et al.*, 2013; Sanna *et al.*, 2018) whereas calcium and magnesium could exert neuroprotective capacities. (Enders *et al.*, 2020; Goldberg *et al.*, 1986) Dietary mineral intake was no different between MS and healthy controls, but it is still possible that bacterial pathways (such as Phytate degradation) modulate the bioavailability of these minerals, thus contributing to disease pathogenesis. *Myo*-inositol, a simple carbohydrate produced in the body and available in foods such as fruits and cereals, is involved in lipid signaling, osmolarity, glucose, and insulin metabolism (Gonzalez-Uarquin *et al.*, 2020) and utilized as dietary supplementation in different pathological conditions, including diabetes and metabolic disorders. (Pintaudi *et al.*, 2016; Shokrpour *et al.*, 2019) Interestingly, a very early study showed that patients with MS appeared to metabolize *myo*-inositol abnormally, (Holm, 1978) and administered *myo*-inositol was shown to have a positive effect on evoked potential responses in MS (n=9) and controls (n=9). (Young *et al.*, 1986) The role of *Akkermansia* in *myo*-inositol metabolism needs to be further elucidated.

Ruminococcus torques is another potent mucus degrader and may decrease gut barrier integrity. (Cani, 2014; Rajilic-Stojanovic and de Vos, 2014) A recent study showed that *R. torques* was associated with an enhanced MRI T2 signal in multiple motor brain areas and

exacerbated disease in an animal model of amyotrophic lateral sclerosis (ALS). (Blacher et al., 2019) *Ruthenibacterium lactatiformans*, a lactate-producing species, was previously associated with an increased EDSS and decreased lower extremity motor function in RRMS and progressive MS. (Cox et al., 2021)

Overall, 7 species were significantly reduced in untreated MS. *Faecalibacterium prausnitzii*, one of the main butyrate producers found in the intestine, has anti-inflammatory properties that were partly associated with secreted metabolites that block NF- κ B activation, IL-8 production and upregulate regulatory T cell production. (Lopez-Siles et al., 2017) It can also attenuate the severity of inflammation through release of metabolites that enhance intestinal barrier function. (Carlsson et al., 2013; Martin et al., 2015) The pyruvate-producing carboxylates metabolism pathways, contributed by *F. prausnitzii*, were found to be significantly reduced in untreated MS patients. Altogether, we found a depletion of potentially beneficial bacteria in untreated MS patients compared to healthy controls, which in turn disturbed key metabolic pathways that might be expected to worsen the inflammation of MS. These findings could lead to the development of “designer probiotics” that can restore the healthy composition and function of the gut microbiome.

We next tested whether these altered bacteria also associated with disease severity, and found that only *Streptococcus thermophilus*, *Azospirillum sp. 47_25* and *Rhodospirillum sp. UNK.MSG-17* were. However, correlations were positive for RRMS and negative for PMS patients. This implies that the change of gut microbial community may be linked to the onset of disease and stabilized during the disease course, a hypothesis which requires further investigation by longitudinal studies. Several other species were found to be correlated exclusively with MS severity (e.g. not with disease status). For example, *Butyrivibrio*, *Clostridium* and *Ruminococcus* species, which are short-chain fatty acid (SCFA) producers, correlated with lower MS severity in PMS. It's well known that SCFAs play a critical role in immunoregulation with well-characterized anti-inflammatory effects on both epithelium and peripheral immune cells. This implies potentially beneficial effects of these bacteria by producing anti-inflammatory metabolites. On the other hand, *Collinsella aerofaciens*, a species showed to increase disease severity in collagen-induced arthritis mice, (Chen et al., 2016b) was associated with a higher MSSS in RRMS patients probably via the pathway “PWY-4981: L-proline biosynthesis II (from arginine)”. *Prevotella* species such as *P. buccalis*, *P. corporis*, *P. disiens*, *P. copri* were associated with higher MSSS. Although *Prevotella* species have been associated with health-beneficial properties, several studies have shown associations with autoimmune diseases, insulin resistance and diabetes, and gut inflammation. (Leite et al., 2017; Pedersen et al., 2016; Scher et al., 2013) Intriguingly, we found pathway “PWY-5097: L-lysine biosynthesis VI”, a decreased pathway in MS versus HHCs, was associated with a lower disease severity. Several commensal bacteria participants in this pathway, including *Bacteroides*, *Faecalibacterium* and *Eubacterium* species. L-lysine has been shown to have anti-inflammatory in rat with chronic lung injury, (Zhang et al., 2019) and may play a neuroprotective role in intracerebral hemorrhage injury, (Cheng et al., 2020) thus suggesting a potential usage of L-lysine to suppress the disease progression. Based on these observations, we speculate that the role of gut bacteria in disease progression/severity is multi-faceted and individual-dependent.

The use of DMTs resulted in a decrease in the relative abundance of specific taxa that are not MS-associated, potentially by their innate antimicrobial properties.(Maier *et al.*, 2018; Storm-Larsen et al., 2019) Specifically, *Bacteroides*, *Blautia* and *Clostridium* species were significantly reduced in response to oral medications and species like *Faecalibacterium prausnitzii*, *Dialister invisus CAG:218* and *Roseburia intestinalis* were reduced in individual receiving injectables. Furthermore, infused therapies resulted in a decrease of *Bifidobacterium adolescentis*, which was shown to promote Th17 cell accumulation and exacerbated autoimmune arthritis in a mouse model, arguing for its pathological relevance.(Tan et al., 2016) On the other hand, we found several species that were increased by DMTs, in particular *Ruthenibacterium lactatiformans* and *Ruminococcus torques* (with Fingolimod), *Eubacterium hallii* (with GA) and *Bacteroides uniformis* (with interferon). Intriguingly, sequence-based analysis suggested the oral drug Fingolimod would induce the most metabolic changes compared to other medications, a finding validated by metabolomic analysis. Specifically, the depletion of microbial “CARBOXYLATES-DEG” pathways (which produces pyruvate) may explain the low level of pyruvate observed in feces and serum samples from RRMS patients treated with Fingolimod, and the depletion of *F. prausnitzii* (a major SCFA producing bacteria) could account for the lower levels of acetate and propionate found in MS. We also found several microbe-derived fecal metabolites were remarkably lower in treated RRMS patients, implying a particularly important effect of these medications, likely through direct interactions with gut microbiota. Of interest, a significant increase of serum propionic acid was found in RRMS patients treated with interferon. Propionate supplementation in MS patients has been associated with an increased Treg/Th17 ratio, leading to long-term clinical improvement.(Duscha *et al.*, 2020) Based on our findings, we propose the increased absorption of microbially-derived propionate via upregulation of the SCFA transporter MCT1 (SLC16A1) as contributing mechanism of action for IFN- β . Our results provide compelling evidence that DMTs have considerable effects on gut microbiota, not only compositionally but functionally, that may highlight therapeutic mechanisms requiring further investigation. Additional larger and longitudinal follow-up studies will help to evaluate these effects more precisely.

A healthier diet associates with higher microbial diversity, but diet may not the only factor at play. Some bacteria remained unaffected by dietary change depending on host phenotype and the preexisting microbiota composition.(Flandroy et al., 2018) In addition, local environment (i.e. air, soil and water) could also influence diversity of the gut microbiota by horizontal transmission of environmental microbes.(Tasnim et al., 2017) Vitamin D deficiency has long been associated with MS, and higher levels of vitamin D were associated with reduced clinical activity in established MS.(Munger and Ascherio, 2011) Unsurprisingly, we observed that MS patients took more vitamin D, but showed no significant influence on gut microbiome composition.

Limitations of the study

Shotgun metagenomics sequencing was limited to ~500,000 reads per sample. While this coverage is adequate to classify bacterial communities with higher resolution than 16S RNA gene sequencing, and to provide some insight into the metabolic potential of the communities, higher sequencing depth will be needed to resolve most strains, clades

and DNA polymorphisms. We cannot exclude power limitations due to stratification by treatment. As a consequence of the paired household design, the majority of the pairs are spouses, thus leading to an uneven sex distribution of MS (69.4% of the MS participants are female, in keeping with the expected demographics for MS (Langer-Gould et al., 2013)). However, our model adjusted for the effect of sex on gut microbiome.

In summary, this is a large, multi-center gut microbiome study in MS patients and HHC. The findings strongly support the presence of specific gut microbiome associations both with MS disease course and progression, and functional changes in response to treatment. The origin and biological relevance of these associations remain to be elucidated. Nevertheless, our study supports the possibility that microbial manipulation and dietary intervention could be used as preventive and therapeutic strategies in MS.

STAR Methods

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Sergio E. Baranzini (Sergio.Baranzini@ucsf.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability—Shotgun and 16S rRNA amplicon sequencing datasets generated from human fecal DNAs are available in the EMBL-ENA repository with accession number ERP115476. See Table S1 for a complete list of sequenced samples. Processed data for 16S rRNA, metagenomics and metabolomics profiles, clinical data and diet data are available at Dryad (<https://doi.org/10.7272/Q60C4T26>). All original code is available at https://github.com/BaranziniLab/iMSMS_study. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Recruitment and inclusion criteria—A total of 576 MS patients and their household healthy controls were included in this study. See Table S1 for phenotypes of all participants. The first 128 MS-control pairs were recruited as Cohort 1 (Zhou *et al.*, 2020) and the subsequent 448 pairs were recruited as the Cohort 2. Participants were recruited through MS clinics at UCSF (San Francisco, CA), Brigham and Women's Hospital (Boston, MA), Mount Sinai (New York, NY), the Anne Rowling Clinic (Edinburgh, UK), University of Pittsburgh (Pittsburgh, PA), Biodonostia Health Research Institute (San Sebastián, Spain) and FLENI (Buenos Aires, Argentina).

Inclusion criteria required that participants carry a diagnosis of MS; (McDonald et al., 2001) be of White (Hispanic or non-Hispanic) ethnicity (i.e. to match characteristic genetic risk profile of MS (Baranzini and Oksenberg, 2017)); and be enrolled with a genetically unrelated household control with cohabitation for at least six months. Exclusion criteria for MS and control subjects included the presence of other autoimmune disorders, gastrointestinal

infections, and other neurological disorders. Participants were excluded if they received oral antibiotics within the past three months, corticosteroids within the past 30 days, or were on a disease modifying therapy for less than three months.

Ethics approval and consent to participate—Each collaborating site obtained human subject research approval through their respective ethics review committees, following a master protocol established at UCSF (protocol no. 15–17061). All participants provided written informed consent and signed a HIPAA Authorization allowing for the use of their medical record for research purposes.

METHOD DETAILS

Specimen collection—Participants were provided with a stool sample collection kit and instructed to obtain two consecutive stool samples in the privacy of their own homes. Each stool sample time point included 3 collection vials - a Q-tip (Q, dry), a snap frozen vial (S, wet), and a vial filled with Luria-Bertani broth (LB) and 30% glycerol. Participants were instructed to freeze the samples for at least 12 hours and ship them frozen with the ice pack included in the kit. Samples were returned to each site via overnight shipping in a thermal envelope. Blood samples were collected at the initial visit only and stored at -80°C upon further processing. All participants were required to complete a clinical survey to report the disease status and treatment, and a subject survey to report demographic, medication, lifestyle and physiology factors. Clinical outcomes included the Expanded Disability Status Scale (EDSS), (Kurtzke, 1983) and the Multiple Sclerosis Functional Composite (MSFC). (Fischer et al., 1999) All data were collected and stored through secure REDCap™ questionnaires.

Stool sample preparation and 16S rRNA sequencing—For the first cohort, Q-tip samples (i.e. dry) and snap frozen (i.e. wet) samples were processed using the QIAamp PowerFecal DNA Kit (ref 12830–50). After lysis solution was added to bead beating tubes, dry samples were transferred by grinding the Q-tips into the bottom while snap frozen samples were chipped to an appropriate size for the kit. Sample processing was done on a QIAcube platform according to the protocols generated by the manufacturer (QIAGEN). DNA sample quantity and purity were measured by NanoDrop spectrophotometry (Thermo Scientific™). The second cohort samples were processed using the MagAttract PowerSoil DNA EP Kit (ref 27100–4-EP). After lysis solution was added to the bead beating plate, samples were added to each well in in the same manner used previously for bead beating tubes. Physical lysis was executed using a mixer mill and subsequent steps were automated using the EpMotion platform. Sample quality and quantity were assessed with the same method used for the first cohort. To test whether the DNA processing method changes microbial composition, we extracted DNAs from the same 20 samples using both QIAcube and epMotion platforms. A subset of 40 samples prepared in Cohort 1 were re-sequenced in Cohort 2 to test the impact of sequencing runs on microbial composition. As the impact of sample collection method on microbial composition is negligible, (Zhou *et al.*, 2020) sequencing counts of samples from each participant were summed. ASVs were filtered to retain only the ones covering at least 10 total reads and present in at least 5% of samples for downstream analyses (Table S4).

Microbial diversity—Both weighted and unweighted UniFrac(Lozupone and Knight, 2005) distances were computed between all samples (Table S4), and principal coordinates analysis (PCoA) was applied to visualize the β -diversity. All these analyses were performed with QIIME2. Bray-Curtis(Bray JR, 1957) dissimilarities were calculated to compare gut microbiome among individuals in terms of geographic distance. Since the MS and control subjects within household are often of different sex, the random comparisons between households utilized only cross-sexual comparisons to control for the effect of gender. Statistical significance was determined by ANOVA. The PERMANOVA test (McArdle and Anderson, 2001) was used to assess the effect of host metadata categories (confounders): demography, lifestyle, diseases, medication and physiology, on the variation of microbiome abundance (Table S4). The test was performed by using the “adonis” function implemented in R package *vegan*(Zapala and Schork, 2006) and tested on weighted UniFrac distances of paired MS and HHC samples with reported host factors. The variance of microbial abundance between MS and control or between treated/untreated MS and controls were tested by specifying “strata” as household to control the within house comparison. The empirical p-value was obtained by running 999 permutations. When appropriate, statistical p-values were adjusted by false discovery rate (FDR).

Shallow whole metagenome shotgun sequencing (WMGS) and data

processing—For samples with less than 1 ng DNA, a maximum volume of 3.5 μ l input was used. Library concentration was determined with triplicate readings of the Kapa Illumina Library Quantification Kit (cohort 1, ref 07962428001) or Pico Green Quantification Kit (cohort2, ref P11496); 20 fmol of sample libraries were pooled and size selected for fragments between 300 and 800 bp on the Sage Science PippinHT to exclude primer dimers.

Demultiplexed shallow shotgun metagenomic sequences were processed using Atropos (v1.1.24) (Didion et al., 2017) to remove adapters (forward “GATCGGAAGAGCACACGTCTGAACTCCAGTCAC”, reverse “GATCGGAAGAGCGTCGTGTAGGGAAAGGAGTGT”) and filter reads with lower quality score than 15 and length less than 80 base pairs. For taxonomic assignment reads were aligned to the Web of Life(Zhu et al., 2019) of 10,575 bacterial and archaeal genomes using SHOGUN (Hillmann *et al.*, 2018) in the Bowtie2 alignment mode. Species-level functional profiling was performed using HUMAnN2 default. (Franzosa *et al.*, 2018) Sequencing counts of samples from each participant were summed (Table S5). To deal with sparse microbial data in the analysis, we focused on species present in at least 5% of samples, covering at least 10 total reads. This provided a list of 1,677 species for use in the statistical analysis. The relative abundances of gene families were characterized from UniProt Reference Clusters (UniRef90) using HUMAnN2 (V2.8.2),(Franzosa *et al.*, 2018) which were further mapped to microbial pathways and high-classes based on pathway hierarchy from the MetaCyc metabolic pathway database.(Caspi *et al.*, 2016; Caspi et al., 2018) 490 pathways present in at least 5% of samples were retained for statistical analysis. Microbial gene families present in more than 5% samples were used to link with select fecal metabolites. The phylogenetic diversity of *Akkermansia muciniphila* was measured using Metagenomic Intra-species Diversity Analysis System (MIDAS)(Nayfach *et al.*, 2016) with its default parameters.

Microbial co-abundance network—Significant co-abundance was controlled at FDR 0.05 level using $100 \times$ permutation (Table S6). In each permutation, the abundance of each microbial factor was randomly shuffled across samples. To keep the co-abundances with high correlations in a dense microbial network, we filtered co-abundances with a lower absolute correlation than 0.4 and subnetworks with only two species.

To test whether the microbial co-abundance relationships showed case or control specificity, i.e. whether the effect size of co-abundance in MS group was very different from that in healthy control, we applied the IQR (interquartile ranges) based the outlier detection method as adapted in paper.(Chen et al., 2020) The effect size for co-abundance was measured by the SparCC correlation coefficient in our analysis. The effect sizes were ranked from low to high and extracted corresponding 25%, 50% and 75% quartile values (Q1, Q2 and Q3, respectively). IQR was then calculated as $Q3 - Q1$. The specific co-abundance was defined in each corresponding MS or healthy group if the effect size fell outside of $Q1 - 2 \times IQR$ (smallest) or $Q3 + 2 \times IQR$ (largest).

Metabolite measurement in stool and serum samples—Blood samples were centrifuged at 2200g for 20 minutes. The serum layers were aspirated and moved into 2mL cryovials. The serum samples were stored at -80°C before metabolomics profiling. Fecal (150g/sample) and serum (150ul/sample) samples were shipped on dry ice to Metabolon Inc. (Durham, North Carolina) and maintained at -80°C until processed following their published protocols.(Evans et al., 2009; Long et al., 2017; McMurdie et al., 2022)

Global metabolomic profiling—Samples were prepared using the automated MicroLab STAR[®] system from Hamilton Company. Several recovery standards were added prior to the first step in the extraction process for QC purposes. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five fractions: two for analysis by two separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI, and one sample was reserved for backup. Samples were placed briefly on a TurboVap[®] (Zymark) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before preparation for analysis.

Several types of controls were analyzed in concert with the experimental samples: a pooled matrix sample generated by taking a small volume of each experimental sample (or alternatively, use of a pool of well-characterized human plasma) served as a technical replicate throughout the data set; extracted water samples served as process blanks; and a cocktail of QC standards that were carefully chosen not to interfere with the measurement of endogenous compounds were spiked into every analyzed sample, allowed instrument performance monitoring and aided chromatographic alignment. Instrument variability was determined by calculating the median relative standard deviation (RSD) for the standards that were added to each sample prior to injection into the mass spectrometers. Overall process variability was determined by calculating the median RSD for all endogenous

metabolites (i.e., non-instrument standards) present in 100% of the pooled matrix samples. Experimental samples were randomized across the platform run with QC samples spaced evenly among the injections.

All methods utilized a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. The sample extract was dried then reconstituted in solvents compatible to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18–2.1×100 mm, 1.7 μm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analyzed using acidic positive ion conditions, however it was chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same afore mentioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, however with 6.5mM Ammonium Bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1×150 mm, 1.7 μm) using a gradient consisting of water and acetonitrile with 10mM Ammonium Formate, pH 10.8. The MS analysis alternated between MS and data-dependent MS_n scans using dynamic exclusion. The scan range varied slightly between methods but covered 70–1000 m/z. Raw data files are archived and extracted as described below.

Raw data was extracted, peak-identified and QC processed using Metabolon's hardware and software. These systems are built on a web-service platform utilizing Microsoft's .NET technologies, which run on high-performance application servers and fiber-channel storage arrays in clusters to provide active failover and load-balancing. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolon maintains a library based on authenticated standards that contains the retention time/index (RI), mass to charge ratio (m/z), and chromatographic data (including MS/MS spectral data) on all molecules present in the library. Furthermore, biochemical identifications are based on three criteria: retention index within a narrow RI window of the proposed identification, accurate mass match to the library +/- 10 ppm, and the MS/MS forward and reverse scores between the experimental data and authentic standards. The MS/MS scores are based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. While there may be similarities between these molecules based on one of these factors, the use of all three data points can be utilized to distinguish and differentiate biochemicals. More than 3300 commercially available purified standard compounds have been acquired and registered into LIMS for analysis on all platforms for determination of their analytical characteristics. Additional mass spectral entries have been created for structurally unnamed biochemicals, which have

been identified by virtue of their recurrent nature (both chromatographic and mass spectral). These compounds have the potential to be identified by future acquisition of a matching purified standard or by classical structural analysis.

A variety of curation procedures were carried out to ensure that a high-quality data set was made available for statistical analysis and data interpretation. The QC and curation processes were designed to ensure accurate and consistent identification of true chemical entities, and to remove those representing system artifacts, mis-assignments, and background noise. Metabolon data analysts use proprietary visualization and interpretation software to confirm the consistency of peak identification among the various samples. Library matches for each compound were checked for each sample and corrected if necessary.

Targeted short-chain fatty acid profiling—Human feces and human serum samples are analyzed for eight short chain fatty acids: acetic acid (C2), propionic acid (C3), isobutyric acid (C4), butyric acid (C4), 2-methyl- butyric acid (C5), isovaleric acid (C5), valeric acid (C5), and caproic acid (hexanoic acid, C6), with the addition of lactic acid by request, by LC-MS/MS (Metabolon Method TAM135: “LC-MS/MS Method for the Quantitation of Short Chain Fatty Acid (C2 to C6) in Human Feces” and TAM148: “LC-MS/MS Method for the Quantitation of Short Chain Fatty Acid (C2 to C6) in Human Plasma and Serum”). Human feces and human serum samples are spiked with stable labelled internal standards and are homogenized and subjected to protein precipitation with an organic solvent. After centrifugation, an aliquot of the supernatant is derivatized. The reaction mixture is diluted, and an aliquot is injected onto an Agilent 1290/AB Sciex QTrap 5500 LC MS/MS system equipped with a C18 reversed phase UHPLC column. The mass spectrometer is operated in negative mode using electrospray ionization (ESI). The peak area of the individual analyte product ions is measured against the peak area of the product ions of the corresponding internal standards. Quantitation is performed using a weighted linear least squares regression analysis generated from fortified calibration standards prepared immediately prior to each run. LC-MS/MS raw data are collected using AB SCIEX software Analyst 1.6.2 and processed with SCIEX OS-MQ software v1.7.

Differential microbiome features by mixed linear regression analysis—Global metabolite intensity and SCFA concentration were normalized by log transformation. Mixed linear regression model was applied on transformed data to identify differential features (species, pathways and metabolites) by adjusting random effects of house and recruitment site, and fixed effects of age, sex and BMI. The linear regression was performed using *lmer* function from R package “*lme4*” as $\text{lmer}(y \sim \text{disease} + \text{age} + \text{BMI} + \text{sex} + (1|\text{site}) + (1|\text{house}))$. To reduce the effect of zero-inflation in microbiome data, a variance filtering step was applied to remove species features with very low variance ($<1E-5$). The contribution of individual species in a specific pathway was visualized in a bar plot using HUMANn2 “*humann2_barplot*” function. Altered metabolites were linked to gut microbes through reactions (MetaCyc and KEGG) mediated by microbial gene families screened in our WGMS data using HUMANn2. Functional KEGG enrichment analysis of metabolites was performed using MetaboAnalyst 5.0.(Pang *et al.*, 2021)

To identify species associated with disease severity, the updated global Multiple Sclerosis Severity Score (uGMSSS) was calculated by combining the Expanded Disability Status Scale (EDSS) and disease duration using *global_msss* function from R package “*ms.sev*”. We focused on the species with prevalence in more than 50% samples, spearman correlations were calculated and tested adjusting for age and BMI using *pcor.test* function from R package “*ppcor*”.

Diet analysis—A validated Block 2005 food frequency questionnaire (FFQ)(Block, 2005) was set up through an external vendor (NutritionQuest). The intake of foods and nutrients were measured by NutritionQuest in a standardized fashion for all participants based on their responses to the FFQ. 37 nutrient items were summarized and grouped as antioxidants, average intake, B-vitamins, food group servings and minerals (Table S2). Dietary dissimilarity was measured using Jaccard distance of the nutrient intake. The effect of confounders on the variation of diet and the effect of dietary items (covariates) on the variation of gut microbiome were accessed by PERMANOVA (Permutational multivariate analysis of variance).(McArdle and Anderson, 2001) The test was performed by using the “adonis” function implemented in R package *vegan*.(Zapala and Schork, 2006) The empirical p-value was obtained by running 999 permutations. Healthy Eating Index-2015 (HEI-2015(Krebs-Smith et al., 2018)) was used for evaluation of the diet quality and calculated by NutritionQuest (Table S3). The HEI-2015 adequate dietary components include ‘total fruit’, ‘whole fruit’, ‘total vegetables’, ‘greens and beans’, ‘whole grains’, ‘dairy’, ‘total protein’, ‘seafood and plant proteins’, and ‘fatty acids’, which are recommended to be high in a healthy diet. In contrast, moderate dietary components where consumption is recommended to be limited include ‘refined grains’, ‘sodium’, ‘added sugar’ and ‘saturated fatty acids’.(Krebs-Smith *et al.*, 2018) Each component was measured by a maximum point scale. To make all components comparable with maximum point of 10, the points of ‘total fruit’ and ‘whole fruit’ were added as ‘fruit’, ‘total vegetables’ and ‘greens and beans’ were added as ‘vegetables’, ‘total protein’ and ‘seafood and plant proteins’ were added as ‘protein’. Correlation between HEI-2015 and host phenotypes (age and BMI), microbial diversity or microbial relative abundance was measured by Pearson’s correlation. Correlations between each dietary component and MS associated species were measured by coefficients from mixed linear regression model adjusted for age, BMI, sex and recruiting site. Difference of healthy eating index and dietary component points between HHC and MS were tested using paired T-test.

QUANTIFICATION AND STATISTICAL ANALYSIS

Details on statistical tests, n numbers, and significance cutoffs can also be found in the figure legends. Principal coordinates analysis (PCoA) of weighted UniFrac community distance were computed by disease status, treatment status and disease subtype. R^2 and FDR adjusted p values were obtained by PERMANOVA. Effect size (Adonis R^2) of confounders significantly associated with gut microbial variations were shown by weighted UniFrac distance, PERMANOVA (FDR adjusted $p < 0.05$). Community distance of RRMS subjects treated and untreated, and their corresponding household healthy controls were analyzed by PCoA of weighted UniFrac (P values obtained by PERMANOVA)

Mixed linear regression models adjusted for age, BMI, sex, recruiting site and house were used for metagenomics species adjustment for host factors; to define metagenomic pathways altered in untreated MS, untreated RRMS or untreated PMS versus their HHCs; to identify dominant microbial species contributing to “PWY-4702” and “GALACT-GLUCUROCAT-PWY” pathways; to identify metagenomics species and metabolic pathways altered in untreated and treated RRMS; to identify pathways altered in untreated and treated RRMS, and to identify the linear coefficient for 31 microbe-derived metabolites and 8 short chain fatty acids analyzed in untreated and treated RRMS in both stool and serum. Spearman correlations (adjusted for age and BMI) were computed between species and pathways with MS severity scores in untreated RRMS patients (n=112) or untreated progressive MS (n=97). Cohort specific analysis (quantile range outlier) was used to characterize disease status specific co-abundance species in untreated MS and HHCs.

A mixed linear regression model adjusted for age, BMI, sex and recruiting site was also used for identifying correlations between dietary component and MS-associated species. Pearson’s correlation was computed between HEI and microbial α -diversity in healthy and MS individuals, as well as with species significantly correlated with HEI. Jaccard dissimilarity was used to show effect size (Adonis R^2) of confounders associated with dietary variations. Differences in HEI and dietary intake between MS patients and their household healthy controls were evaluated by paired T-test.

Disease duration-adjusted MS severity score (gMSSS) was compared between untreated and treated RRMS by ANOVA. Paired T-test were used to show Arcsine square-root transformed relative abundance of *Akkermansia muciniphila* and *Faecalibacterium prausnitzii* that participate in phytate degradation I pathway (PWY-4702) and superpathway of hexuronide and hexuronate degradation pathway, respectively.

When relevant, further details are found in the method details for the specific measurement in the context of describing sample collections. Analytical methods that were described in the Method details are provided again here.

Microbial diversity—Both weighted and unweighted UniFrac(Lozupone and Knight, 2005) distances were computed between all samples (Table S4), and principal coordinates analysis (PCoA) was applied to visualize the β -diversity. All these analyses were performed with QIIME2. Bray-Curtis(Bray JR, 1957) dissimilarities were calculated to compare gut microbiome among individuals in terms of geographic distance. Since the MS and control subjects within household are often of different sex, the random comparisons between households utilized only cross-sexual comparisons to control for the effect of gender. Statistical significance was determined by ANOVA. The PERMANOVA test (McArdle and Anderson, 2001) was used to assess the effect of host metadata categories (confounders): demography, lifestyle, diseases, medication and physiology, on the variation of microbiome abundance (Table S4). The test was performed by using the “adonis” function implemented in R package *vegan*(Zapala and Schork, 2006) and tested on weighted UniFrac distances of paired MS and HHC samples with reported host factors. The variance of microbial abundance between MS and control or between treated/untreated MS and controls were tested by specifying “strata” as household to control the within house comparison. The

empirical p-value was obtained by running 999 permutations. When appropriate, statistical p-values were adjusted by false discovery rate (FDR).

Shallow whole metagenome shotgun sequencing (WMGS) data

processing—Demultiplexed shallow shotgun metagenomic sequences were processed using Atropos (v1.1.24) (Didion et al., 2017) to remove adapters (forward “GATCGGAAGAGCACACGTCTGAACTCCAGTCAC”, reverse “GATCGGAAGAGCGTCGTGTAGGGAAAGGAGTGT”) and filter reads with lower quality score than 15 and length less than 80 base pairs. For taxonomic assignment reads were aligned to the Web of Life (Zhu et al., 2019) of 10,575 bacterial and archaeal genomes using SHOGUN (Hillmann *et al.*, 2018) in the Bowtie2 alignment mode. Species-level functional profiling was performed using HUMAnN2 default. (Franzosa *et al.*, 2018) Sequencing counts of samples from each participant were summed (Table S5). To deal with sparse microbial data in the analysis, we focused on species present in at least 5% of samples, covering at least 10 total reads. This provided a list of 1,677 species for use in the statistical analysis. The relative abundances of gene families were characterized from UniProt Reference Clusters (UniRef90) using HUMAnN2 (V2.8.2), (Franzosa *et al.*, 2018) which were further mapped to microbial pathways and high-classes based on pathway hierarchy from the MetaCyc metabolic pathway database. (Caspi *et al.*, 2016; Caspi et al., 2018) 490 pathways present in at least 5% of samples were retained for statistical analysis. Microbial gene families present in more than 5% samples were used to link with select fecal metabolites. The phylogenetic diversity of *Akkermansia muciniphila* was measured using Metagenomic Intra-species Diversity Analysis System (MIDAS) (Nayfach *et al.*, 2016) with its default parameters.

Microbial co-abundance network—Significant co-abundance was controlled at FDR 0.05 level using $100 \times$ permutation (Table S6). In each permutation, the abundance of each microbial factor was randomly shuffled across samples. To keep the co-abundances with high correlations in a dense microbial network, we filtered co-abundances with a lower absolute correlation than 0.4 and subnetworks with only two species.

To test whether the microbial co-abundance relationships showed case or control specificity, i.e. whether the effect size of co-abundance in MS group was very different from that in healthy control, we applied the IQR (interquartile ranges) based the outlier detection method as adapted in paper. (Chen et al., 2020) The effect size for co-abundance was measured by the SparCC correlation coefficient in our analysis. The effect sizes were ranked from low to high and extracted corresponding 25%, 50% and 75% quartile values (Q1, Q2 and Q3, respectively). IQR was then calculated as $Q3 - Q1$. The specific co-abundance was defined in each corresponding MS or healthy group if the effect size fell outside of $Q1 - 2 \times IQR$ (smallest) or $Q3 + 2 \times IQR$ (largest).

Global metabolomic profiling analysis—Several types of controls were analyzed in concert with the experimental samples: a pooled matrix sample generated by taking a small volume of each experimental sample (or alternatively, use of a pool of well-characterized human plasma) served as a technical replicate throughout the data set; extracted water samples served as process blanks; and a cocktail of QC standards that were carefully chosen

not to interfere with the measurement of endogenous compounds were spiked into every analyzed sample, allowed instrument performance monitoring and aided chromatographic alignment. Instrument variability was determined by calculating the median relative standard deviation (RSD) for the standards that were added to each sample prior to injection into the mass spectrometers. Overall process variability was determined by calculating the median RSD for all endogenous metabolites (i.e., non-instrument standards) present in 100% of the pooled matrix samples. Experimental samples were randomized across the platform run with QC samples spaced evenly among the injections.

All methods utilized a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. The sample extract was dried then reconstituted in solvents compatible to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18–2.1×100 mm, 1.7 μm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analyzed using acidic positive ion conditions, however it was chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same aforementioned C18 column using methanol, acetonitrile, water, 0.05% PFPA and 0.01% FA and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, however with 6.5mM Ammonium Bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1×150 mm, 1.7 μm) using a gradient consisting of water and acetonitrile with 10mM Ammonium Formate, pH 10.8. The MS analysis alternated between MS and data-dependent MS_n scans using dynamic exclusion. The scan range varied slightly between methods but covered 70–1000 m/z. Raw data files are archived and extracted as described below.

Raw data was extracted, peak-identified and QC processed using Metabolon's hardware and software. These systems are built on a web-service platform utilizing Microsoft's .NET technologies, which run on high-performance application servers and fiber-channel storage arrays in clusters to provide active failover and load-balancing. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. Metabolon maintains a library based on authenticated standards that contains the retention time/index (RI), mass to charge ratio (m/z), and chromatographic data (including MS/MS spectral data) on all molecules present in the library. Furthermore, biochemical identifications are based on three criteria: retention index within a narrow RI window of the proposed identification, accurate mass match to the library +/- 10 ppm, and the MS/MS forward and reverse scores between the experimental data and authentic standards. The MS/MS scores are based on a comparison of the ions present in the experimental spectrum to the ions present in the library spectrum. While there may be similarities

between these molecules based on one of these factors, the use of all three data points can be utilized to distinguish and differentiate biochemicals. More than 3300 commercially available purified standard compounds have been acquired and registered into LIMS for analysis on all platforms for determination of their analytical characteristics. Additional mass spectral entries have been created for structurally unnamed biochemicals, which have been identified by virtue of their recurrent nature (both chromatographic and mass spectral). These compounds have the potential to be identified by future acquisition of a matching purified standard or by classical structural analysis.

A variety of curation procedures were carried out to ensure that a high quality data set was made available for statistical analysis and data interpretation. The QC and curation processes were designed to ensure accurate and consistent identification of true chemical entities, and to remove those representing system artifacts, mis-assignments, and background noise. Metabolon data analysts use proprietary visualization and interpretation software to confirm the consistency of peak identification among the various samples. Library matches for each compound were checked for each sample and corrected if necessary.

Targeted short-chain fatty acid profiling analysis—The peak area of the individual analyte product ions is measured against the peak area of the product ions of the corresponding internal standards. Quantitation is performed using a weighted linear least squares regression analysis generated from fortified calibration standards prepared immediately prior to each run. LC-MS/MS raw data are collected using AB SCIEX software Analyst 1.6.2 and processed with SCIEX OS-MQ software v1.7.

Differential microbiome features by mixed linear regression analysis—Global metabolite intensity and SCFA concentration were normalized by log transformation. Mixed linear regression model was applied on transformed data to identify differential features (species, pathways and metabolites) by adjusting random effects of house and recruitment site, and fixed effects of age, sex and BMI. The linear regression was performed using *lmer* function from R package “*lme4*” as $\text{lmer}(y \sim \text{disease} + \text{age} + \text{BMI} + \text{sex} + (1|\text{site}) + (1|\text{house}))$. To reduce the effect of zero-inflation in microbiome data, a variance filtering step was applied to remove species features with very low variance ($<1\text{E-}5$). The contribution of individual species in a specific pathway was visualized in a bar plot using HUMAnN2 “*humann2_barplot*” function. Altered metabolites were linked to gut microbes through reactions (MetaCyc and KEGG) mediated by microbial gene families screened in our WGMS data using HUMAnN2. Functional KEGG enrichment analysis of metabolites was performed using MetaboAnalyst 5.0.(Pang *et al.*, 2021)

To identify species associated with disease severity, the updated global Multiple Sclerosis Severity Score (uGMSSS) was calculated by combining the Expanded Disability Status Scale (EDSS) and disease duration using *global_msss* function from R package “*ms.sev*”. We focused on the species with prevalence in more than 50% samples, spearman correlations were calculated and tested adjusting for age and BMI using *pcor.test* function from R package “*ppcor*”.

Diet analysis—A validated Block 2005 food frequency questionnaire (FFQ)(Block, 2005) was set up through an external vendor (NutritionQuest). The intake of foods and nutrients were measured by NutritionQuest in a standardized fashion for all participants based on their responses to the FFQ. 37 nutrient items were summarized and grouped as antioxidants, average intake, B-vitamins, food group servings and minerals (Table S2). Dietary dissimilarity was measured using Jaccard distance of the nutrient intake. The effect of confounders on the variation of diet and the effect of dietary items (covariates) on the variation of gut microbiome were accessed by PERMANOVA (Permutational multivariate analysis of variance).(McArdle and Anderson, 2001) The test was performed by using the “adonis” function implemented in R package *vegan*.(Zapala and Schork, 2006) The empirical p-value was obtained by running 999 permutations. Healthy Eating Index-2015 (HEI-2015(Krebs-Smith et al., 2018)) was used for evaluation of the diet quality and calculated by NutritionQuest (Table S3). The HEI-2015 adequate dietary components include ‘total fruit’, ‘whole fruit’, ‘total vegetables’, ‘greens and beans’, ‘whole grains’, ‘dairy’, ‘total protein’, ‘seafood and plant proteins’, and ‘fatty acids’, which are recommended to be high in a healthy diet. In contrast, moderate dietary components where consumption is recommended to be limited include ‘refined grains’, ‘sodium’, ‘added sugar’ and ‘saturated fatty acids’.(Krebs-Smith *et al.*, 2018) Each component was measured by a maximum point scale. To make all components comparable with maximum point of 10, the points of ‘total fruit’ and ‘whole fruit’ were added as ‘fruit’, ‘total vegetables’ and ‘greens and beans’ were added as ‘vegetables’, ‘total protein’ and ‘seafood and plant proteins’ were added as ‘protein’. Correlation between HEI-2015 and host phenotypes (age and BMI), microbial diversity or microbial relative abundance was measured by Pearson’s correlation. Correlations between each dietary component and MS associated species were measured by coefficients from mixed linear regression model adjusted for age, BMI, sex and recruiting site. Difference of healthy eating index and dietary component points between HHC and MS were tested using paired T-test.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Major funding for this work was provided by the Valhalla Foundation. This work was also supported by the Salah Foundation and the UCSF Benioff Center for Microbiome Medicine. RH and LAG are supported by the National MS Society USA (RFA-2104–37464), Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy – ID 390857198), the Gemeinnützige Hertie Stiftung, Bavarian association and national association of the German MS society (DMSG), Dr. Leopold And Carmen Ellinger Foundation, and the association “Verein zur Therapieforschung für MS Kranke e.V. ZX has received research support from the National Institutes of Health and the Department of Defense. LMC is supported by The Race to Erase MS Early Investigator Award. HLW is supported by NIH/NINDS award R01NS087226. SLH is supported by NIH/NINDS award 5R35NS111644. SEB holds the Heidrich Family and Friends endowed chair in Neurology at UCSF. SEB holds the Professorship in Neurology I at UCSF.

REFERENCES

- Arabshahi S, Lahmann PH, Williams GM, Marks GC, and van der Pols JC (2011). Longitudinal change in diet quality in Australian adults varies by demographic, socio-economic, and lifestyle characteristics. *J Nutr* 141, 1871–1879. 10.3945/jn.111.140822. [PubMed: 21865564]
- Baranzini SE, and Oksenberg JR (2017). The Genetics of Multiple Sclerosis: From 0 to 200 in 50 Years. *Trends in genetics : TIG* 33, 960–970. 10.1016/j.tig.2017.09.004.
- Becken B, Davey L, Middleton DR, Mueller KD, Sharma A, Holmes ZC, Dallow E, Remick B, Barton GM, David LA, et al. (2021). Genotypic and Phenotypic Diversity among Human Isolates of *Akkermansia muciniphila*. *Mbio* 12. ARTN e00478–21 10.1128/mBio.00478-21.
- Berer K, Gerdes LA, Cekanaviciute E, Jia X, Xiao L, Xia Z, Liu C, Klotz L, Stauffer U, Baranzini SE, et al. (2017). Gut microbiota from multiple sclerosis patients enables spontaneous autoimmune encephalomyelitis in mice. *Proc Natl Acad Sci U S A* 114, 10719–10724. 10.1073/pnas.1711233114. [PubMed: 28893994]
- Blacher E, Bashiardes S, Shapiro H, Rothschild D, Mor U, Dori-Bachash M, Kleimeyer C, Moresi C, Harnik Y, Zur M, et al. (2019). Potential roles of gut microbiome and metabolites in modulating ALS in mice. *Nature* 572, 474–480. 10.1038/s41586-019-1443-5. [PubMed: 31330533]
- Block G (2005). Block Adult Questionnaire. Retrieved 2007–04-30 from the NutritionQuest Questionnaires and Screener Page:http://www.nutritionquest.com/products/questionnaires_screeners.htm.
- Bray JR CJ (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr* 27, 325–349. 10.2307/1942268.
- Camara-Lemarroy CR, Metz LM, and Yong VW (2018). Focus on the gut-brain axis: Multiple sclerosis, the intestinal barrier and the microbiome. *World J Gastroenterol* 24, 4217–4223. 10.3748/wjg.v24.i37.4217. [PubMed: 30310254]
- Cani PD (2014). Metabolism in 2013: The gut microbiota manages host metabolism. *Nat Rev Endocrinol* 10, 74–76. 10.1038/nrendo.2013.240. [PubMed: 24322652]
- Cani PD, and de Vos WM (2017). Next-Generation Beneficial Microbes: The Case of *Akkermansia muciniphila*. *Front Microbiol* 8, 1765. 10.3389/fmicb.2017.01765. [PubMed: 29018410]
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7, 335–336. 10.1038/nmeth.f.303. [PubMed: 20383131]
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6, 1621–1624. 10.1038/ismej.2012.8. [PubMed: 22402401]
- Carlsson AH, Yakymenko O, Olivier I, Hakansson F, Postma E, Keita AV, and Soderholm JD (2013). Faecalibacterium prausnitzii supernatant improves intestinal barrier function in mice DSS colitis. *Scand J Gastroenterol* 48, 1136–1144. 10.3109/00365521.2013.828773. [PubMed: 23971882]
- Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, et al. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 44, D471–480. 10.1093/nar/gkv1164. [PubMed: 26527732]
- Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Midford PE, Ong Q, Ong WK, et al. (2018). The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res* 46, D633–D639. 10.1093/nar/gkx935. [PubMed: 29059334]
- Castillo-Alvarez F, Perez-Matute P, Oteo JA, and Marzo-Sola ME (2018). The influence of interferon beta-1b on gut microbiota composition in patients with multiple sclerosis. *Neurologia (Engl Ed)* 10.1016/j.nrl.2018.04.006.
- Cekanaviciute E, Yoo BB, Runia TF, Debelius JW, Singh S, Nelson CA, Kanner R, Bencosme Y, Lee YK, Hauser SL, et al. (2017). Gut bacteria from multiple sclerosis patients modulate human T cells and exacerbate symptoms in mouse models. *Proc Natl Acad Sci U S A* 114, 10713–10718. 10.1073/pnas.1711235114. [PubMed: 28893978]

- Chao A (1984). Non-parametric estimation of the number of classes in a population. *Scand J Stat* 11, 265–270.
- Chen J, Chia N, Kalari KR, Yao JZ, Novotna M, Soldan MMP, Luckey DH, Marietta EV, Jeraldo PR, Chen XF, et al. (2016a). Multiple sclerosis patients have a distinct gut microbiota compared to healthy controls. *Sci Rep-Uk* 6. ARTN 28484 10.1038/srep28484.
- Chen J, Wright K, Davis JM, Jeraldo P, Marietta EV, Murray J, Nelson H, Matteson EL, and Taneja V (2016b). An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome Med* 8. 10.1186/s13073-016-0299-7.
- Chen LM, Collij V, Jaeger M, van den Munckhof ICL, Vila AV, Kurilshikov A, Gacesa R, Sinha T, Oosting M, Joosten LAB, et al. (2020). Gut microbial co-abundance networks show specificity in inflammatory bowel disease and obesity. *Nature Communications* 11. ARTN 4018 10.1038/s41467-020-17840-y.
- Cheng J, Tang JC, Pan MX, Chen SF, Zhao D, Zhang Y, Liao HB, Zhuang Y, Lei RX, Wang S, et al. (2020). L-lysine confers neuroprotection by suppressing inflammatory response via microRNA-575/PTEN signaling after mouse intracerebral hemorrhage injury. *Exp Neurol* 327. ARTN 113214 10.1016/j.expneurol.2020.113214.
- Chung WSF, Walker AW, Louis P, Parkhill J, Vermeiren J, Bosscher D, Duncan SH, and Flint HJ (2016). Modulation of the human gut microbiota by dietary fibres occurs at the species level. *Bmc Biol* 14. ARTN 3 10.1186/s12915-015-0224-3.
- Cox LM, Maghzi AH, Liu S, Tankou SK, Dhang FH, Willocq V, Song A, Wasen C, Tauhid S, Chu R, et al. (2021). Gut Microbiome in Progressive Multiple Sclerosis. *Ann Neurol* 89, 1195–1211. 10.1002/ana.26084. [PubMed: 33876477]
- De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT, Piccio L, Raychaudhuri S, Tran D, Aubin C, et al. (2009). Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat Genet* 41, 776–782. 10.1038/ng.401. [PubMed: 19525953]
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, and Andersen GL (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microb* 72, 5069–5072. 10.1128/Aem.03006-05.
- Didion JP, Martin M, and Collins FS (2017). Atropos: specific, sensitive, and speedy trimming of sequencing reads. *PeerJ* 5, e3720. 10.7717/peerj.3720. [PubMed: 28875074]
- du Teil Espina M, Gabarrini G, Harmsen HJM, Westra J, van Winkelhoff AJ, and van Dijk JM (2019). Talk to your gut: the oral-gut microbiome axis and its immunomodulatory role in the etiology of rheumatoid arthritis. *FEMS Microbiol Rev* 43, 1–18. 10.1093/femsre/fuy035. [PubMed: 30219863]
- Duscha A, Gisevius B, Hirschberg S, Yissachar N, Stangl GI, Eilers E, Bader V, Haase S, Kaisler J, David C, et al. (2020). Propionic Acid Shapes the Multiple Sclerosis Disease Course by an Immunomodulatory Mechanism. *Cell* 180, 1067–1080 e1016. 10.1016/j.cell.2020.02.035. [PubMed: 32160527]
- Enders M, Heider T, Ludwig A, and Kuerten S (2020). Strategies for Neuroprotection in Multiple Sclerosis and the Role of Calcium. *Int J Mol Sci* 21. 10.3390/ijms21051663.
- Esmail Amini M, Shomali N, Bakhshi A, Rezaei S, Hemmatzadeh M, Hosseinzadeh R, Eslami S, Babi F, Aslani S, Torkamandi S, and Mohammadi H (2020). Gut microbiome and multiple sclerosis: New insights and perspective. *Int Immunopharmacol* 88, 107024. 10.1016/j.intimp.2020.107024. [PubMed: 33182024]
- Evans AM, DeHaven CD, Barrett T, Mitchell M, and Milgram E (2009). Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal Chem* 81, 6656–6667. 10.1021/ac901536h. [PubMed: 19624122]
- Ferreira KPZ, Oliveira SR, Kallaur AP, Kaimen-Maciel DR, Lozovoy MAB, de Almeida ERD, Morimoto HK, Mezzaroba L, Dichi I, Reiche EMV, and Simao ANC (2017). Disease progression and oxidative stress are associated with higher serum ferritin levels in patients with multiple sclerosis. *J Neurol Sci* 373, 236–241. 10.1016/j.jns.2016.12.039. [PubMed: 28131195]

- Fischer JS, Rudick RA, Cutter GR, and Reingold SC (1999). The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. *Mult Scler* 5, 244–250. 10.1177/135245859900500409. [PubMed: 10467383]
- Flandroy L, Poutahidis T, Berg G, Clarke G, Dao MC, Decaestecker E, Furman E, Hahtela T, Massart S, Plovier H, et al. (2018). The impact of human activities and lifestyles on the interlinked microbiota and health of humans and of ecosystems. *Sci Total Environ* 627, 1018–1038. 10.1016/j.scitotenv.2018.01.288. [PubMed: 29426121]
- Franzosa EA, McIver LJ, Rahnava G, Thompson LR, Schirmer M, Weingart G, Lipson KS, Knight R, Caporaso JG, Segata N, and Huttenhower C (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* 15, 962–968. 10.1038/s41592-018-0176-y. [PubMed: 30377376]
- Friedman J, and Alm EJ (2012). Inferring correlation networks from genomic survey data. *PLoS Comput Biol* 8, e1002687. 10.1371/journal.pcbi.1002687. [PubMed: 23028285]
- Ganesh BP, Klopffleisch R, Loh G, and Blaut M (2013). Commensal *Akkermansia muciniphila* exacerbates gut inflammation in *Salmonella* Typhimurium-infected gnotobiotic mice. *Plos One* 8, e74963. 10.1371/journal.pone.0074963. [PubMed: 24040367]
- Garcia-Ribera S, Amat-Bou M, Climent E, Llobet M, Chenoll E, Corripio R, Ibanez L, Ramon-Krauel M, and Lerin C (2020). Specific Dietary Components and Gut Microbiota Composition are Associated with Obesity in Children and Adolescents with Prader-Willi Syndrome. *Nutrients* 12. ARTN 1063 10.3390/nu12041063.
- Gaulke CA, and Sharpton TJ (2018). The influence of ethnicity and geography on human gut microbiome composition. *Nat Med* 24, 1495–1496. 10.1038/s41591-018-0210-8. [PubMed: 30275567]
- Geva-Zatorsky N, Sefik E, Kua L, Pasman L, Tan TG, Ortiz-Lopez A, Yanortsang TB, Yang L, Jupp R, Mathis D, et al. (2017). Mining the Human Gut Microbiota for Immunomodulatory Organisms. *Cell* 168, 928–943 e911. 10.1016/j.cell.2017.01.022. [PubMed: 28215708]
- Goldberg P, Fleming MC, and Picard EH (1986). Multiple-Sclerosis - Decreased Relapse Rate through Dietary Supplementation with Calcium, Magnesium and Vitamin-D. *Med Hypotheses* 21, 193–200. Doi 10.1016/0306-9877(86)90010-1. [PubMed: 3537648]
- Gonzalez A, Navas-Molina JA, Kosciolk T, McDonald D, Vazquez-Baeza Y, Ackermann G, DeReus J, Janssen S, Swafford AD, Orchanian SB, et al. (2018). Qiita: rapid, web-enabled microbiome meta-analysis. *Nat Methods* 15, 796–798. 10.1038/s41592-018-0141-9. [PubMed: 30275573]
- Gonzalez-Uarquin F, Rodehutsord M, and Huber K (2020). Myo-inositol: its metabolism and potential implications for poultry nutrition-a review. *Poult Sci* 99, 893–905. 10.1016/j.psj.2019.10.014. [PubMed: 32036985]
- Guo X, Warden BA, Paeratakul S, and Bray GA (2004). Healthy Eating Index and obesity. *Eur J Clin Nutr* 58, 1580–1586. 10.1038/sj.ejcn.1601989. [PubMed: 15162130]
- Hametner S, Wimmer I, Haider L, Pfeifenbring S, Bruck W, and Lassmann H (2013). Iron and neurodegeneration in the multiple sclerosis brain. *Ann Neurol* 74, 848–861. 10.1002/ana.23974. [PubMed: 23868451]
- Hatton GB, Madla CM, Rabbie SC, and Basit AW (2018). All disease begins in the gut: Influence of gastrointestinal disorders and surgery on oral drug performance. *Int J Pharm* 548, 408–422. 10.1016/j.ijpharm.2018.06.054. [PubMed: 29969711]
- Hillmann B, Al-Ghalith GA, Shields-Cutler RR, Zhu Q, Gohl DM, Beckman KB, Knight R, and Knights D (2018). Evaluating the Information Content of Shallow Shotgun Metagenomics. *mSystems* 3. 10.1128/mSystems.00069-18.
- Himmelstein DS, Lizee A, Hessler C, Brueggeman L, Chen SL, Hadley D, Green A, Khankhanian P, and Baranzini SE (2017). Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *Elife* 6. 10.7554/eLife.26726.
- Holm V (1978). Inositol in multiple sclerosis. *Arch Neurol* 35, 478. 10.1001/archneur.1978.00500310080018.
- Honda K, and Littman DR (2016). The microbiota in adaptive immune homeostasis and disease. *Nature* 535, 75–84. 10.1038/nature18848. [PubMed: 27383982]

- Jangi S, Gandhi R, Cox LM, Li N, von Glehn F, Yan R, Patel B, Mazzola MA, Liu S, Glanz BL, et al. (2016). Alterations of the human gut microbiome in multiple sclerosis. *Nat Commun* 7, 12015. 10.1038/ncomms12015. [PubMed: 27352007]
- Kadowaki A, and Quintana FJ (2020). The Gut-CNS Axis in Multiple Sclerosis. *Trends Neurosci* 43, 622–634. 10.1016/j.tins.2020.06.002. [PubMed: 32650957]
- Kanehisa M, and Goto S (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28, 27–30. [PubMed: 10592173]
- Karcher N, Nigro E, Puncochar M, Blanco-Miguez A, Ciciani M, Manghi P, Zolfo M, Cumbo F, Manara S, Golzato D, et al. (2021). Genomic diversity and ecology of human-associated *Akkermansia* species in the gut microbiome revealed by extensive metagenomic assembly. *Genome Biol* 22, 209. 10.1186/s13059-021-02427-7. [PubMed: 34261503]
- Kirmiz N, Galindo K, Cross KL, Luna E, Rhoades N, Podar M, and Flores GE (2020). Comparative Genomics Guides Elucidation of Vitamin B12 Biosynthesis in Novel Human-Associated *Akkermansia* Strains. *Appl Environ Microbiol* 86. 10.1128/AEM.02117-19.
- Kozhieva M, Naumova N, Alikina T, Boyko A, Vlassov V, and Kabilov MR (2019). Primary progressive multiple sclerosis in a Russian cohort: relationship with gut bacterial diversity. *Bmc Microbiol* 19. ARTN 309 10.1186/s12866-019-1685-2.
- Krebs-Smith SM, Pannucci TE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, Wilson MM, and Reedy J (2018). Update of the Healthy Eating Index: HEI-2015. *J Acad Nutr Diet* 118, 1591–1602. 10.1016/j.jand.2018.05.021. [PubMed: 30146071]
- Kurtz ZD, Muller CL, Miraldi ER, Littman DR, Blaser MJ, and Bonneau RA (2015). Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput Biol* 11, e1004226. 10.1371/journal.pcbi.1004226. [PubMed: 25950956]
- Kurtzke JF (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 33, 1444–1452. 10.1212/wnl.33.11.1444. [PubMed: 6685237]
- Langer-Gould A, Brara SM, Beaver BE, and Zhang JL (2013). Incidence of multiple sclerosis in multiple racial and ethnic groups. *Neurology* 80, 1734–1739. 10.1212/WNL.0b013e3182918cc2. [PubMed: 23650231]
- Larsen JM (2017). The immune response to *Prevotella* bacteria in chronic inflammatory disease. *Immunology* 151, 363–374. 10.1111/imm.12760. [PubMed: 28542929]
- Lazzarino G, Amorini AM, Petzold A, Gasperini C, Ruggieri S, Quartuccio ME, Lazzarino G, Di Stasio E, and Tavazzi B (2017). Serum Compounds of Energy Metabolism Impairment Are Related to Disability, Disease Course and Neuroimaging in Multiple Sclerosis. *Mol Neurobiol* 54, 7520–7533. 10.1007/s12035-016-0257-9. [PubMed: 27826747]
- Leite AZ, Rodrigues NC, Gonzaga MI, Paiolo JCC, de Souza CA, Stefanutto NAV, Omori WP, Pinheiro DG, Brisotti JL, Matheucci Junior E, et al. (2017). Detection of Increased Plasma Interleukin-6 Levels and Prevalence of *Prevotella copri* and *Bacteroides vulgatus* in the Feces of Type 2 Diabetes Patients. *Front Immunol* 8, 1107. 10.3389/fimmu.2017.01107. [PubMed: 28966614]
- Liu SR, Rezende RM, Moreira TG, Tankou SK, Cox LM, Wu M, Song A, Dhang FH, Wei ZY, Costamagna G, and Weiner HL (2019). Oral Administration of miR-30d from Feces of MS Patients Suppresses MS-like Symptoms in Mice by Expanding *Akkermansia muciniphila*. *Cell Host Microbe* 26, 779–+. 10.1016/j.chom.2019.10.008. [PubMed: 31784260]
- Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, Andrews E, Ajami NJ, Bonham KS, Brislawn CJ, et al. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 569, 655–662. 10.1038/s41586-019-1237-9. [PubMed: 31142855]
- Long T, Hicks M, Yu HC, Biggs WH, Kirkness EF, Menni C, Zierer J, Small KS, Mangino M, Messier H, et al. (2017). Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat Genet* 49, 568–578. 10.1038/ng.3809. [PubMed: 28263315]
- Lopez-Almela I, Romani-Perez M, Bullich-Vilarrubias C, Benitez-Paez A, Gomez Del Pulgar EM, Frances R, Liebisch G, and Sanz Y (2021). *Bacteroides uniformis* combined with fiber amplifies metabolic and immune benefits in obese mice. *Gut Microbes* 13, 1–20. 10.1080/19490976.2020.1865706.

- Lopez-Siles M, Duncan SH, Garcia-Gil LJ, and Martinez-Medina M (2017). *Faecalibacterium prausnitzii*: from microbiology to diagnostics and prognostics. *Isme J* 11, 841–852. 10.1038/ismej.2016.176. [PubMed: 28045459]
- Louis S, Tappu RM, Damms-Machado A, Huson DH, and Bischoff SC (2016). Characterization of the Gut Microbial Community of Obese Patients Following a Weight-Loss Intervention Using Whole Metagenome Shotgun Sequencing. *Plos One* 11. ARTN e0149564 10.1371/journal.pone.0149564.
- Lozupone C, and Knight R (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71, 8228–8235. 10.1128/AEM.71.12.8228-8235.2005. [PubMed: 16332807]
- Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, Brochado AR, Fernandez KC, Dose H, Mori H, et al. (2018). Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* 555, 623–628. 10.1038/nature25979. [PubMed: 29555994]
- Martin R, Miquel S, Chain F, Natividad JM, Jury J, Lu J, Sokol H, Theodorou V, Bercik P, Verdu EF, et al. (2015). *Faecalibacterium prausnitzii* prevents physiological damages in a chronic low-grade inflammation murine model. *Bmc Microbiol* 15, 67. 10.1186/s12866-015-0400-1. [PubMed: 25888448]
- McArdle BH, and Anderson MJ (2001). Fitting Multivariate Models to Community Data: A Comment on Distance-Based Redundancy Analysis. *Ecology* 82, 290–297. 10.1890/0012-9658(2001)082[0290:FMMTCD]2.0.CO;2.
- McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, McFarland HF, Paty DW, Polman CH, Reingold SC, et al. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 50, 121–127. [PubMed: 11456302]
- McMurdie PJ, Stoeva MK, Justice N, Nemchek M, Sieber CMK, Tyagi S, Gines J, Skennerton CT, Souza M, Kolterman O, and Eid J (2022). Increased circulating butyrate and ursodeoxycholate during probiotic intervention in humans with type 2 diabetes. *Bmc Microbiol* 22, 19. 10.1186/s12866-021-02415-8. [PubMed: 34996347]
- Miller PG, Bonn MB, Franklin CL, Ericsson AC, and McKarns SC (2015). TNFR2 Deficiency Acts in Concert with Gut Microbiota To Precipitate Spontaneous Sex-Biased Central Nervous System Demyelinating Autoimmune Disease. *J Immunol* 195, 4668–4684. 10.4049/jimmunol.1501664. [PubMed: 26475926]
- Miyauchi S, Gopal E, Fei YJ, and Ganapathy V (2004). Functional identification of SLC5A8, a tumor suppressor down-regulated in colon cancer, as a Na(+)-coupled transporter for short-chain fatty acids. *J Biol Chem* 279, 13293–13296. 10.1074/jbc.C400059200. [PubMed: 14966140]
- Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, Reyes JA, Shah SA, LeLeiko N, Snapper SB, et al. (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 13, R79. 10.1186/gb-2012-13-9-r79. [PubMed: 23013615]
- Munger KL, and Ascherio A (2011). Prevention and treatment of MS: studying the effects of vitamin D. *Mult Scler* 17, 1405–1411. 10.1177/1352458511425366. [PubMed: 21998006]
- Munger KL, Levin LI, Hollis BW, Howard NS, and Ascherio A (2006). Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. *JAMA* 296, 2832–2838. 10.1001/jama.296.23.2832. [PubMed: 17179460]
- Nayfach S, Rodriguez-Mueller B, Garud N, and Pollard KS (2016). An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res* 26, 1612–1625. 10.1101/gr.201863.115. [PubMed: 27803195]
- Nelson CA, Acuna AU, Paul AM, Scott RT, Butte AJ, Cekanaviciute E, Baranzini SE, and Costes SV (2021). Knowledge Network Embedding of Transcriptomic Data from Spaceflown Mice Uncovers Signs and Symptoms Associated with Terrestrial Diseases. *Life (Basel)* 11. 10.3390/life11010042.
- Nerurkar SN, Goh D, Cheung CCL, Nga PQY, Lim JCT, and Yeong JPS (2020). Transcriptional Spatial Profiling of Cancer Tissues in the Era of Immunotherapy: The Potential and Promise. *Cancers (Basel)* 12. 10.3390/cancers12092572.
- Olsson A, Gustavsen S, Nguyen TD, Nyman M, Langkilde AR, Hansen TH, Sellebjerg F, Oturai AB, and Sondergaard HB (2021). Serum Short-Chain Fatty Acids and Associations With Inflammation

- in Newly Diagnosed Patients With Multiple Sclerosis and Healthy Controls. *Front Immunol* 12. ARTN 661493 10.3389/fimmu.2021.661493.
- Ottoboni L, Keenan BT, Tamayo P, Kuchroo M, Mesirov JP, Buckle GJ, Khoury SJ, Hafler DA, Weiner HL, and De Jager PL (2012). An RNA profile identifies two subsets of multiple sclerosis patients differing in disease activity. *Sci Transl Med* 4, 153ra131. 10.1126/scitranslmed.3004186.
- Pang ZQ, Chong J, Zhou GY, Morais DAD, Chang L, Barrette M, Gauthier C, Jacques PE, Li SZ, and Xia JG (2021). *MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights*. *Nucleic Acids Res* 49, W388–W396. 10.1093/nar/gkab382. [PubMed: 34019663]
- Pedersen HK, Gudmundsdottir V, Nielsen HB, Hyotylainen T, Nielsen T, Jensen BA, Forslund K, Hildebrand F, Prifti E, Falony G, et al. (2016). Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* 535, 376–381. 10.1038/nature18646. [PubMed: 27409811]
- Pintaudi B, Di Vieste G, and Bonomo M (2016). The Effectiveness of Myo-Inositol and D-Chiro Inositol Treatment in Type 2 Diabetes. *Int J Endocrinol* 2016. ArtN 9132052 10.1155/2016/9132052.
- Probstel AK, and Baranzini SE (2018). The Role of the Gut Microbiome in Multiple Sclerosis Risk and Progression: Towards Characterization of the “MS Microbiome”. *Neurotherapeutics* 15, 126–134. 10.1007/s13311-017-0587-y. [PubMed: 29147991]
- Probstel AK, Zhou XY, Baumann R, Wischniewski S, Kutza M, Rojas OL, Sellrie K, Bischof A, Kim K, Ramesh A, et al. (2020). Gut microbiota-specific IgA(+) B cells traffic to the CNS in active multiple sclerosis. *Sci Immunol* 5. ARTN eabc7191 10.1126/sciimmunol.abc7191.
- Rajilic-Stojanovic M, and de Vos WM (2014). The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol Rev* 38, 996–1047. 10.1111/1574-6976.12075. [PubMed: 24861948]
- Reynders T, Devolder L, Valles-Colomer M, Van Remoortel A, Joossens M, De Keyser J, Nagels G, D’hooghe M, and Raes J (2020). Gut microbiome variation is associated to Multiple Sclerosis phenotypic subtypes. *Ann Clin Transl Neur* 7, 406–419. 10.1002/acn3.51004.
- Ritzhaupt A, Wood IS, Ellis A, Hosie KB, and Shirazi-Beechey SP (1998). Identification and characterization of a monocarboxylate transporter (MCT1) in pig and human colon: its potential to transport L-lactate as well as butyrate. *J Physiol* 513 (Pt 3), 719–732. 10.1111/j.1469-7793.1998.719ba.x. [PubMed: 9824713]
- Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, Costea PI, Godneva A, Kalka IN, Bar N, et al. (2018). Environment dominates over host genetics in shaping human gut microbiota. *Nature* 555, 210–215. 10.1038/nature25973. [PubMed: 29489753]
- Routy B, Le Chatelier E, Derosa L, Duong CPM, Alou MT, Daillere R, Fluckiger A, Messaoudene M, Rauber C, Roberti MP, et al. (2018). Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* 359, 91–97. 10.1126/science.aan3706. [PubMed: 29097494]
- Runia TF, Hop WC, de Rijke YB, Buljevac D, and Hintzen RQ (2012). Lower serum vitamin D levels are associated with a higher relapse risk in multiple sclerosis. *Neurology* 79, 261–266. 10.1212/WNL.0b013e31825fdec7. [PubMed: 22700811]
- Rusinova I, Forster S, Yu S, Kannan A, Masse M, Cumming H, Chapman R, and Hertzog PJ (2013). *INTERFEROME v2.0: an updated database of annotated interferon-regulated genes*. *Nucleic Acids Res* 41, D1040–D1046. 10.1093/nar/gks1215. [PubMed: 23203888]
- Sand IK, Zhu YJ, Ntranos A, Clemente JC, Cekanaviciute E, Brandstadter R, Crabtree-Hartman E, Singh S, Bencosme Y, Debelius J, et al. (2019). Disease-modifying therapies alter gut microbial composition in MS. *Neurol-Neuroimmunol* 6. ARTN e517 10.1212/NXI.0000000000000517.
- Sanna A, Firinu D, Zavattari P, and Valera P (2018). Zinc Status and Autoimmunity: A Systematic Review and Meta-Analysis. *Nutrients* 10. 10.3390/nu10010068.
- Scher JU, Sczesnak A, Longman RS, Segata N, Ubeda C, Bielski C, Rostron T, Cerundolo V, Pamer EG, Abramson SB, et al. (2013). Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *eLife* 2, e01202. [PubMed: 24192039]
- Shannon CE (1997). The mathematical theory of communication. 1963. *MD Comput* 14, 306–317. [PubMed: 9230594]

- Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, Gupta PB, Hao T, Silver SJ, Root DE, et al. (2009). A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* 139, 1255–1267. 10.1016/j.cell.2009.12.018. [PubMed: 20064372]
- Shokrpour M, Foroozanfard F, Ebrahimi FA, Vahedpoor Z, Aghadavod E, Ghaderi A, and Asemi Z (2019). Comparison of myo-inositol and metformin on glycemic control, lipid profiles, and gene expression related to insulin and lipid metabolism in women with polycystic ovary syndrome: a randomized controlled clinical trial. *Gynecol Endocrinol* 35, 406–411. 10.1080/09513590.2018.1540570. [PubMed: 30608001]
- Sokal RR (1982). Citation Classic - Biometry - the Principles and Practice of Statistics in Biological-Research. *Cc/Agr Biol Environ*, 22–22.
- Storm-Larsen C, Myhr KM, Farbu E, Midgard R, Nyquist K, Broch L, Berg-Hansen P, Buness A, Holm K, Ueland T, et al. (2019). Gut microbiota composition during a 12-week intervention with delayed-release dimethyl fumarate in multiple sclerosis - a pilot trial. *Mult Scler J Exp Transl Clin* 5, 2055217319888767. 10.1177/2055217319888767. [PubMed: 31798939]
- Takewaki D, Suda W, Sato W, Takayasu L, Kumar N, Kimura K, Kaga N, Mizuno T, Miyake S, Hattori M, and Yamamura T (2020). Alterations of the gut ecological and functional microenvironment in different stages of multiple sclerosis. *Proc Natl Acad Sci U S A* 117, 22402–22412. 10.1073/pnas.2011703117. [PubMed: 32839304]
- Tan TG, Sefik E, Geva-Zatorsky N, Kua L, Naskar D, Teng F, Pisman L, Ortiz-Lopez A, Jupp R, Wu HJ, et al. (2016). Identifying species of symbiont bacteria from the human gut that, alone, can induce intestinal Th17 cells in mice. *Proc Natl Acad Sci U S A* 113, E8141–E8150. 10.1073/pnas.1617460113. [PubMed: 27911839]
- Tasnim N, Abulizi N, Pither J, Hart MM, and Gibson DL (2017). Linking the Gut Microbial Ecosystem with the Environment: Does Gut Health Depend on Where We Live? *Front Microbiol* 8, 1935. 10.3389/fmicb.2017.01935. [PubMed: 29056933]
- Thingholm LB, Ruhlemann MC, Koch M, Fuqua B, Laucke G, Boehm R, Bang C, Franzosa EA, Hubenthal M, Rahnavard A, et al. (2019). Obese Individuals with and without Type 2 Diabetes Show Different Gut Microbial Functional Capacity and Composition. *Cell Host Microbe* 26, 252–+. DOI 10.1016/j.chom.2019.07.004. [PubMed: 31399369]
- Thorpe MG, Milte CM, Crawford D, and McNaughton SA (2019). Education and lifestyle predict change in dietary patterns and diet quality of adults 55 years and over. *Nutr J* 18, 67. 10.1186/s12937-019-0495-6. [PubMed: 31699092]
- Venegas DP, De la Fuente MK, Landskron G, Gonzalez MJ, Quera R, Dijkstra G, Harmsen HJM, Faber KN, and Hermoso MA (2019). Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases (vol 10, 277, 2019). *Front Immunol* 10. ARTN 1486 10.3389/fimmu.2019.01486.
- Villoslada P, Alonso C, Agirrezabal I, Kotelnikova E, Zubizarreta I, Pulido-Valdeolivas I, Saiz A, Comabella M, Montalban X, Villar L, et al. (2017). Metabolomic signatures associated with disease severity in multiple sclerosis. *Neurol Neuroimmunol Neuroinflamm* 4, e321. 10.1212/NXI.0000000000000321. [PubMed: 28180139]
- Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, and Belkaid Y (2020). Host variables confound gut microbiota studies of human disease. *Nature* 587, 448–454. 10.1038/s41586-020-2881-9. [PubMed: 33149306]
- Wang J, Jelcic I, Muhlenbruch L, Haunerding V, Toussaint NC, Zhao Y, Cruciani C, Faigle W, Naghavian R, Foege M, et al. (2020). HLA-DR15 Molecules Jointly Shape an Autoreactive T Cell Repertoire in Multiple Sclerosis. *Cell* 183, 1264–1281 e1220. 10.1016/j.cell.2020.09.054. [PubMed: 33091337]
- Wang K, Liao M, Zhou N, Bao L, Ma K, Zheng Z, Wang Y, Liu C, Wang W, Wang J, et al. (2019). Parabacteroides distasonis Alleviates Obesity and Metabolic Dysfunctions via Production of Succinate and Secondary Bile Acids. *Cell Rep* 26, 222–235 e225. 10.1016/j.celrep.2018.12.028. [PubMed: 30605678]
- Young GB, Hader WJ, Hiscock M, Warren KG, and Logan D (1986). The role of myo-inositol in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 49, 265–272. 10.1136/jnnp.49.3.265. [PubMed: 3007676]

- Zahoor I, Rui B, Khan J, Datta I, and Giri S (2021). An emerging potential of metabolomics in multiple sclerosis: a comprehensive overview. *Cell Mol Life Sci* 78, 3181–3203. 10.1007/s00018-020-03733-2. [PubMed: 33449145]
- Zapala MA, and Schork NJ (2006). Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables. *Proc Natl Acad Sci U S A* 103, 19430–19435. 10.1073/pnas.0609333103. [PubMed: 17146048]
- Zeng Q, Junli G, Liu X, Chen C, Sun X, Li H, Zhou Y, Cui C, Wang Y, Yang Y, et al. (2019). Gut dysbiosis and lack of short chain fatty acids in a Chinese cohort of patients with multiple sclerosis. *Neurochem Int* 129, 104468. 10.1016/j.neuint.2019.104468. [PubMed: 31108132]
- Zhang CY, He YJ, and Shen YF (2019). L-Lysine protects against sepsis-induced chronic lung injury in male albino rats. *Biomed Pharmacother* 117. ARTN 109043 10.1016/j.biopha.2019.109043.
- Zhou XY, Singh S, Baumann R, Barba P, Landefeld J, Casaccia P, Sand IK, Xia ZQ, Weiner H, Chitnis T, et al. (2020). Household paired design reduces variance and increases power in multi-city gut microbiome study in multiple sclerosis. *Mult Scler J. Artn 1352458520924594* 10.1177/1352458520924594.
- Zhu QY, Mai U, Pfeiffer W, Janssen S, Asnicar F, Sanders JG, Belda-Ferre P, Al-Ghalith GA, Kopylova E, McDonald D, et al. (2019). Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. *Nature Communications* 10. ARTN 5477 10.1038/s41467-019-13443-4.

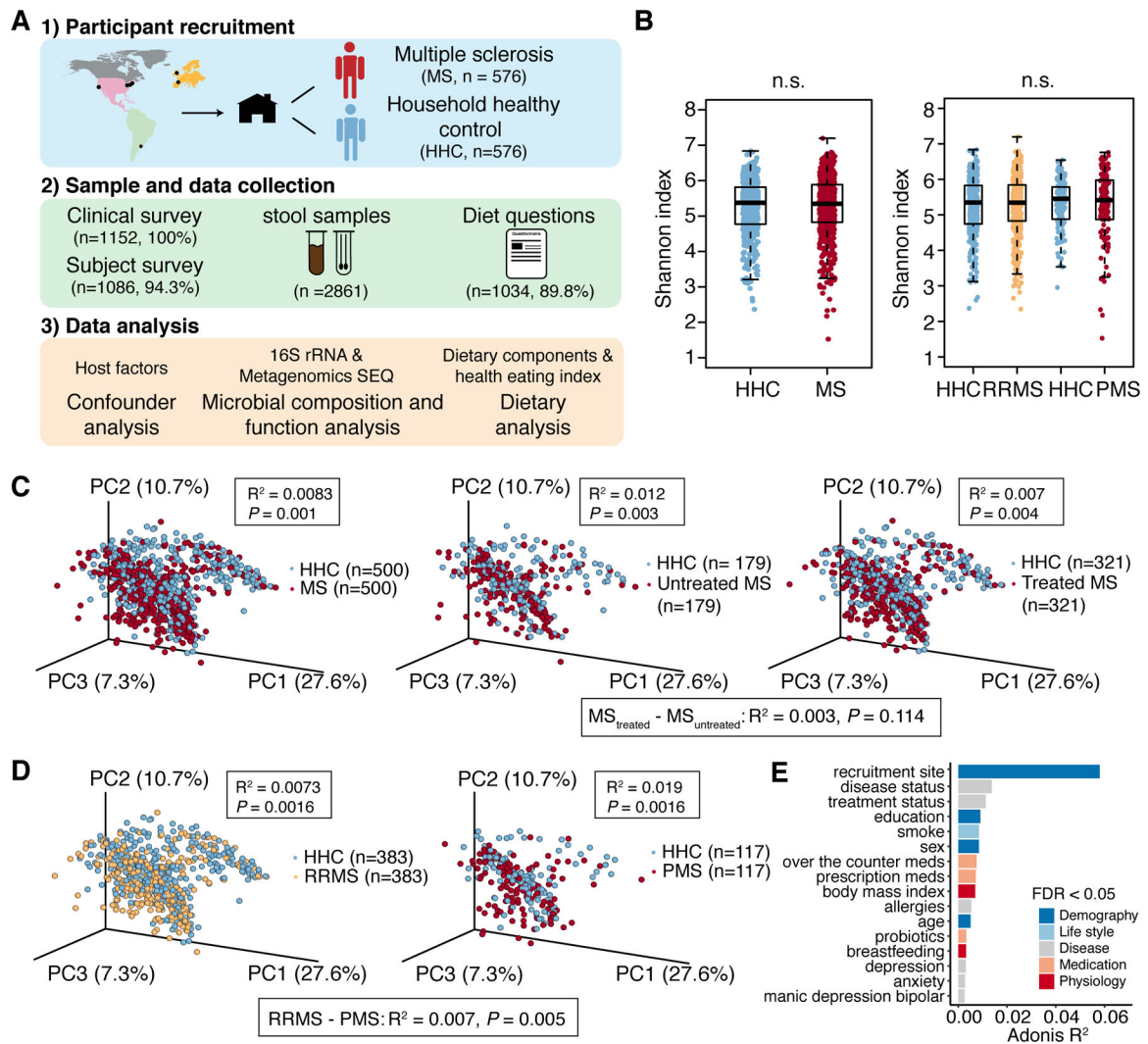


Figure 1. Study summary and overall strategy.

(A) Workflow of microbiome study in 576 MS patients and their household healthy controls. (B) Boxplot of microbiome α -diversity in MS, RRMS, PMS and their HHCs (ANOVA, n.s., not significant). (C-D) PCoA of weighted UniFrac community distance by disease and treatment status (C) and disease subtype (D) (R^2 and FDR adjusted p values were tested by PERMANOVA). (E) Bar plot showing the effect size (Adonis R^2) of confounders significantly associated with gut microbial variations (weighted UniFrac distance, PERMANOVA, FDR adjusted p < 0.05).

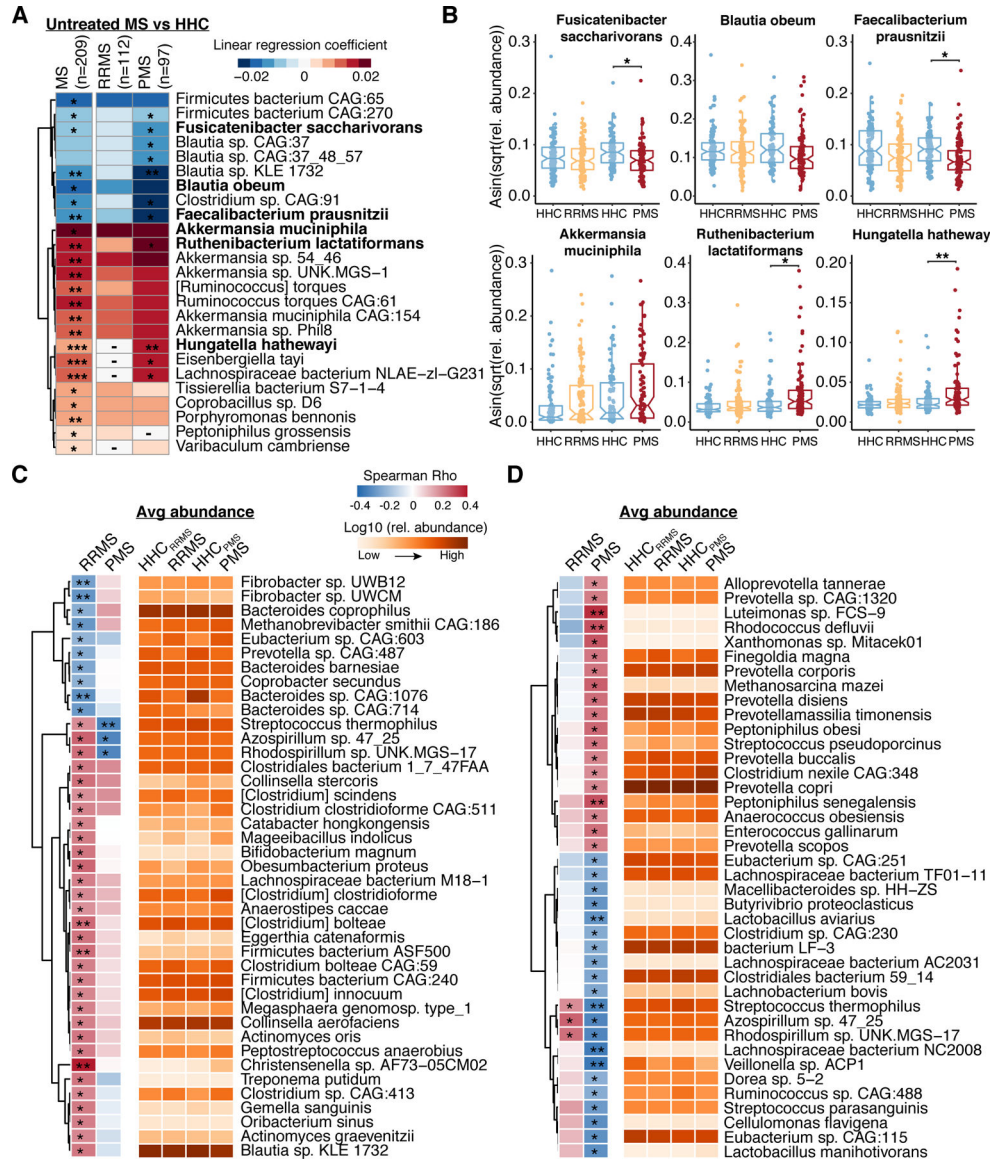


Figure 2. Microbial taxa alterations between MS and HHC.

(A) Taxa altered in untreated MS (n= 209), untreated RRMS (n=112) or untreated PMS (n=97) versus their HHCs (mixed linear regression model adjusted for age, BMI, sex, recruiting site and house). “-” indicates species with lower variance across samples were filtered out and not included in linear regression. *FDR < 0.05, **FDR< 0.01, ***FDR< 0.001. (B) Arcsine square-root transformed relative abundance of 3 decreased species and 3 increased species in untreated MS versus HHCs. (C-D) Species were significantly correlated with MS Severity Scores (MSSS) in untreated RRMS patients (n=112, C) or in untreated PMS (n=97, D). Spearman correlations were adjusted for age and body mass index. *p < 0.05, **p < 0.01. Averaged abundance of significant species are shown in untreated RRMS untreated PMS and their corresponding HHCs.

pathways altered in treated and untreated RRMS (mixed linear regression model adjusted for age, BMI, sex, recruiting site and house, *FDR < 0.05, **FDR < 0.01, ***FDR < 0.001). (F) Pathways were significantly correlated with MSSS in untreated RRMS patients (RRMS, n=112, top panel) or in untreated PMS (PMS, n=97, bottom panel). Spearman correlations are adjusted for age and body mass index. *p < 0.05, **p < 0.01. Averaged abundances of significant pathways are shown in untreated RRMS and untreated PMS compared to their corresponding HHCs.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

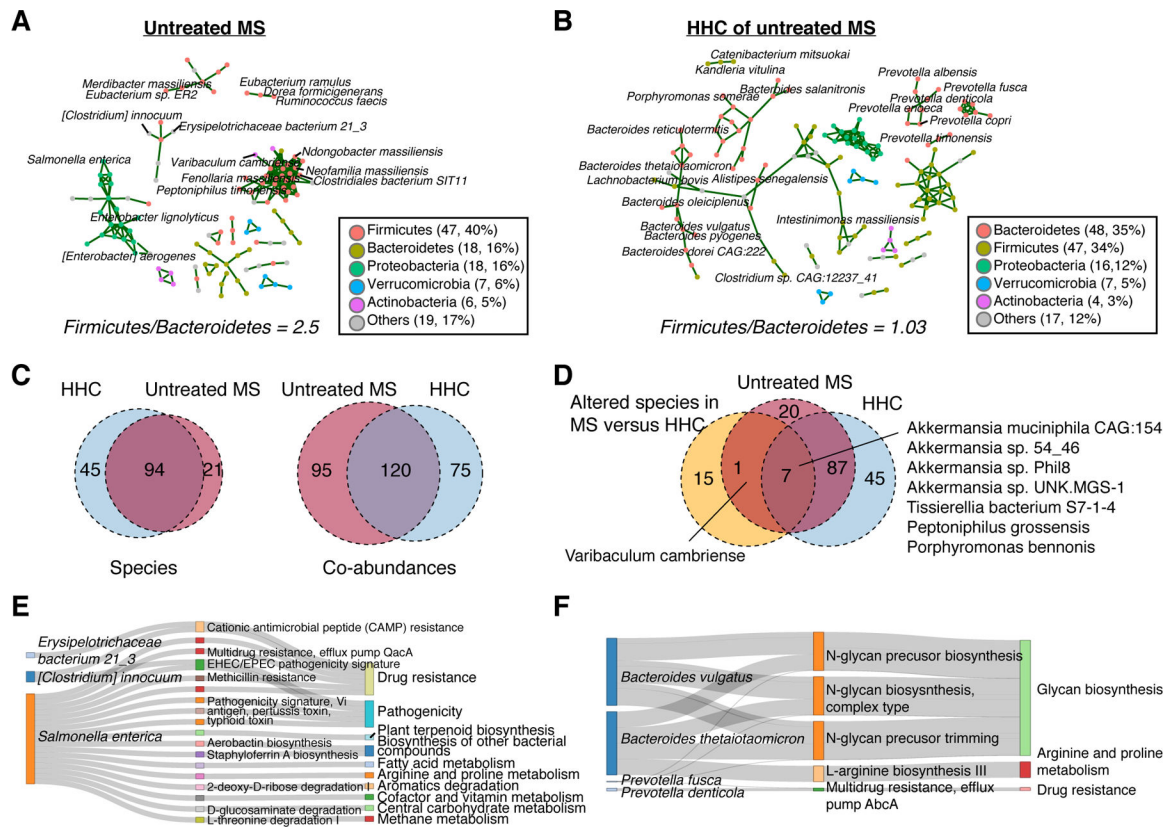


Figure 4. Disease status specific co-abundance species.

Microbial co-abundance communities specific for (A) untreated MS and (B) HHCs by cohort specific analysis (quantile range outlier). Each node indicates one species and color indicates the phylum classification. Each edge represents a significant species-species co-abundance relationship. (C) Overlapped counts of species and co-abundances in untreated MS specific and HHC specific networks. (D) Differential species in untreated MS versus HHC were overlapped with cohort specific species. (E-F) Functional pathways unique to the species highlighted in untreated MS- (E) or HHC (F) specific networks. Line size indicates betweenness centrality of a species in the cohort specific co-abundance network.

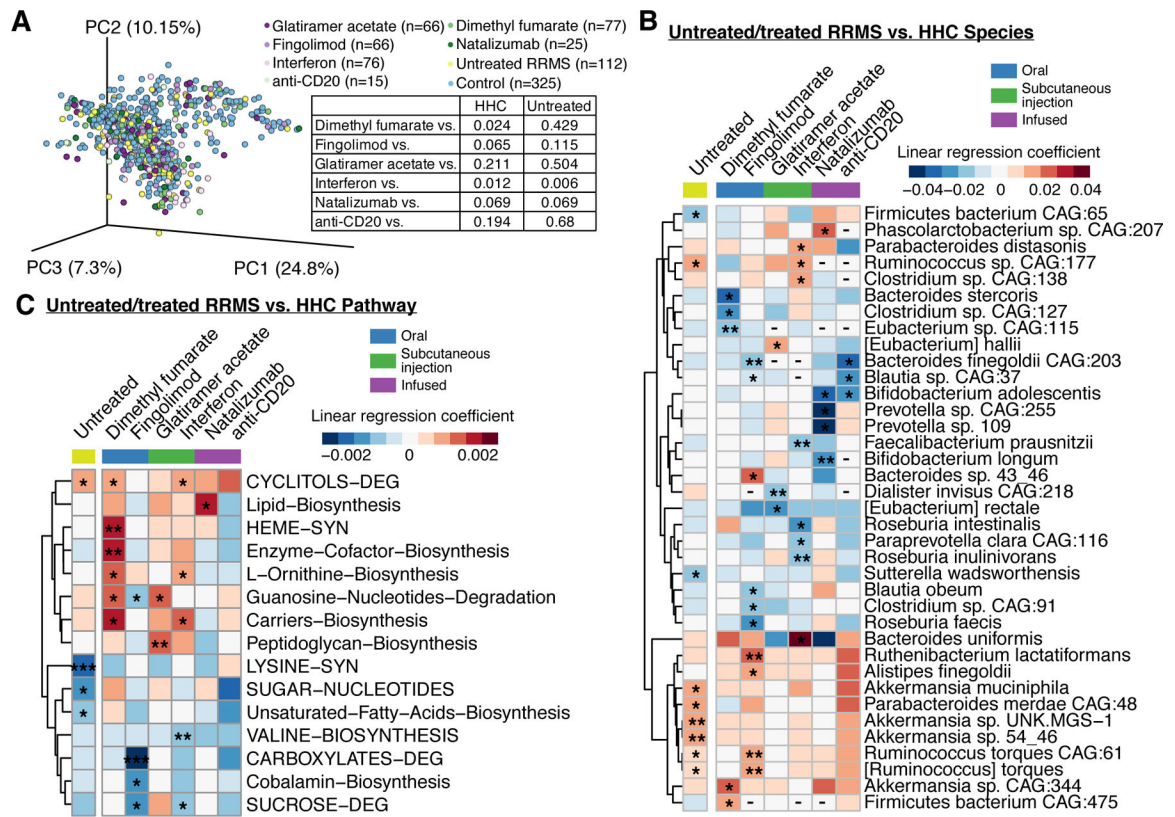


Figure 5. Treatment -associated metagenomic changes in RRMS patients.

(A) PCoA of weighted UniFrac community distance of RRMS subjects treated and untreated, and their corresponding household healthy controls (P values were obtained by PERMANOVA). (B) metagenomics species (C) metabolic pathways altered in treated and untreated RRMS (mixed linear regression model adjusted for age, BMI, sex, recruiting site and house). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and linear coefficient upper 5% or coefficient lower 5%.

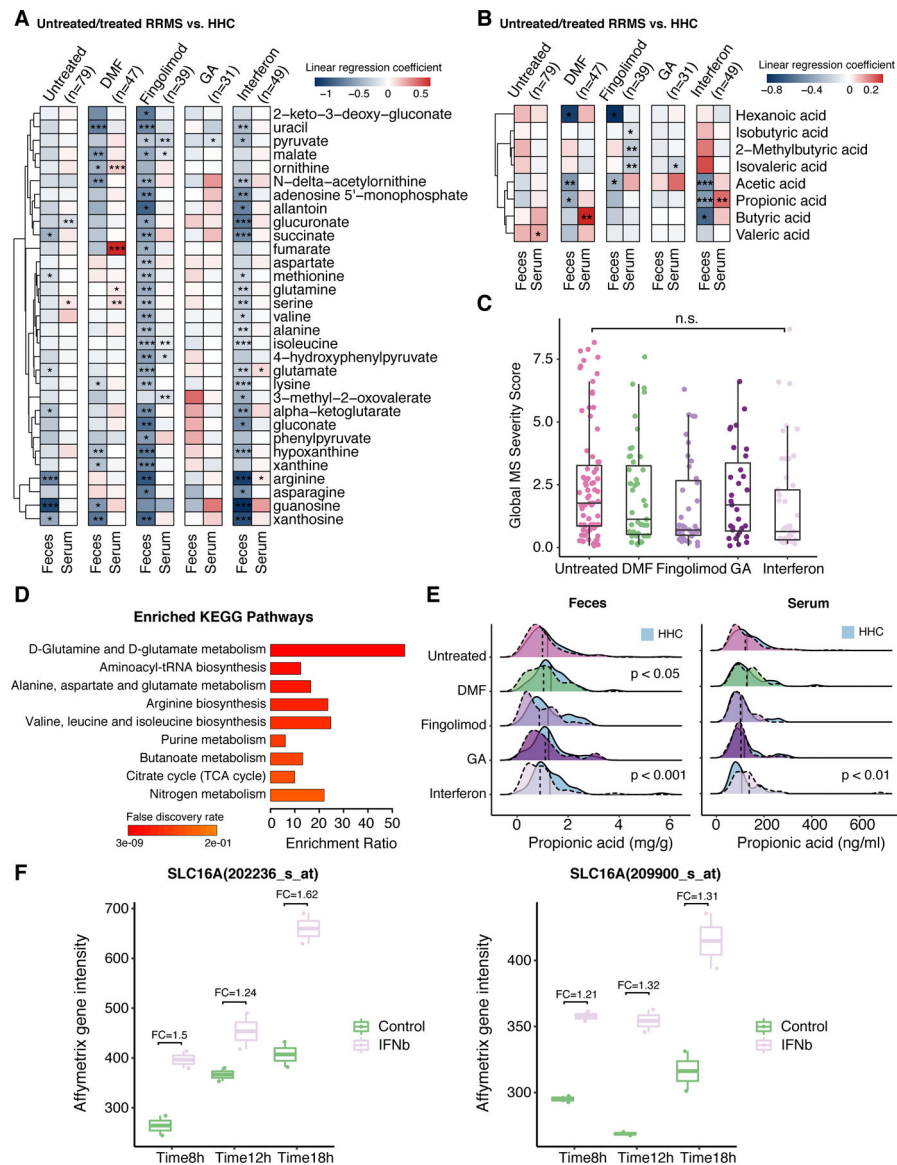


Figure 6. Treatment-associated metabolomic alterations in RRMS patients.

(A) 31 microbe-derived metabolites and (B) 8 short chain fatty acids in treated and untreated RRMS in both stool and serum. Linear coefficient was measured by mixed linear regression model adjusted for age, BMI, sex, recruiting site and house. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (C) Disease duration adjusted MS severity score (gMSSS) was compared between untreated and treated RRMS (ANOVA). (D) KEGG pathways enriched by 23 microbe-derived metabolites in response to interferon (FDR < 0.05). (E) Concentration of propionic acid in feces (left) and serum (right) compared for treated and untreated RRMS, compared to their respective HHCs. DMF, dimethyl fumarate, GA, glatiramer acetate. (F) Expression of SLC16A in human bronchial epithelial cells stimulated by IFN- β from study by Shapira, S. D. et al. The SLC16A gene was represented by two probes (202236_at and 209900_s_at) of Affymetrix HT Human Genome U133 Arrays.

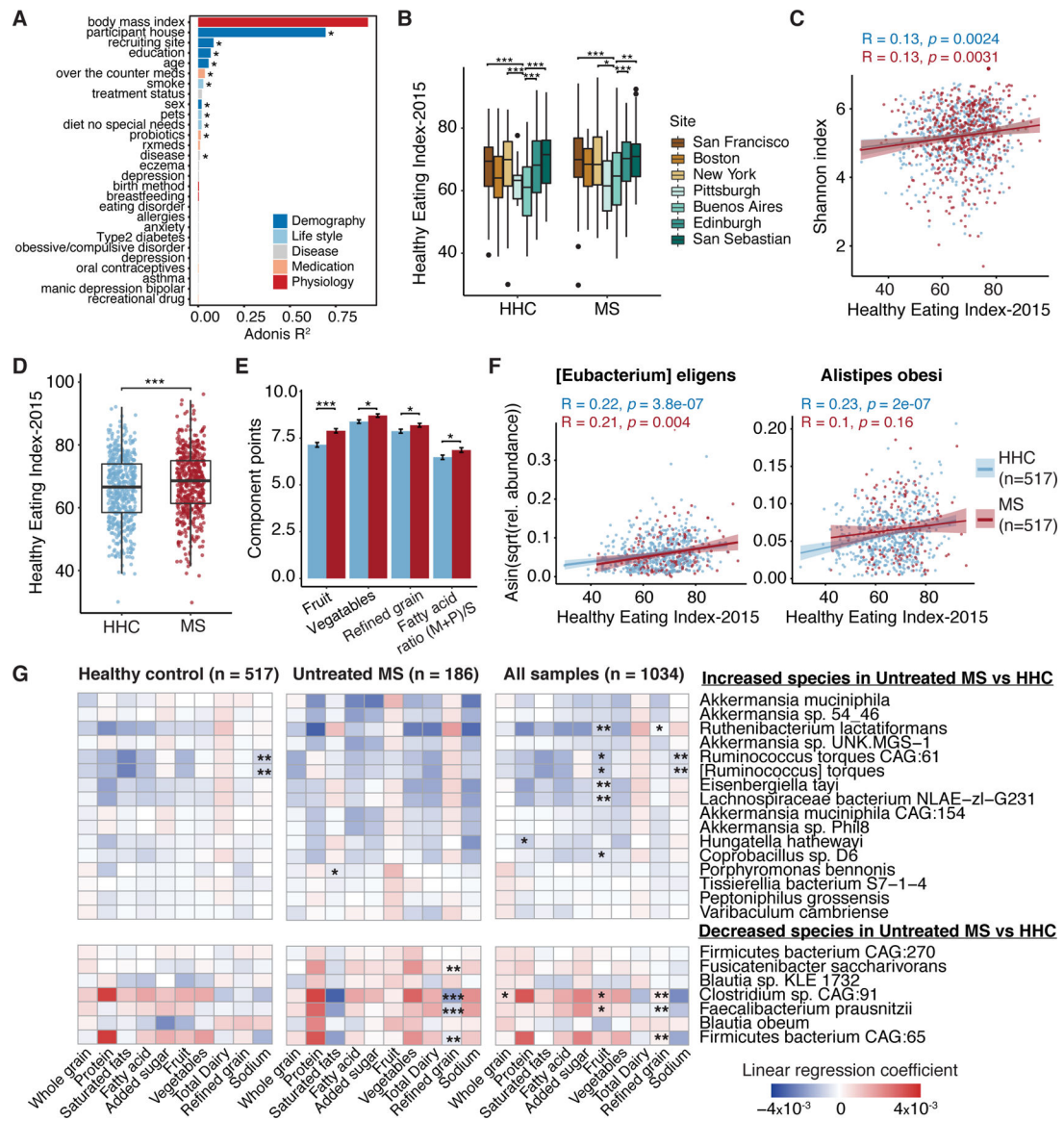


Figure 7. Diet and gut microbes.

(A) Bar plot showing the effect size (Adonis R^2) of confounders associated with dietary variations (Jaccard dissimilarity). Confounders showing a significant impact on gut microbiome were labeled (PERMANOVA, *FDR ≤ 0.05). (B) Boxplot of healthy eating index measured in the participants from each recruiting site. (C) Pearson's correlation between healthy eating index and microbial α -diversity in healthy (blue) and MS (red) individuals. (D) Boxplot of healthy eating index measured in MS patients and their household healthy controls (paired T-test, *** $p < 0.001$). (E) Difference in dietary components between MS and HHC individuals (paired T-test, * $p < 0.05$, *** $p < 0.001$). (F) Species significantly correlated with HEI (Pearson's correlation with FDR < 0.05). (G) Correlations between dietary component and MS-associated species measured in healthy controls, untreated MS and all samples, respectively (mixed linear regression model adjusted for age, BMI, sex and recruiting site, *FDR < 0.05 , **FDR < 0.01 , ***FDR < 0.001).

Table 1.

Sample characteristics for 576 pairs of MS and their household healthy controls.

	HHC		MS		RRMS		PMS	
	n	%	n	%	n	%	n	%
Number	576	50	576	50	437	75.9	139	24.1
Age (y)	50.6 (40.8–61)		48.9 (40–59)		45.8 (37–55)		58.6 (54–65)	
Female	201	34.9	400	69.4	312	71.4	88	63.3
BMI	26.9 (23.5–29)		25.4 (21.8–27.6)		25.4 (21.8–27.5)		25.3 (21.8–27.8)	
EDSS			2.6 (1–4)		1.77 (0–2.5)		5.21 (3.75–6.5)	
Disease duration (y)			14.2 (6–21)		12.5 (5–18)		19.6 (9.5–28.5)	
MSSS			3.37 (0.86–5.57)		2.5 (0.655–3.65)		6.11 (4.74–7.53)	
Untreated			209	36.3	112	25.6	97	69.8
Treated			367	63.7	325	74.4	42	30.2
Treatment								
Fingolimod			71	12.3	66	15.1	5	3.6
Dimethyl fumarate			86	14.9	77	17.6	9	6.5
Glatiramer acetate			68	11.8	66	15.1	2	1.4
Interferon			87	15.1	76	17.4	11	7.9
anti-CD20			28	4.9	15	3.4	13	9.4
Natalizumab			27	4.7	25	5.7	2	1.4
Recruiting site								
San Francisco	164	28.5	164	28.5	110	25.2	54	38.8
Boston	42	7.3	42	7.3	35	8.0	7	5.0
New York	59	10.2	59	10.2	45	10.3	14	10.1
Pittsburgh	12	2.1	12	2.1	12	2.7	0	0.0
Buenos Aires	129	22.4	129	22.4	120	27.5	9	6.5
Edinburgh	131	22.7	131	22.7	82	18.8	49	35.3
San Sebastián	39	6.8	39	6.8	33	7.6	6	4.3

Data are presented as mean (interquartile range, IQR); y, year; BMI, Body Mass Index; EDSS, Expanded Disability Status Scale; MSSS, Multiple Sclerosis Severity Score.

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Human stool samples	This paper	n/a
Human serum samples	This paper	n/a
Critical commercial assays		
QIAamp PowerFecal DNA Kit	QIAGEN	12830–50
MagAttract PowerSoil DNA EP Kit	QIAGEN	27100–4-EP
Kapa Illumina Library Quantification Kit	Roche	07962428001
Pico Green Quantification Kit	Invitrogen/Thermo	P11496
Deposited data		
Raw microbiome data	This paper	ENA: ERP115476
Code used for data analysis	This paper	https://github.com/BaranziniLab/iMSMS_study
Dryad datasets, supplementary data I–VI. Code used for data analysis	This paper	https://doi.org/10.7272/Q60C4T26
Web of Life	Knight Lab	https://biocore.github.io/wol/
Software and algorithms		
R	The R foundation	https://www.r-project.org/
SHOGUN	Knight lab	https://github.com/knights-lab/SHOGUN
QIITA	Knight lab	
QIIME2	Knight lab	https://qiime2.org/
SPOKE	Baranzini lab	https://spoke.ucsf.edu/
MetaCyc	(Caspi et al., 2016)	https://metacyc.org/
KEGG	(Kanehisa and Goto, 2000)	https://www.kegg.jp/
MetaboAnalyst 5.0.	(Pang et al., 2021)	https://www.metaboanalyst.ca/
HUMAN2	(Franzosa et al., 2018)	https://pypi.org/project/humann2/