PAIN Publish Ahead of Print
DOI: 10.1097/j.pain.0000000000001640

Altered microbiome composition in individuals with fibromyalgia

Authors: Amir Minerbi[1], Emmanuel Gonzalez[2,3], Nicholas J.B. Brereton[4], Abraham Anjarkouchian[5], Ken Dewar[3,6], Mary-Ann Fitzcharles[1,7], Stéphanie Chevalier[5,8,9], Yoram Shir[1]

[1] Alan Edwards Pain Management Unit, McGill University Health Centre, Montreal, Quebec, Canada

[2] Canadian Center for Computational Genomics, McGill University and Genome Quebec Innovation Center, Montreal, Quebec, Canada

[3] Department of Human Genetics, McGill University, Montreal, Quebec, Canada

[4] Institut de recherche en biologie végétale, University of Montreal, Montreal, QC H1X 2B2, Canada

[5] School of Human Nutrition, McGill University, 21111 Lakeshore Rd, Ste-Anne-de-Bellevue, QC H9X 3V9, Canada

[6] McGill University and Genome Quebec Innovation Center, McGill University, Montreal, Quebec, Canada

[7] Division of Rheumatology, McGill University Health Centre, Quebec, Canada

[8] Research Institute of the McGill University Health Centre, 1001 Décarie Blvd, Montreal, QC H4A 3J1, Canada

[9] Department of Medicine, McGill University, 845 Sherbrooke St. W, Montreal, QC H3A 0G4, Canada

Toatl number of pages: 43

Number of figures: 5

Number of supplementary figures: 3

Number of supplementary tables: 6

Address of corresponding author: Amir Minerbi, Montreal General Hospital, McGill University Health Centre, 1650 Cedar ave., Montreal, Quebec, H3G 1A4 Tel: 514-9348222 Fax: 514-934-8096 E-mail: minerbi@gmail.com

Abstract

Fibromyalgia (FM) is a prevalent syndrome, characterised by chronic widespread pain, fatigue and impaired sleep, that is challenging to diagnose and difficult to treat. The microbiomes of 77 women with FM and that of 79 control participants were compared using 16S rRNA gene amplification and whole genome sequencing. When comparing FM patients to unrelated controls using differential abundance analysis, significant differences were revealed in several bacterial taxa. Variance in the composition of the microbiomes was

explained by FM-related variables more than by any other innate or environmental variable and correlated with clinical indices of FM. In line with observed alteration in butyrate metabolising species, targeted serum metabolite analysis verified differences in the serum levels of butyrate and propionate in FM patients. Using machine learning algorithms, the microbiome composition alone allowed for the classification of patients and controls (ROC AUC 87.8%). To the best of our knowledge, this is the first demonstration of gut microbiome alteration in non-visceral pain. This observation paves the way for further studies, elucidating the pathophysiology of FM, developing diagnostic aids and possibly allowing for new treatment modalities to be explored.

Keywords: Fibromyalgia; widespread-pain; microbiome; gut-brain-axis

Introduction

Fibromyalgia is one of the most common forms of chronic widespread pain, with estimated prevalence of 2-4% of the adult population [15,31,72]. Characterised by pain, physical exhaustion, sleep disturbance and cognitive symptoms, FM leads to a significant impairment in quality of life of affected individuals [31,53,72,80]. The pathophysiology of FM is not well understood; with multiple hypotheses being suggested including impaired central nervous system nociceptive processing, altered peripheral nociception and systemic inflammation [80].

The diagnosis of FM is based on the recognition of a typical cluster of symptoms, while excluding other potential sources of pain [31]. The diagnostic criteria for FM rely on self-reported symptoms [94]. The lack of objective diagnostic criteria is a source of frustration among patients and clinicians, and adds to the possible reasons for inaccurate diagnoses. In two retrospective cohorts of patients diagnosed with FM, the rate of false positive diagnoses

was estimated at 66-73% [23,90]. Even when provided with best available treatment many patients continue to suffer from significant symptoms [31]. In recent years, evidence is mounting on the critical role of the gut microbiota in a variety of pathologies, including, but not limited to, metabolic, cardiovascular, oncologic, neurologic and psychiatric disorders [34,50]. The scientific literature is rich in studies on different aspects of the interactions between the host and the microbiome; however, data on the possible role of the gut microbiota in the pathophysiology of chronic pain outside of the gastrointestinal tract is still scarce. Our increasing understanding of the interactions between the gut microbiota and the central nervous system, also known as the 'gut-brain axis', makes reasonable the hypothesis that it may also affect pain processing and perception [6,7,16,22,25,54]. Recently, this hypothesis has been supported by several animal studies, which have shown that gut microbiota play an important role in the development of visceral pain [49,65,66], of chemotherapy-induced neuropathic pain [77] and of opioid tolerance [36]. Human studies have thus far focused on alterations of the gut microbiota in several visceral pain disorders, showing consistent alterations in individuals with irritable bowel syndrome (IBS) [17,20,71,84,97], and in patients suffering from chronic dysfunctional pelvic pain [5,10,78]. Similarly, patients with chronic fatigue syndrome, which shares some symptomatic features with FM, were shown to have altered gut microbiome and metabolomic profiles [26,28,61,62]. Finally, in several rheumatologic diseases, including rheumatoid arthritis and spondyloarthropathies, microbiome alterations have been reported [33,93,95,96].

Indirect evidence hints that the gut microbiome may be altered in FM patients: Altered small intestinal permeability was reported in a cohort of FM and complex regional pain syndrome (CRPS) patients [29]; in a small cross-sectional study of FM patients, a distinct urine

metabolomic signature was demonstrated, which could be attributed to gut microbiome alterations [52].

Despite recent reports of gut microbiome alterations in certain inflammatory rheumatic diseases, IBS and IC, there is still no evidence of such alterations in non-visceral pain. Here, variation in the composition of the gut microbiome between FM patients and control participants is investigated for the first time.

Methods

Study design and oversight

The study took place at the Alan Edwards Pain Management Unit (AEPMU) of the McGill University Health Centre, Montreal (MUHC), and at the West Island Rheumatology Clinic, in Montreal, Quebec, Canada. The study was approved by the MUHC institutional review board. All participants were given a detailed explanation of the study and signed an informed consent form.

Patient recruitment and clinical evaluation

Individuals with FM and control participants were recruited as follows: patients were recruited at the AEPMU and at the West Island Rheumatology Clinic. Patients were contacted either by their treating physician, or by advertisements in the local media and a dedicated website. Inclusion criteria for patients were: female sex, age 30-60 years, widespread pain index and symptoms severity scores compatible with the 2016 diagnostic criteria for FM and ability to give informed consent in French or in English. Three groups of control participants were recruited: 1) First degree female relatives of patients participating in the study were recruited as genetic controls.

This group included adult women; 2) Household members of patients participating in the study were recruited as environment controls. This group included women and men aged 30-70 years; 3) Unrelated healthy women aged 30-60 years.

Exclusion criteria were: major chronic illness (e.g. malignancy, active inflammatory disease, metabolic disease), antibiotics treatment in the preceding two months, any acute illness in the preceding month, change in regularly taken medication in the preceding month and substantial dietary alterations in the preceding month.

All patients with chronic pain were interviewed by a specialized pain physician for a thorough assessment. Only individuals whose diagnosis of FM was confirmed were deemed eligible to participate in the study as patients. All patient and control participants were interviewed by a specialized pain physician, and data were collected regarding their demographics, anthropometrics, co-morbidities (including a specific evaluation for irritable bowel syndrome, based on ROME IV criteria [45,56]), medications, dietary intake, smoking and alcohol consumption. Comorbidities were recorded based on a systematic clinical interview by a pain medicine physician and a review of participants' regular medications. Participants then filled-in the following questionnaires: The FM Survey Diagnostic Criteria and Severity Scale (FSDC) questionnaire, assessing symptom severity, pain location, fatigue, sleep quality and cognitive and somatic complaints in FM patients, based on the 2016 criteria for the diagnosis of FM [94]; The FM Impact Questionnaire (FIQ), a 10-item questionnaire evaluating physical functioning, work difficulty, pain, fatigue, morning tiredness, stiffness, anxiety and depression [12]; Physical activity assessment using The Physical Activity self-Administered Questionnaire, a 22-item questionnaire evaluating the average individual physical activity level [89]; Sleep Scale from the Medical Outcomes Study, a 12-item sleep quality evaluation questionnaire [59]. All

questionnaires were offered in English or in French and have been validated in both languages [14,24,69,89]. Participants were then interviewed by a dietitian, who supervised them as they filled-in the web-based NIH ASA24-Canada dietary recall, validated in English and in French [87].

Dietary intake assessment

Dietary intake was assessed using the NIH Automated Self-Administered 24-hour recall (ASA24-Canada version 2016, English or French). This tool has been extensively validated [67,87] During the on-site visit, participants were trained and supervised by a dietitian as they filled the first day of recall on-line. They were asked to fill two additional separate days during the following week, allowing for a total of 3-day dietary assessment, including two weekdays and one weekend day. Datasets with at least two complete days were kept (n=146) and daily nutrient averages were calculated. Nutrient intakes were compared among groups as absolute intakes and relative to energy or body weight. Under-reporting was investigated by applying a ratio of reported energy intake over resting energy expenditure estimated with the Mifflin-St. Jeor equation [57]; participants with a ratio below 0.9 were deemed as under-reporters and excluded from analysis. The Healthy Eating Index (HEI-2015) was calculated to assess diet quality [40].

Sample acquisition and handling

Patients were asked to collect a stool sample at home, using the Omnigen Gut OM-200 kit (DNA Genotek, Ottawa, Canada). Stool was fresh-frozen at -20C, and delivered by the participants to the study facility within ten days of its collection. Proper delivery conditions were assured by using a dedicated thermal bag (Thermos, IL, USA), containing an ice pack and monitored using

an adhesive thermal indicator (Warm Mark, MesaLabs, CO, USA), which was attached to the sample bag. Upon arrival to the study facility, thermal indicators were read to assure proper delivery conditions and the stool samples were frozen at -80C.

DNA extraction, 16S ribosomal RNA gene amplification and Illumina Sequencing

DNA was extracted from all 156 stool samples using the QIAamp PowerFecal DNA kit (Qiagen, Netherlands) following the manufacturer's instructions. The V5-V6 region (based on *Escherichia coli*) of the 16S ribosomal RNA (rRNA) was targeted for amplification by PCR using the forward primer: S-D-Bact-0785-a-S-18, GGMTTAGATACCCBDGTA and reverse primer: S-*-Univ-1100-a-A-15, GGGTYKCGCTCGTTR [39]. The CS1 (ACACTGACGACATGGTTCTACA) and CS2 (TACGGTAGCAGAGACTTGGTCT) tags were used to add a barcode and Illumina adapters. Amplification was performed using Q5 High Fidelity DNA polymerase (New England BioLabs) with PCR cycles as follows: initial denaturation step of 98°C, for 30 secs, before 23 cycles of 98°C, for 10 secs, 58°C, for 15 secs and 72°C, for 30 secs, with the final extension at 72°C, for 2 min. The MiSeq250 platform was used for 2 x 250 nucleotides (nt) paired-end sequencing of the resulting PCR products. The ANCHOR pipeline was used to process and annotate sequence reads (https://github.com/gonzalezem/ANCHOR). Briefly, sequences were aligned and dereplicated using Mothur [75] before selection of OTUs using a count threshold of 48 across all samples. Annotation used four sequence repositories with strict BLASTn criteria (>99% identity and coverage): NCBI curated bacterial and Archaea RefSeq, NCBI nr/nt, SILVA, Ribosomal Database Project (RDP) (NCBI curated bacterial and Archaea RefSeq is given a priority when at 100% identity and coverage). When the highest identity/coverage is shared amongst different putative annotation, all are retained and reported (borrowed from the idea of secondary

annotation in metatranscriptomics [30]). Amplicons with low-counts (<48) are binned to high-count sequences in a second BLASTn, using a lower threshold of >98% identity/coverage.

Whole genome sequencing, analysis using MetaPhlAn2 and mapping to 16S rRNA gene OTUs

Genomic DNA was quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies). Libraries were generated using the NEB Ultra II DNA kit (NEB) as per the manufacturer's recommendations. TruSeq adapters with unique dual indices and PCR primers were purchased from IDT. Size selection of libraries at 360nt was performed using SPRIselect beads (Beckman Coulter). Libraries were quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). Average size fragment was determined using a LabChip GX (PerkinElmer) instrument. The libraries were normalized, pooled and denatured in 0.05N NaOH and then neutralized using HT1 buffer. ExAMP was added to the mix to bring the final concentration to 200 pM following the manufacturer's instructions. A phiX library was used as a control and mixed with libraries at 1% level. The pool was loaded on a Illumina cBot and the flowcell was ran on a HiSeq 4000 for 2x100 cycles (150 nt, paired-end mode). The Illumina control software was HCS HD 3.4.0.38, the real-time analysis program was RTA v. 2.7.7. Program bcl2fastq2 v2.18 was then used to demultiplex samples and generate fastq reads. Trimmomatic [44] was used to trim nucleotides of poor quality and reads < 100 nt were removed (parameters: LEADING:25 TRAILING:25 MINLEN:100). WGS reads were analysed in two ways: by directly mapping them to the 16S rRNA gene ANCHOR OTU sequences and using MetaPhlAn2 [76]. Bowtie2 [42] was used to align WGS reads against MetaPhlAn2 v2.2.0 (default parameters). Mapping of WGS reads directly to the ANCHOR OTU table was performed using BLASTn (100% identity and 100% query coverage thresholds).

Canonical correspondence analysis and prediction

Canonical correspondence analysis (CCA) was performed on raw OTU counts of FM patients and unrelated controls (PERMANOVA, $p < 0.001$). Household and family controls were projected on the axes post-hoc (predict.cca function, R Vegan package [63]).

Microbiome explained variance analysis

Unconstrained ordination was calculated using Nonmetric Multidimensional Scaling (NMDS) transformation on Bray-Curtis distances (metaMDS function, R Vegan package). Environmental factors were projected onto ordination diagram and each variable regression was independently tested by Monte Carlo permutation test (envfit function, R Vegan package) [63].

DESeq2 differential abundance analysis

Differential abundance analysis was performed using 16S rRNA in accordance with DESeq2 [48,85], which has been shown to perform well when applied to the uneven library sizes and sparsity common in 16S rRNA gene marker data [92]. DA selection parameter of false discovery rate (FDR; Benjamini-Hochberg procedure) < 0.1 was applied [4,48,47] (Supplementary File, available at http://links.lww.com/PAIN/A839). Regularized log transformation was applied to raw counts across samples (rlog function, R phyloseq package). A sparsity and low-count cut-offs were used whereby an OTU count in a single sample is < 90% of the count in all samples, and OTU count must be > 2 in 40% of the samples.

Correlation between taxa abundance and clinical indices

Spearman's correlation was calculated between log-transformed OTU abundance and continuous or rank clinical variables. Individual p value was calculated for comparison, and correction for multiple comparisons was done using Benjamini-Hochberg FDR, with a cut-off value of 0.05. A

colour coded correlation matrix representing taxa abundance vs. clinical indices was elaborated, and sorted based on hierarchical clustering of DA OTUs using MathWorks MATLAB.

OTU-OTU correlation

Pearson's correlation was calculated between normalized OTU counts. Correlation matrices were sorted based on hierarchical clustering of OTUs and represented using R package corrplot [91].

Machine learning instances for FM prediction

The L1-penalized regression or LASSO (least absolute shrinkage and selection operator) [86] was used to predict FM diagnosis from the ANCHOR DA OTU count table and was run using glmnet R package [27]. Several count tables were tested as pre-selection datasets for the LASSO including full and reduced (sparsity filtered, DESeq2 DA OTUs), each with different count transformation or normalization strategies (raw, log, quantile, square root; supplementary figure 5, available at http://links.lww.com/PAIN/A839). DESeq2 DA OTUs (72 features) performed better in diagnosis prediction than sparsity filtered counts (1071 features), so were considered a good pre-selection dataset in line with previous research [99]. Random OTU count table subsets of equivalent size (each with ~72 OTUs) were also compared against the DESeq2 DA OTU subset, with DA OTUs again showing better prediction power (Supplementary Figure 3, available at http://links.lww.com/PAIN/A839). Raw counts outperformed transformed data (log, square root) and normalized counts (quantile) in their prediction and were used for subsequent analysis. LASSO's integrated feature selection (i.e. selection of sets of features leading to best prediction) was run on DESeq2 DA count table (72 OTUs) using a 10-fold multinomial cross-validation and repeated 10,000 times. A total of 123 high predictive DA OTU sets, i.e. sets of features selected multiple times by LASSO with a high prediction (median AUC>0.9) or selected

a single time with a perfect prediction (AUC of 100%), were selected for Leave-one-out cross-validation (LOOCV) multinomial LASSO. LOOCV was performed on training datasets composed of 90% of the participants and was repeated 1000 times. Prediction was evaluated on the remaining 10% of participants (test dataset). ROC (Receiver Operating Characteristic) curves (pROC R package [74]) were used to illustrate the discriminative ability of the classifier while prediction performance was evaluated by the area under the curve (AUC) metric. The related support vector machine (SVM) technique [35] was used based on the best predicting DA OTUs and 90% of the samples and the diagnosis prediction was evaluated (AUC on 10% on the total number of participant).

Targeted serum metabolite analysis

Serum levels of butyric acid, isobutyric acid and propionic acid were measured as previously described [18]: Targeted metabolite quantification approach was used to analyze the serum samples using a combination of direct injection mass spectrometry with a reverse-phase LC-MS/MS custom assay. The method used combines the derivatization and extraction of analytes, and the selective mass-spectrometric detection using multiple reaction monitoring (MRM) pairs. Isotope-labeled internal standards and other internal standards are used for metabolite quantification. The custom assay contains a 96 deep-well plate with a filter plate attached with sealing tape, and reagents and solvents used to prepare the plate assay. First 14 wells were used for one blank, three zero samples, seven standards and three quality control samples. 150 uL of ice-cold methanol and 10 uL of isotope-labeled internal standard mixture was added to 50 uL of serum sample for overnight protein precipitation. Then it was centrifuged at 13000x g for 20 min. 50 uL of supernatant was loaded into the center of wells of a 96-deep well plate, followed

by the addition of 3-nitrophenylhydrazine (NPH) reagent. After incubation for 2h, BHT stabilizer and water were added before LC-MS injection.

Mass spectrometric analysis was performed on an API4000 Qtrap® tandem mass spectrometry instrument (Applied Biosystems/MDS Analytical Technologies, Foster City, CA) equipped with a solvent delivery system. Data analysis was performed using Analyst 1.6.2.

General statistical considerations

Clinical indices, including demographics, anthropometrics, co-morbidities and medications, in the study groups were compared using IBM® SPSS® Statistics version 23 and MathWorks MATLAB® version 2018b, using ANOVA for univariate comparisons and MANOVA for multivariate comparisons. Metabolite concentrations were compared similarly.

Dietary analysis was done as follows: normality of distribution was explored using Kolmogorov-Smirnov and Shapiro-Wilk tests. Most dietary data were assessed with one-way ANOVA and post-hoc Games-Howell test (no assumption of equal variance and appropriate for unequal sample size). Some nutrients were not normally distributed and analysed with non-parametric Kruskal-Wallis test. Analyses were conducted using IBM® SPSS® Statistics version 23.

Formal power analysis for sample-size calculation was not performed, but sample size was not smaller than similar previous studies of microbiome characterization in other medical conditions [95].

Data availability

All data generated as part of this study are included in this published article and its supplementary file or are available from the corresponding authors on reasonable

request. Supplementary file is available at: https://github.com/gonzalezem/Fibromyalgia. Genome sequences can be found at: https://www.ncbi.nlm.nih.gov/sra/PRJNA521398 (16S rRNA sequences) and https://www.ncbi.nlm.nih.gov/sra/PRJNA521587 (WGS sequences).

Results

Participant characteristics

Two hundred potential participants were screened for this study. Twenty-nine participants were excluded during the screening phase and 15 participants, who did not meet the diagnostic criteria for FM, were excluded following a medical interview with two specialized pain physicians. Of the 156 participants, 77 were FM patients and 79 were control participants. Control participants included 11 first degree relatives of participating FM patients, 20 household members of participating patients and 48 unrelated controls. Other than gender, marital status and occupational status, differences in demographic and anthropometric characteristics among groups were not statistically significant (Supplementary Table 1, available at http://links.lww.com/PAIN/A839). One patient in the first-degree relative control group was diagnosed with well-controlled rheumatoid arthritis, with no evidence of active arthritis or elevated inflammatory markers. All participants were recruited and all samples collected between October 2017 and June 2018. No seasonal variation was observed between groups when comparing sample collection season (ANOVA p=0.30, F=1.24).

FM patients were women (mean age 46±8 years), who on average had been diagnosed with FM 12 (SD 7.7) years earlier. The vast majority of participants were of Caucasian ethnicity. Mean widespread pain index in the patients' group was 11 (SD 3.2), and their mean symptom severity score was 9.3 (SD 1.8). The 2016 FM diagnostic criteria scores and FM Impact Questionnaire

scores were significantly higher in patients as compared with all control groups (Pillai's trace $p<0.0001$, $F=4.53$; Supplementary Table 2, available at http://links.lww.com/PAIN/A839). Participants' sleep quality was evaluated using the Insomnia Severity Index questionnaire. FM patients reported more difficulty falling asleep, maintaining sleep, and early awakening (Pillai's trace $p<0.0001$, $F=7.75$, see univariate analysis in Supplementary Table 3, available at http://links.lww.com/PAIN/A839). The frequency of depression and anxiety scores of 5/10 or higher among FM patients was 59% and 68% respectively, in line with previous reports [2,32].

Participants' physical activity was evaluated using the Physical Activity Self-administered Questionnaire. No consistent trend towards higher or lower level of activity was observed for any group (Supplementary Table 3, available at http://links.lww.com/PAIN/A839).

Dietary assessment was performed for all participants. Among participants with complete dietary reports, 16 patients and 3 unrelated controls were excluded for under-reporting. Energy and macronutrient intake (total and relative) were not different between patients and controls (Supplementary Table 4, available at http://links.lww.com/PAIN/A839). In addition, vitamin, mineral, different fatty acids, alcohol, caffeine, sugar and fiber intakes were not different among groups; nor was overall diet quality score (data not shown).

Overall gut microbiome composition

To explore the composition of the gut microbiome of FM patients and controls, we studied a cohort of 156 adults. Stool samples (77 FM patients, 79 controls) were collected and sequenced using both 16S rRNA and metagenome methods, and analyzed at multiple taxonomic levels. 16S rRNA sequencing produced 12,137,453 reads, with an average of 77,804±6,832 reads per sample.

A total of 1,620 operational taxonomic units (OTUs) were identified capturing 8,474,421 reads. These could be classified at various taxonomic levels: 8 at class, 25 at order, 179 at family, 716 at genus and 349 as species, with 343 OTUs remaining as unclassified or unknown (Figure 1). Predominant phyla across all participants included bacteroidetes (48% of total raw counts), firmicutes (40%), proteobacteria (4%) and actinobacteria (2%). From the 349 OTUs annotated at species taxonomic level, 312 were unique whereas 37 shared a sequence common to multiple species (Additional 16S and WGS data are provided in the supplementary file). The ten most abundant OTUs annotated at species resolution across all samples were from the Bacteroides genera, including *B. dorei*, *B. uniformis*, *B. stercoris*, *B. ovatus*, as well as *Prevotella copri*, *Alistipes putredinis* and *Faecalibacterium prausnitzii*.

To validate the taxa identified using 16S rRNA gene amplification, whole genome sequencing (WGS) data were directly mapped to the 16S rRNA gene OTU table as well as analysed using MetaPhlAn2. WGS generated 1,289,867,794 reads with an average of 8,268,383 reads per sample. All 1620 OTUs were successfully recapitulated by WGS reads at 100% identity. The median coverage was 99.1% (SD 4.2) with 1015 OTUs mapped uniquely and 605 sharing reads. Independent analysis using MetaPhlAn2 identified 196 species, including all but 3 species found using 16S rRNA gene amplification with the exception of those not yet present in the MetaPhlAn2 marker database (80 species, Figure 1H).

Gut microbiome composition is altered in FM patients

Shannon and inverse Simpson alpha-diversity indices showed non-significant differences in sample diversity between FM patients and household, relatives and unrelated control groups (ANOVA $p > 0.05$; Figure 1F). Bray-Curtis dissimilarity between pairs of participants showed a small but significant difference in mean dissimilarity metrics among FM-FM pairs as compared

to FM-control pairs ($p<0.0001$, $F=48.44$), or control-control pairs ($p<0.0001$, $F=56.71$, Figure 1E).

Canonical correspondence analysis suggests samples segregate by FM patients and unrelated controls in non-overlapping clusters (Figure 2A). Multivariate analysis showed between-group variance to be significantly greater than within-group variance ($p < 0.01$) between FM patients and unrelated controls. Household members and first-degree relative controls clustered between FM patients and unrelated controls (Figure 2B).

Differential abundance analysis was performed to compare FM patients (77) to unrelated control participants (48) using DESeq2. Seventy-two OTUs were identified as significantly differentially abundant (DA), 53 higher in FM and 18 higher in unrelated controls. These included 47 which could be annotated at either genera or species taxonomic levels and included taxa from firmicutes, bacteroidetes and proteobacteria. The largest positive fold-change difference in DA putative species observed in FM patients were from the OTUs Parabacteroides_merdae_4 and Clostridium_scindens_1; whereas the largest negative fold-change difference in FM patient were in Prevotella_copri_1 and Bacteroides_uniformis_3 (Figure 3).

When relative abundance of OTUs were correlated, DA species generally clustered according to their abundance in participant groups, whereby species found in higher abundance in FM patients clustered together while those found in higher abundance in controls clustered separately (Supplementary Figure 1A, available at http://links.lww.com/PAIN/A839). Among DA species, the majority of correlations were statistically significant ($p<0.1$), showing positive correlations between a majority of DA species within FM group and within unrelated control group (Pearson correlation>0) and significant inverse correlations of DA species between FM and unrelated control groups (Pearson correlation<0, Supplementary figure 1B, available at

http://links.lww.com/PAIN/A839). This observation serves as an independent support that DA taxa represent a valid difference between FM patients and controls.

To control for possible confounding effects of host-related and environmental variables, the amount of variance (r^2) explained by various possible confounders was evaluated using a general linear model. The highest amounts of microbiome variance were explained by the diagnosis of FM and disease-specific measures, while all other covariates, including diet, medications, co-morbidities and anthropometric variables had a smaller effect on the observed variance (See Supplementary Figure 2 and Supplementary Tables 5 and 6, available at http://links.lww.com/PAIN/A839). This suggests that at least some of the observed differences in microbiome composition are likely to be attributed to FM.

Altered serum levels of fermentation end products in FM patients

Consistent alterations in the abundance of butyrate-metabolism-related bacteria were observed: *F. prausnitzii* and *B. uniformis* were found in lower relative abundance in FM patients, while higher relative abundance was observed for *Intestinimonas butyriciproducens*, *Flavonifractor plautii*, *Butyricoccus desmolans*, *Eisenbergiella tayi* and *Eisenbergiella massiliensis*. To explore the possible metabolic effect of these alterations, a targeted metabolite approach was used to measure the serum concentrations of butyric acid, isobutyric acid, propionic acid and lactic acid. Serum levels of butyric acid in FM patients (n=73) were higher compared to unrelated controls (n=46, p=0.005), while levels of propionic acid were lower (p=0.006) and a trend towards lower levels of isobutyric acid was also observed (p=0.056). No significant differences in the serum levels of lactic acid was observed (Figure 4B). Multivariate analysis showed a significant between-group difference (Pillai's Trace, F=8.97, p<0.0001).

Gut microbiome is associated with clinical indices of FM

To explore the relationship between DA taxa abundance and symptoms severity, the relative abundance of DA OTUs was correlated against disease-specific and independent variables. Spearman's rank correlation coefficient demonstrated a statistically significant association between the abundance of several taxa and disease severity measures, including pain intensity, widespread pain index, dyscognition and fatigue (Benjamini-Hochberg FDR < 0.05, Figure 4A). The abundance of these taxa showed no consistent relationship with patients' demographic, anthropometric, or physical activity variables. Thus, abundance of some DA taxa could be correlated with FM symptom severity.

Microbiome-based machine-learning identification of FM patients

To explore the utility of gut microbiome composition in identifying FM patients, we utilised machine learning using LASSO (least absolute shrinkage and selection operator) and SVM (support vector machine) algorithms to define the OTUs most predictive of the diagnosis of FM (Figure 5A). The full and sparsity filtered count tables, as well as OTU subsets (DA and random) were tested to determine the most useful dataset for classification. The DA OTU dataset yielded the highest diagnostic accuracy. A 10-fold cross-validation (10,000 iterations) yielded 123 DA OTU sets with the highest diagnostic accuracy. LOOCV (leave-one-out cross-validation) identified a set of OTUs which provided the best overall prediction accuracy (AUC-ROC [Area Under the Curve – Receiver-Operating-Characteristics] of 87.2%; Figure 5B,C). Finally, this best predicting OTU set was used in a support vector machine (SVM) trained on 90% of participants. The SVM yielded a prediction accuracy of ROC-AUC 87.8% (Figure 5D) using cohorts of 10% of participants (100 iterations).

The superior utility of the DA OTU dataset reiterates the notion that DA taxa represent a true difference between FM patients and controls. Furthermore, these results suggest that the composition of the microbiome could be indicative of the diagnosis of FM.

Discussion

The overall population structure and diversity of the gut microbiome in patients with FM and a matched cohort of healthy controls was relatively similar. When explored at a higher resolution, however, gut microbiome composition showed significant alterations in FM patients. These differences, independently evident using multiple analytic approaches, revealed a distinct pattern of the fecal microbiome in FM. The diagnosis of FM and its clinical features (pain, fatigue, cognitive symptoms) explained more microbiome variance than did any other covariate, including demographic, anthropometric and dietary variables, as well as co-morbidities and medication consumption.

We observed a quantitative association between the abundance of several taxa and the severity of FM-related symptoms, including pain intensity, pain distribution, fatigue, sleep disturbances and cognitive symptoms. The abundance of these taxa correlated selectively with disease-related symptoms, but not with disease-independent variables.

Having established that gut microbiome is altered in FM patients, LASSO machine learning algorithm showed high prediction accuracy of patients from controls, based only on individual microbiome features [68]. Further study, including larger cohorts, greater ethnic sampling and cross-cohort assessment is required before a confident contribution can be made to improved FM diagnostics; however, the ability of our model to accurately identify FM patients in this study,

indicates that gut microbiome may have the potential to contribute to the clinically challenging diagnosis of FM.

When considering the nineteen specific species identified as significantly differentially abundant between FM patients and unrelated control participants, there was a broad range in how well-characterised these species were. Those species putatively depleted (lower in relative abundance) in FM were relatively well-characterised and included *Faecalibacterium prausnitzii, Bacteroides uniformis, Prevotella copri* and *Blautia faecis*. *Faecalibacterium prausnitzii* is one of the most abundant and well-studied butyrate producing bacteria in the human gut [46]. This species has been reported to be depleted in multiple intestinal diseases and was therefore suggested as a potential marker for gut disorders. Within the gut, *F. prausnitzii* has been reported to exert anti-nociceptive as well as anti-inflammatory effects [58,79] and to enhance the intestinal barrier function [46]. Similar to our observations, *F. prausnitzii* was also reported to be depleted in patients with CFS [61].

Bacteroides uniformis, is one of several species which have recently been reported as having altered relative abundance in patients with inflammatory arthritis, along with *H. parainfluenza, P. copri* and others [96]. *B. uniformis* and *H. parainfluenza* were detected in synovial tissues of osteoarthritic joints, whereas *P. copri* and *H. parainfluenza* were detected in rheumatoid arthritis synovial fluid. *P. copri* is thought to mediate inflammatory response via Th17 activation [37,43], and was also shown to induce arthritis in an animal model of arthritis-prone mice [51]. In this study, these species were found in *lower abundance* in FM patients. Although FM is often considered to be a rheumatologic disease, it seems that at least some species previously found at higher abundance in inflammatory rheumatic diseases are depleted in FM.

In contrast to the depletion of butyrate producers *F. prausnitzii* and *B. uniformis* in FM patients, we observed significant higher relative abundance of a number of other known intestinal butyrate producers: *Intestinimonas butyriciproducens*, *Flavonifractor plautii*, *Butyricoccus desmolans*, *Eisenbergiella tayi* and the recently identified *Eisenbergiella massiliensis* . Alterations in butyrate and propionate metabolizing species were further supported by alterations in serum levels of these short chain fatty acids. Coherent with this putative shift in the butyrate producing community of FM patients, *Parabacteroides merdae* was also significantly higher in relative abundance in FM patients. Recently, *P. merdae* has been reported by Olson et al. to be one of two key mediators of the anti-epileptic effect of the ketogenic diet [64]: in a mouse model, ketogenic diet can drive an increase in the abundance of *P. merdae*, which in turn, by regulation of amino acid γ-glutamylation leads to an increase in hippocampal γ-aminobutyric acid (GABA)/glutamate ratio. The increase in the inhibitory to excitatory neurotransmitter ratio in the brain is thought to protect against seizures. The second key mediator species of the ketogenic diet effect on seizures reported by Olson et al. was *Akkermansia muciniphila*, which in our cohort was also found at higher abundance in FM patients, although this observation did not reach statistical significance (p=0.042, Benjamini-Hochberg FDR=0.27).

Clostridium scindens and *B. desmolans*, two bacterial species capable of converting cortisol to androgens by 20α-hydroxysteroid dehydrogenase activity [73,8,81,60,19] were found in higher abundance in FM patients. Interestingly, abnormal regulation of hypothalamic-pituitary-adrenal (HPA) axis has been reported in FM patients, although the direction of dysregulation is controversial [21,80].

FM shares clinical features with several syndromes, including irritable bowel syndrome (IBS), chronic fatigue syndrome (CFS) and interstitial cystitis (IC). Specific alterations in gut

microbiota were reported in all of these syndromes (IBS [70,71,83,98], CFS [61], IC [10]). While several taxa, found here to be differentially abundant in FM, were also reported to be differentially abundant in other overlapping syndromes, other taxa seem to be unique to FM. As an example, *C. scindens* was found at higher abundance both in FM and in CFS. *H. parainfluenza* and *F. prausnitzii* were found at decreased abundance in FM and in CFS, and the latter was also reported to be depleted in IC patients [10]. Several genera, including *Bacteroides*, *Parabacteroides* and *Clostridium* were found differentially abundant both in IBS and in FM (independent of IBS) [71,83,98]. Nevertheless, a few taxa showed an inverse differential abundance pattern: e.g. *P. merdae*, significantly higher in FM, was reported to be depleted in CFS [61]. Furthermore, FM is associated with a high frequency of mental health comorbidity, mainly anxiety and depression affecting 40 to 80% of individuals [2,32]. Anxiety, depression and emotional stress have been associated with alterations in gut microbiome both in the general population [41] and in the context of specific other diagnoses. Recent research by Peter et al (2018) has revealed the potential importance of mental health as a factor in irritable bowel syndrome, suggesting that further cross-disciplinary research is needed to fully understand the microbiome differences identified here [70]. It thus appears that FM shares some microbiome characteristics with several overlapping syndromes, while retaining some unique features.

In this study effort was made to validate the diagnosis of each patient by a thorough medical evaluation by experienced pain specialists, thus excluding over 15% of interviewed patients. We consider the meticulous characterization of patient phenotypes an important strength of this study. A full dietary intake evaluation was completed for nearly 80% of participants and analyzed by a team of nutritionists. To the best of our knowledge, this type of in-depth evaluation is not routinely done in similar studies. Finally, each participant underwent a full

medical history, including documentation of their co-morbidities (with a specific evaluation for IBS), medication consumption, physical activity and sleep. These measures resulted in a well-balanced case-control cohort, almost identical in demographics, anthropometrics, co-morbidities, non-pain medication intake, and dietary intake. Furthermore, the extent of collected data allowed us to evaluate the effect of each variable on the observed variance in microbiome composition and to detect potential confounding factors. The cohort of 156 participants, provided sound statistical power. We have employed two sequencing methods – 16S rRNA gene amplification and whole genome sequencing metagenomics analyses – each with its own advantages and drawbacks, to allow for an independent validation of the microbiome composition. The combination of large cohort size with high depth of coverage allowed for the identification of low-prevalence differentially abundant species, such as *I. butyriciproducens*. Targeted serum metabolomics allowed an independent support of some of the alterations observed in the gut microbiome. Results of this study should be considered in light of several limitations: first, the symptoms of FM include chronic pain, fatigue and other somatic symptoms, precluding the attribution of specific microbiome alteration to a specific clinical presentation. Second, over 90% of participants self-reported as Caucasians, possibly not reflecting microbiome patterns in other ethnic groups. Third, participants' diagnoses were recorded based on a systematic clinical interview which could potentially lead to under-diagnosis. Mental health conditions may be underrepresented in FM patients (Supplementary Table 5, available at http://links.lww.com/PAIN/A839) as compared to self-reported symptoms (Supplementary Table 2, participants were not evaluated by a mental health care specialist, available at http://links.lww.com/PAIN/A839). Finally, this study was designed to find an *association*

between microbiome composition and FM, not allowing direct conclusions regarding *causality* or mechanisms of action to be drawn.

The results reported here are, to the best of our knowledge, the first to demonstrate gastrointestinal tract microbiota alteration in patients with somatic (non-visceral) pain in general, and FM in particular. Some of the differentially abundant taxa reported here are known to be involved in metabolic pathways whose effect on symptoms of FM is biologically plausible. Moreover, there seems to be a quantitative correlation between the abundance of several taxa and symptoms severity. Finally, differentially abundant taxa in FM patients share some commonalities with overlapping dysfunctional syndromes - IBS, CFS, IC – while retaining some unique differentially abundant taxa.

Results of this study provide evidence for alterations of gut microbiome alterations in FM. Further studies are needed to measure possible microbiome changes in other chronic pain conditions, and to explore potential causal correlations between the gut microbiome and FM. Our findings may offer important opportunities to improve our understanding of FM, aid in its diagnosis and perhaps outline future therapeutic modalities. First, exploring the mechanisms by which the gut microbiota may affect nociception could offer new insights into the pathogenesis and pathophysiology of FM, and possibly suggest leads to explore the possible effects of trigger events, personal susceptibility and individual prognosis. Second, the accuracy of microbiome-based machine learning classification suggests that the microbiome holds the potential to aid in the diagnosis of FM in the future. Finally, if a causal relationship between gut microbiome alterations and FM is established, the way may be paved for the development of new treatment modalities targeting this fascinating community of intestinal microorganisms.

Acknowledgments

The authors thankfully acknowledge the generous financial support of the Louise and Alan Edwards Foundation. We thank the Israeli Society for Musculoskeletal Medicine for their financial support. We would like to thank Mrs. Sylvie Toupin, Mrs. Linda Ferguson, Dr. Pierre Lepage (Génome Québec), Dr. Rupa Mandal (TMIC), Ms. Audrey Moyen and Mr. Stéphan Attara for their invaluable help in the realization of this project. We are grateful to L'association de la Fibromyalgie de Montréal and to all participants for their partnership. We thank David Anderson for advising on figure design. DNA sequencing was performed by Génome Québec, Montreal. Metabolite analysis was performed by the metabolomic innovation centre (TMIC), Edmonton, Alberta. The Canadian Center for Computational Genomics (C3G) is a Genomics Technology Platform (GTP) supported by the Canadian Government through Genome Canada.

Author contributions

AM and YS conceived, designed and supervised the study. YS, AM, MAF, KD and SC were involved in the design of the study. AM and MAF interviewed and evaluated FM patients. AA performed the nutritional evaluations. AM collected all biological samples. Genomic data was analysed by EG, NJB and AM. Dietary analysis was done by SC and AA. Metabolomic data was analysed by AM. KD, SC, EG, NJB and AM were involved in data analysis, evaluation and integration. The manuscript was written by AM, YS, NJB, EG and SC.

Competing interests

The authors declare no competing interests.

Supplemental video content

A video abstract associated with this article can be found at http://links.lww.com/PAIN/A840.

References

[1] Ahn S, Jin T-E, Chang D-H, Rhee M-S, Kim HJ, Lee SJ, Park D-S, Kim B-C. Agathobaculum butyriciproducens gen. nov. sp. nov., a strict anaerobic, butyrate-producing gut bacterium isolated from human faeces and reclassification of Eubacterium desmolans as Agathobaculum desmolans comb. nov. Int J Syst Evol Microbiol 2016;66:3656–3661.

[2] Alciati A, Sgiarovello P, Atzeni F, Sarzi-Puttini P. Psychiatric problems in fibromyalgia: clinical and neurobiological links between mood disorders and fibromyalgia. Reumatismo 2012;64:268–274.

[3] Amir I, Bouvet P, Legeay C, Gophna U, Weinberger A. Eisenbergiella tayi gen. nov., sp. nov., isolated from human blood. Int J Syst Evol Microbiol 2014;64:907–914.

[4] Anders S, McCarthy DJ, Chen Y, Okoniewski M, Smyth GK, Huber W, Robinson MD. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. Nat Protoc 2013;8:1765–1786.

[5] Arora HC, Eng C, Shoskes DA. Gut microbiome and chronic prostatitis/chronic pelvic pain syndrome. Ann Transl Med 2017;5. doi:10.21037/atm.2016.12.32.

[6] Attar N. Microbiome: Good for the gut, good for the brain. Nat Rev Microbiol 2016;14:269.

[7] Bauer KC, Huus KE, Finlay BB. Microbes and the mind: emerging hallmarks of the gut microbiota-brain axis. Cell Microbiol 2016;18:632–644.

[8] Bloem LM, Storbeck K-H, Schloms L, Swart AC. 11β-hydroxyandrostenedione returns to the steroid arena: biosynthesis, metabolism and function. Mol Basel Switz 2013;18:13228–13244.

[9] Bollepalli S, Korhonen T, Kaprio J, Ollikainen M, Anders S. EpiSmokEr: A robust classifier to determine smoking status from DNA methylation data. bioRxiv 2018:487975.

[10] Braundmeier-Fleming A, Russell NT, Yang W, Nas MY, Yaggie RE, Berry M, Bachrach L, Flury SC, Marko DS, Bushell CB, Welge ME, White BA, Schaeffer AJ, Klumpp DJ. Stool-based biomarkers of interstitial cystitis/bladder pain syndrome. Sci Rep 2016;6:26083.

[11] Bui TPN, Shetty SA, Lagkouvardos I, Ritari J, Chamlagain B, Douillard FP, Paulin L, Piironen V, Clavel T, Plugge CM, de Vos WM. Comparative genomics and physiology of the butyrate-producing bacterium Intestinimonas butyriciproducens. Environ Microbiol Rep 2016;8:1024–1037.

[12] Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. J Rheumatol 1991;18:728–733.

[13] Carlier J-P, Bedora-Faure M, K'ouas G, Alauzet C, Mory F. Proposal to unify Clostridium orbiscindens Winter et al. 1991 and Eubacterium plautii (Séguin 1928) Hofstad and Aasjord 1982, with description of Flavonifractor plautii gen. nov., comb. nov., and reassignment of Bacteroides capillosus to Pseudoflavonifractor capillosus gen. nov., comb. nov. Int J Syst Evol Microbiol 2010;60:585–590.

[14] Chahoud M, Chahine R, Salameh P, Sauleau EA. Reliability, factor analysis and internal

consistency calculation of the Insomnia Severity Index (ISI) in French and in English among Lebanese adolescents. eNeurologicalSci 2017;7:9–14.

[15] Clauw DJ. Fibromyalgia: a clinical review. JAMA 2014;311:1547–1555.

[16] Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nat Rev Neurosci 2012;13:701–712.

[17] De Palma G, Lynch MDJ, Lu J, Dang VT, Deng Y, Jury J, Umeh G, Miranda PM, Pigrau Pastor M, Sidani S, Pinto-Sanchez MI, Philip V, McLean PG, Hagelsieb M-G, Surette MG, Bergonzelli GE, Verdu EF, Britz-McKibbin P, Neufeld JD, Collins SM, Bercik P. Transplantation of fecal microbiota from patients with irritable bowel syndrome alters gut function and behavior in recipient mice. Sci Transl Med 2017;9.

[18] Deng L, Chang D, Foshaug RR, Eisner R, Tso VK, Wishart DS, Fedorak RN. Development and Validation of a High-Throughput Mass Spectrometry Based Urine Metabolomic Test for the Detection of Colonic Adenomatous Polyps. Metabolites 2017;7.

[19] Devendran S, Méndez-García C, Ridlon JM. Identification and characterization of a 20β-HSDH from the anaerobic gut bacterium Butyricicoccus desmolans ATCC 43058. J Lipid Res 2017;58:916–925.

[20] Distrutti E, Monaldi L, Ricci P, Fiorucci S. Gut microbiota role in irritable bowel syndrome: New therapeutic strategies. World J Gastroenterol 2016;22:2219–2241.

[21] Eller-Smith OC, Nicol AL, Christianson JA. Potential Mechanisms Underlying Centralized Pain and Emerging Therapeutic Interventions. Front Cell Neurosci 2018;12:35.

[22] Felice VD, Quigley EM, Sullivan AM, O'Keeffe GW, O'Mahony SM. Microbiota-gut-brain signalling in Parkinson's disease: Implications for non-motor symptoms. Parkinsonism

Relat Disord 2016.

[23] Fitzcharles M-A, Boulos P. Inaccuracy in the diagnosis of fibromyalgia syndrome: analysis of referrals. Rheumatol Oxf Engl 2003;42:263–267.

[24] Fitzcharles M-A, Ste-Marie PA, Panopalis P, Ménard H, Shir Y, Wolfe F. The 2010 American college of rheumatology fibromyalgia survey diagnostic criteria and symptom severity scale is a valid and reliable tool in a French speaking fibromyalgia cohort. BMC Musculoskelet Disord 2012;13:179.

[25] Forsythe P, Kunze W, Bienenstock J. Moody microbes or fecal phrenology: what do we know about the microbiota-gut-brain axis? BMC Med 2016;14:58.

[26] Frémont M, Coomans D, Massart S, De Meirleir K. High-throughput 16S rRNA gene sequencing reveals alterations of intestinal microbiota in myalgic encephalomyelitis/chronic fatigue syndrome patients. Anaerobe 2013;22:50–56.

[27] Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw 2010;33:1–22.

[28] Giloteaux L, Goodrich JK, Walters WA, Levine SM, Ley RE, Hanson MR. Reduced diversity and altered composition of the gut microbiome in individuals with myalgic encephalomyelitis/chronic fatigue syndrome. Microbiome 2016;4:30.

[29] Goebel A, Buhner S, Schedel R, Lochs H, Sprotte G. Altered intestinal permeability in patients with primary fibromyalgia and in patients with complex regional pain syndrome. Rheumatol Oxf Engl 2008;47:1223–1227.

[30] Gonzalez E, Pitre FE, Pagé AP, Marleau J, Guidi Nissim W, St-Arnaud M, Labrecque M, Joly S, Yergeau E, Brereton NJB. Trees, fungi and bacteria: tripartite metatranscriptomics of a

root microbiome responding to soil contamination. Microbiome 2018;6:53.

[31] Häuser W, Ablin J, Fitzcharles M-A, Littlejohn G, Luciano JV, Usui C, Walitt B. Fibromyalgia. Nat Rev Dis Primer 2015;1:15022.

[32] Häuser W, Fitzcharles M-A. Facts and myths pertaining to fibromyalgia. Dialogues Clin Neurosci 2018;20:53–62.

[33] Hernandez CJ. The Microbiome and Bone and Joint Disease. Curr Rheumatol Rep 2017;19:77.

[34] Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature 2012;486:207–214.

[35] Jaggi M. An Equivalence between the Lasso and Support Vector Machines. ArXiv13031152 Cs Stat 2013. Available: http://arxiv.org/abs/1303.1152. Accessed 20 Feb 2019.

[36] Kang M, Mischel RA, Bhave S, Komla E, Cho A, Huang C, Dewey WL, Akbarali HI. The effect of gut microbiome on tolerance to morphine mediated antinociception in mice. Sci Rep 2017;7:42658.

[37] Kim D, Kim W-U. Editorial: Can Prevotella copri Be a Causative Pathobiont in Rheumatoid Arthritis? Arthritis Rheumatol Hoboken NJ 2016;68:2565–2567.

[38] Kläring K, Hanske L, Bui N, Charrier C, Blaut M, Haller D, Plugge CM, Clavel T. Intestinimonas butyriciproducens gen. nov., sp. nov., a butyrate-producing bacterium from the mouse intestine. Int J Syst Evol Microbiol 2013;63:4606–4612.

[39] Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation

sequencing-based diversity studies. Nucleic Acids Res 2013;41:e1.

[40] Krebs-Smith SM, Pannucci TE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, Wilson MM, Reedy J. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet 2018;118:1591–1602.

[41] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–359.

[42] Larsen JM. The immune response to Prevotella bacteria in chronic inflammatory disease. Immunology 2017;151:363–374.

[43] Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res 2012;40:W622-627.

[44] Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. Functional bowel disorders. Gastroenterology 2006;130:1480–1491.

[45] Lopez-Siles M, Duncan SH, Garcia-Gil LJ, Martinez-Medina M. Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. ISME J 2017;11:841–852.

[46] Love MI, Anders S, Kim V, Huber W. RNA-Seq workflow: gene-level exploratory analysis and differential expression. F1000Research 2015;4:1070.

[47] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15. doi:10.1186/s13059-014-0550-8.

[48] Luczynski P, Tramullas M, Viola M, Shanahan F, Clarke G, O'Mahony S, Dinan TG, Cryan JF. Microbiota regulates visceral pain in the mouse. eLife 2017;6.

[49] Lynch SV, Pedersen O. The Human Intestinal Microbiome in Health and Disease. N Engl J Med 2016;375:2369–2379.

[50] Maeda Y, Kurakawa T, Umemoto E, Motooka D, Ito Y, Gotoh K, Hirota K, Matsushita M, Furuta Y, Narazaki M, Sakaguchi N, Kayama H, Nakamura S, Iida T, Saeki Y, Kumanogoh A, Sakaguchi S, Takeda K. Dysbiosis Contributes to Arthritis Development via Activation of Autoreactive T Cells in the Intestine. Arthritis Rheumatol Hoboken NJ 2016;68:2646–2661.

[51] Malatji BG, Meyer H, Mason S, Engelke UFH, Wevers RA, van Reenen M, Reinecke CJ. A diagnostic biomarker profile for fibromyalgia syndrome based on an NMR metabolomics study of selected patients and controls. BMC Neurol 2017;17:88.

[52] Mansfield KE, Sim J, Croft P, Jordan KP. Identifying patients with chronic widespread pain in primary care. Pain 2017;158:110–119.

[53] Mayer EA, Tillisch K, Gupta A. Gut/brain axis and the microbiota. J Clin Invest 2015;125:926–938.

[54] McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PloS One 2013;8:e61217.

[55] Mearin F, Lacy BE, Chang L, Chey WD, Lembo AJ, Simren M, Spiller R. Bowel Disorders. Gastroenterology 2016.

[56] Mifflin MD, St Jeor ST, Hill LA, Scott BJ, Daugherty SA, Koh YO. A new predictive equation for resting energy expenditure in healthy individuals. Am J Clin Nutr 1990;51:241–247.

[57] Miquel S, Martín R, Lashermes A, Gillet M, Meleine M, Gelot A, Eschalier A, Ardid D, Bermúdez-Humarán LG, Sokol H, Thomas M, Theodorou V, Langella P, Carvalho FA. Anti-nociceptive effect of Faecalibacterium prausnitzii in non-inflammatory IBS-like models. Sci Rep

2016;6. doi:10.1038/srep19399.

[58] Morin CM. Insomnia: Psychological assessment and management. New York, NY, US: Guilford Press, 1993 p.

[59] Morris DJ, Ridlon JM. Glucocorticoids and gut bacteria: “The GALF Hypothesis” in the metagenomic era. Steroids 2017;125:1–13.

[60] Nagy-Szakal D, Williams BL, Mishra N, Che X, Lee B, Bateman L, Klimas NG, Komaroff AL, Levine S, Montoya JG, Peterson DL, Ramanan D, Jain K, Eddy ML, Hornig M, Lipkin WI. Fecal metagenomic profiles in subgroups of patients with myalgic encephalomyelitis/chronic fatigue syndrome. Microbiome Lond 2017;5. doi:http://dx.doi.org/10.1186/s40168-017-0261-y.

[61] Newberry F, Hsieh S-Y, Wileman T, Carding SR. Does the microbiome and virome contribute to myalgic encephalomyelitis/chronic fatigue syndrome? Clin Sci Lond Engl 1979 2018;132:523–542.

[62] Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. vegan: Community Ecology Package. 2019 p. Available: https://CRAN.R-project.org/package=vegan. Accessed 7 Feb 2019.

[63] Olson CA, Vuong HE, Yano JM, Liang QY, Nusbaum DJ, Hsiao EY. The Gut Microbiota Mediates the Anti-Seizure Effects of the Ketogenic Diet. Cell 2018;173:1728-1741.e13.

[64] O’Mahony SM, Felice VD, Nally K, Savignac HM, Claesson MJ, Scully P, Woznicki J, Hyland NP, Shanahan F, Quigley EM, Marchesi JR, O’Toole PW, Dinan TG, Cryan JF.

Disturbance of the gut microbiota in early-life selectively affects visceral pain in adulthood without impacting cognitive or anxiety-related behaviors in male rats. Neuroscience 2014;277:885–901.

[65] O' Mahony SM, Dinan TG, Cryan JF. The gut microbiota as a key regulator of visceral pain. Pain 2017;158 Suppl 1:S19–S28.

[66] Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, Baer DJ, Midthune D, Troiano RP, Bowles H, Subar AF. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr 2018;107:80–93.

[67] Pasolli E, Truong DT, Malik F, Waldron L, Segata N. Machine Learning Meta-analysis of Large Metagenomic Datasets: Tools and Biological Insights. PLoS Comput Biol 2016;12:e1004977.

[68] Perrot S, Dumont D, Guillemin F, Pouchot J, Coste J, French Group for Quality of Life Research. Quality of life in women with fibromyalgia syndrome: validation of the QIF, the French version of the fibromyalgia impact questionnaire. J Rheumatol 2003;30:1054–1059.

[69] Peter J, Fournier C, Durdevic M, Knoblich L, Keip B, Dejaco C, Trauner M, Moser G. A Microbial Signature of Psychological Distress in Irritable Bowel Syndrome. Psychosom Med 2018;80:698–709.

[70] Pozuelo M, Panda S, Santiago A, Mendez S, Accarino A, Santos J, Guarner F, Azpiroz F, Manichanh C. Reduction of butyrate- and methane-producing microorganisms in patients with Irritable Bowel Syndrome. Sci Rep 2015;5:12693.

[71] Rahman A, Underwood M, Carnes D. Fibromyalgia. BMJ 2014;348:g1224.

[72] Ridlon JM, Ikegawa S, Alves JMP, Zhou B, Kobayashi A, Iida T, Mitamura K, Tanabe G, Serrano M, De Guzman A, Cooper P, Buck GA, Hylemon PB. Clostridium scindens: a human gut microbe with a high potential to convert glucocorticoids into androgens. J Lipid Res 2013;54:2437–2449.

[73] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.

[74] Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 2009;75:7537–7541.

[75] Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods 2012;9:811–814.

[76] Shen S, Lim G, You Z, Ding W, Huang P, Ran C, Doheny J, Caravan P, Tate S, Hu K, Kim H, McCabe M, Huang B, Xie Z, Kwon D, Chen L, Mao J. Gut microbiota is critical for the induction of chemotherapy-induced pain. Nat Neurosci 2017.

[77] Shoskes DA, Wang H, Polackwich AS, Tucky B, Altemus J, Eng C. Analysis of Gut Microbiome Reveals Significant Differences between Men with Chronic Prostatitis/Chronic Pelvic Pain Syndrome and Controls. J Urol 2016;196:435–441.

[78] Sitkin S, Pokrotnieks J. Clinical Potential of Anti-inflammatory Effects of Faecalibacterium prausnitzii and Butyrate in Inflammatory Bowel Disease. Inflamm Bowel Dis

2018.

[79] Sluka KA, Clauw DJ. Neurobiology of fibromyalgia and chronic widespread pain. Neuroscience 2016.

[80] Swart AC, Storbeck K-H. 11β-Hydroxyandrostenedione: Downstream metabolism by 11βHSD, 17βHSD and SRD5A produces novel substrates in familiar pathways. Mol Cell Endocrinol 2015;408:114–123.

[81] Takada T, Watanabe K, Makino H, Kushiro A. Reclassification of Eubacterium desmolans as Butyricicoccus desmolans comb. nov., and description of Butyricicoccus faecihominis sp. nov., a butyrate-producing bacterium from human faeces. Int J Syst Evol Microbiol 2016;66:4125–4131.

[82] Tap J, Derrien M, Törnblom H, Brazeilles R, Cools-Portier S, Doré J, Störsrud S, Le Nevé B, Öhman L, Simrén M. Identification of an Intestinal Microbiota Signature Associated With Severity of Irritable Bowel Syndrome. Gastroenterology 2017;152:111-123.e8.

[83] Theodorou V, Ait Belgnaoui A, Agostini S, Eutamene H. Effect of commensals and probiotics on visceral sensitivity and pain in irritable bowel syndrome. Gut Microbes 2014;5:430–436.

[84] Thorsen J, Brejnrod A, Mortensen M, Rasmussen MA, Stokholm J, Al-Soud WA, Sørensen S, Bisgaard H, Waage J. Large-scale benchmarking reveals false discoveries and count transformation sensitivity in 16S rRNA gene amplicon data analysis methods used in microbiome studies. Microbiome 2016;4:62.

[85] Tibshirani R. Regression Shrinkage and Selection via the Lasso. J R Stat Soc Ser B Methodol 1996;58:267–288.

[86] Timon CM, van den Barg R, Blain RJ, Kehoe L, Evans K, Walton J, Flynn A, Gibney ER. A review of the design and validation of web- and computer-based 24-h dietary recall tools. Nutr Res Rev 2016;29:268–280.

[87] Togo AH, Khelaifia S, Bittar F, Maraninchi M, Raoult D, Million M. “Eisenbergiella massiliensis”, a new species isolated from human stool collected after bariatric surgery. New Microbes New Infect 2016;13:15–16.

[88] Vol S, Bedouet M, Gusto G, Leglu C, Beslin E, Decou P, Nègre E, Planage B, Chazelle E, Mercier F, Lantieri O, Tichet J. Evaluating physical activity: The AQAP questionnaire and its interpretation software. Ann Phys Rehabil Med 2011;54:478–495.

[89] Walitt B, Katz RS, Bergman MJ, Wolfe F. Three-Quarters of Persons in the US Population Reporting a Clinical Diagnosis of Fibromyalgia Do Not Satisfy Fibromyalgia Criteria: The 2012 National Health Interview Survey. PLoS ONE 2016;11. doi:10.1371/journal.pone.0157235.

[90] Wei T, Simko V, Levy M, Xie Y, Jin Y, Zemla J. corrplot: Visualization of a Correlation Matrix. 2017 p. Available: https://CRAN.R-project.org/package=corrplot. Accessed 7 Feb 2019.

[91] Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, Hyde ER, Knight R. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 2017;5:27.

[92] Wen C, Zheng Z, Shao T, Liu L, Xie Z, Le Chatelier E, He Z, Zhong W, Fan Y, Zhang L, Li H, Wu C, Hu C, Xu Q, Zhou J, Cai S, Wang D, Huang Y, Breban M, Qin N, Ehrlich SD. Quantitative metagenomics reveals unique gut microbiome biomarkers in ankylosing spondylitis. Genome Biol 2017;18:142.

[93] Wolfe F, Clauw DJ, Fitzcharles M-A, Goldenberg DL, Häuser W, Katz RL, Mease PJ, Russell AS, Russell IJ, Walitt B. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–329.

[94] Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D, Wu X, Li J, Tang L, Li Y, Lan Z, Chen B, Li Y, Zhong H, Xie H, Jie Z, Chen W, Tang S, Xu X, Wang X, Cai X, Liu S, Xia Y, Li J, Qiao X, Al-Aama JY, Chen H, Wang L, Wu Q-J, Zhang F, Zheng W, Li Y, Zhang M, Luo G, Xue W, Xiao L, Li J, Chen W, Xu X, Yin Y, Yang H, Wang J, Kristiansen K, Liu L, Li T, Huang Q, Li Y, Wang J. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. Nat Med 2015;21:895–905.

[95] Zhao Y, Chen B, Li S, Yang L, Zhu D, Wang Y, Wang H, Wang T, Shi B, Gai Z, Yang J, Heng X, Yang J, Zhang L. Detection and characterization of bacterial nucleic acids in culture-negative synovial tissue and fluid samples from rheumatoid arthritis or osteoarthritis patients. Sci Rep 2018;8:14305.

[96] Zhou X-Y, Li M, Li X, Long X, Zuo X-L, Hou X-H, Cong Y-Z, Li Y-Q. Visceral hypersensitive rats share common dysbiosis features with irritable bowel syndrome patients. World J Gastroenterol 2016;22:5211–5227.

[97] Zhuang X, Xiong L, Li L, Li M, Chen M. Alterations of gut microbiota in patients with irritable bowel syndrome: A systematic review and meta-analysis. J Gastroenterol Hepatol 2017;32:28–38.

[98] Zuo Y, Cui Y, Yu G, Li R, Ressom HW. Incorporating prior biological knowledge for network-based differential gene expression analysis using differentially weighted graphical LASSO. BMC Bioinformatics 2017;18:99.

Figure legends

Figure 1: Overall microbiome composition of 156 study participants. **a.** Flower diagram of 1620 OTUs colour coded by phyla. An interactive version is available on http://www.computationalgenomics.ca/pub/fibromyalgia_microbiome/flower_diagram/ ; **b.** Phyla-level taxonomic composition of the gut microbiome among FM patients (n=77), relative controls (RC, n=11), household members (HC, n=20) and unrelated controls (UC, n=48); **c.** distribution of OTUs by taxonomic level of identification; **d.** distribution of relative abundance by the level of taxonomic identification; **e.** comparison of Bray-Curtis dissimilarity between all pairs of 156 participants, demonstrating that pair-wise distances between FM patients (FM-FM, 1482 pairs) are smaller as compared to distances between pairs of patients-controls (FM-C, 6083 pairs) or controls-controls (C-C, 1560); **f.** comparison of alpha diversity indices for each participant group; **g.** Study design flowchart. **h.** comparison of species identification using 16S rRNA genes (ANCHOR; yellow) and WGS (Metaphlan2; blue) technologies.

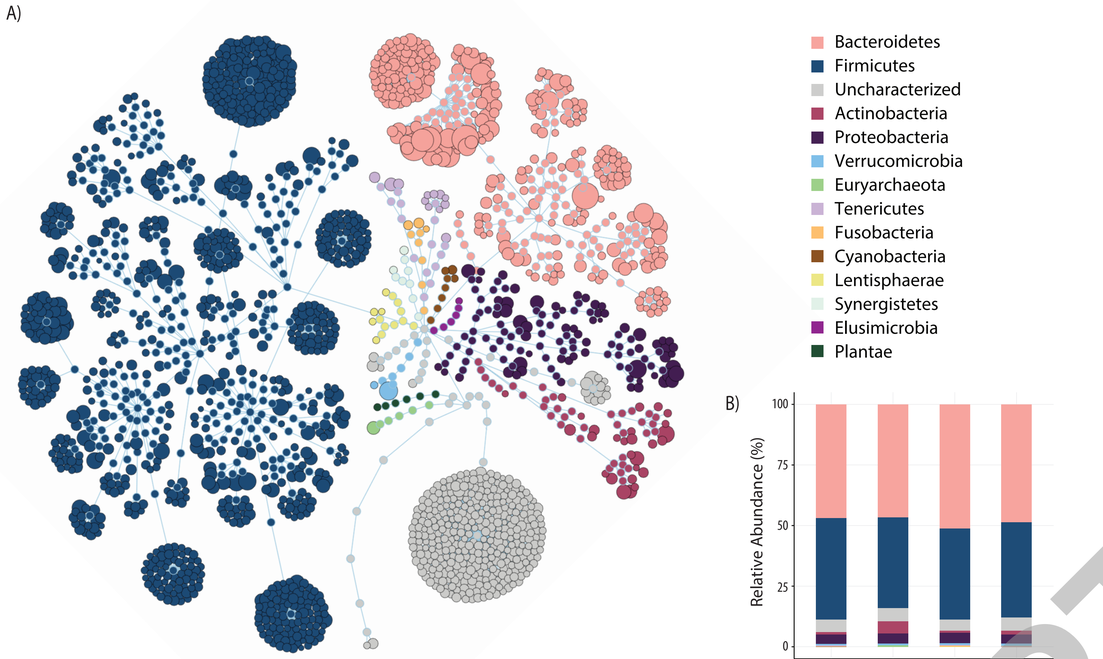
Figure 2: a. Canonical component analysis (x axis) vs. principal component analysis (y axis) of taxa abundance (raw counts) in FM patients vs. unrelated controls (UC). Amount of explained variance is provided ($p<0.001$, ANOVA) **b.** relatives (RC) and household controls (HC) cluster mid-way between FM patients and unrelated controls.

Figure 3: a. Differentially abundant OTUs between the FM and unrelated control groups. Fold change (FC log2) denotes relative differences in relative abundance between groups (DESeq2).

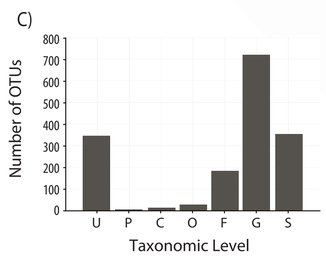
Species are grouped by phylum per comparison; **b.** a diagram of species-level differentially abundant bacteria in FM patients (left) and in healthy controls (right).

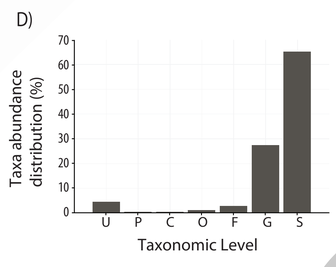
Figure 4: a. Serum concentrations of butyric acid, isobutyric acid, propionic acid and lactic acid (ANOVA ** $p<0.01$, ˙$p<0.1$).. **b.** Abundance of some DA OTUs is associated with clinical indices: Heat map of a univariate Spearman correlation matrix between DA taxa abundance (log_2; x-axis) vs. covariates (y-axis): demographics and anthropometrics; disease severity metrics (FMDC: 2016 FM diagnostic criteria); quality of life scores (FIQ); physical activity measures (PA); and sleep quality (sleep) scores. Heatmap is sorted base on a hierarchical clustering of DA OTUs. Blue shades indicate positive correlations while red shades indicate negative correlations (-0.5<rho<0.5). Statistically significant correlations are marked by a black dot (Benjamini-Hochberg FDR<0.05). Note the selective significant correlation of taxa abundance with symptom severity but not with other independent variables.

Figure 5: Machine learning classifier algorithm: **a.** Two-step LASSO[95] workflow adapted from Bollepalli et al.[9]. OTUs which resulted in the best prediction accuracy in LASSO were used in a support vector machine (SVM). **b.** Species-level annotated OTUs most often selected by LASSO; **c.** ROC curve for 1,000 iterations of LASSO-LOOCV feature prioritisation using 123 previously selected OTU sets; **d.** ROC curve for 100 iterations of SVM validation test using most important group of OTUs (group_103). The Model was trained on 90% of the initial number of samples and tested on the remaining. CV: cross-validation, LOOCV: leave-one-out cross-validation, AUC: area under curve, ROC: receiver operating characteristic.

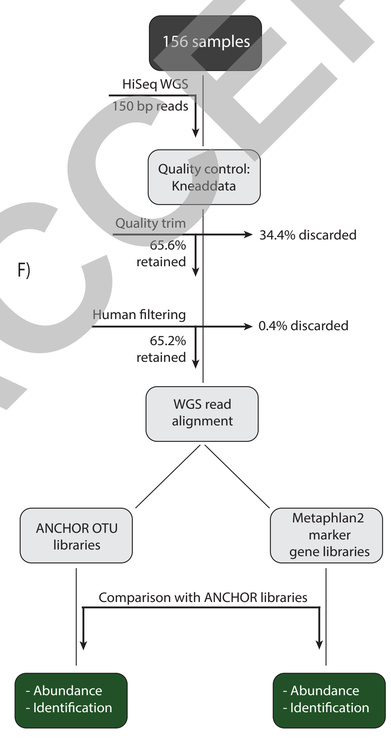

A)
Bacteroidetes
Firmicutes
Uncharacterized
Actinobacteria
Proteobacteria
Verrucomicrobia
Euryarchaeota
Tenericutes
Fusobacteria
Cyanobacteria
Lentisphaerae
Synergistetes
Elusimicrobia
Plantae

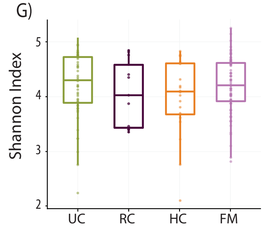
B)
Relative Abundance (%)
100
75
50
25
0
UC
RC
HC
FM

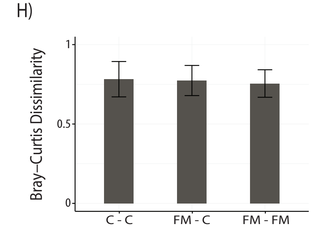

C)
Number of OTUs
800
700
600
500
400
300
200
100
0
U
P
C
O
F
G
S
Taxonomic Level


D)
Taxa abundance distribution (%)
70
60
50
40
30
20
10
0
U
P
C
O
F
G
S
Taxonomic Level

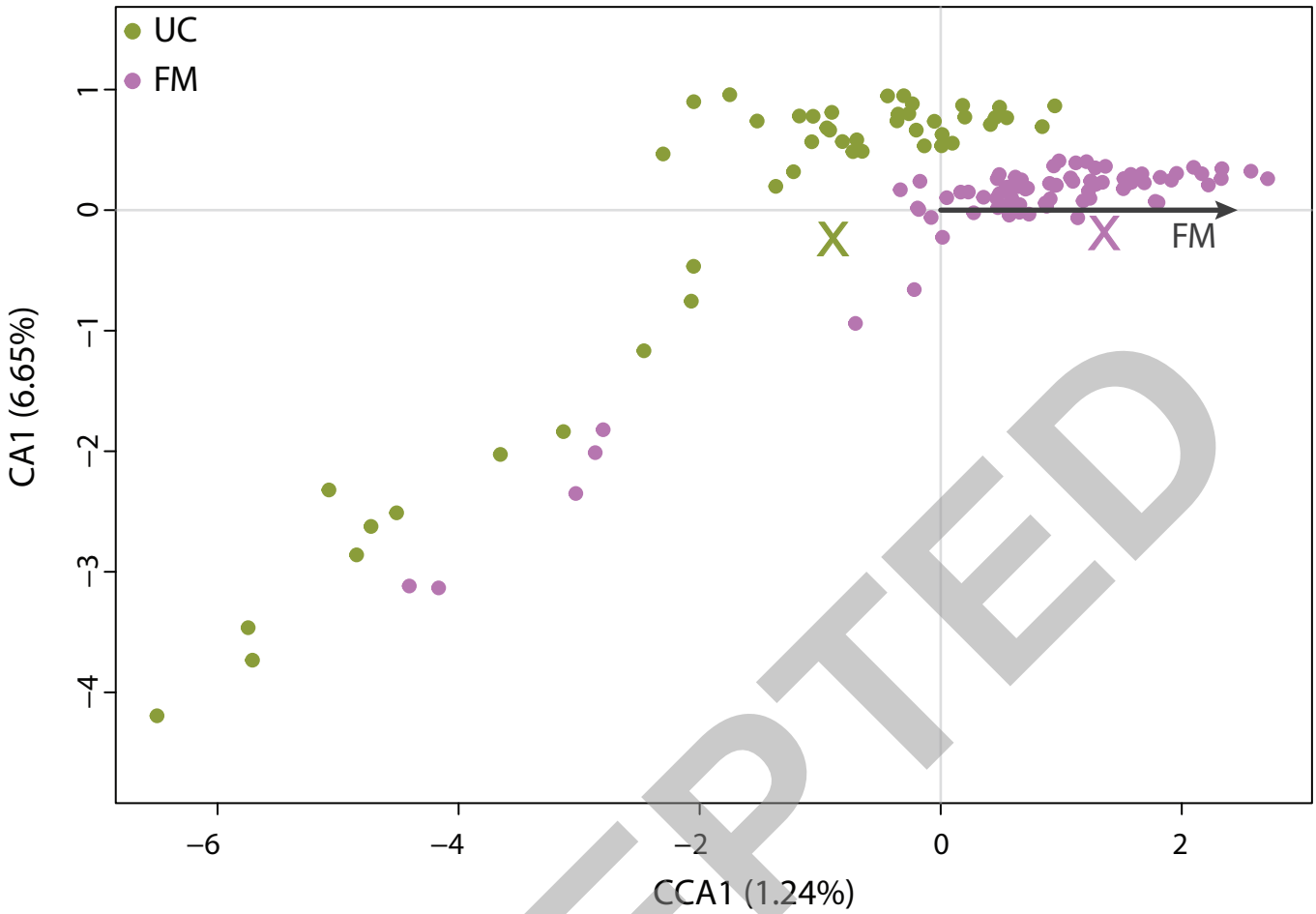

E)
16S rRNA gene (ANCHOR)
83*
100
111
WGS (Metaphlan2)


F)
156 samples
HiSeq WGS
150 bp reads
Quality control: Kneaddata
Quality trim
65.6% retained
34.4% discarded
Human filtering
65.2% retained
0.4% discarded
WGS read alignment
ANCHOR OTU libraries
Metaphlan2 marker gene libraries
Comparison with ANCHOR libraries
- Abundance
- Identification
- Abundance
- Identification

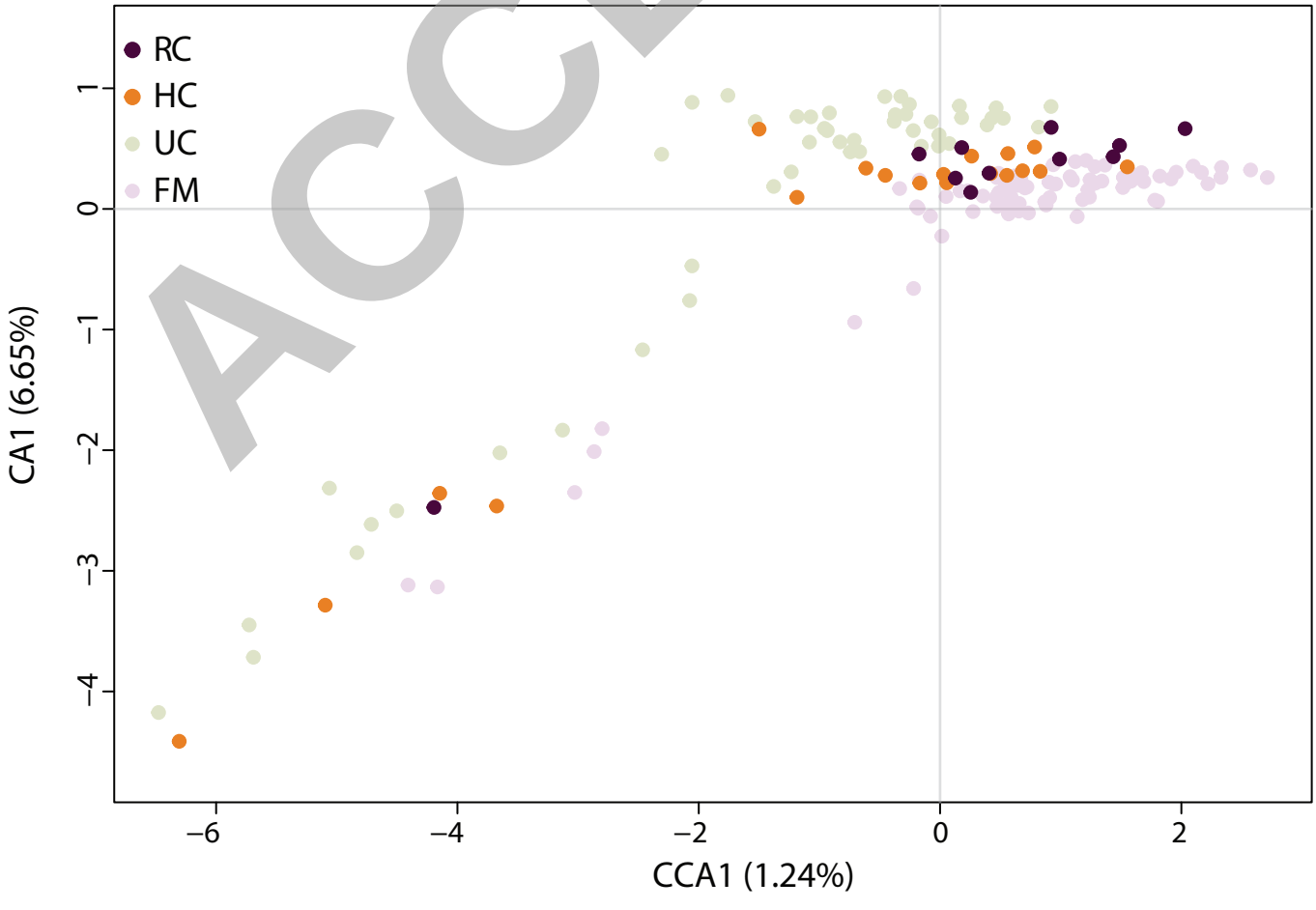

G)
Shannon Index
5
4
3
2
UC
RC
HC
FM


H)
Bray–Curtis Dissimilarity
1
0.5
0
C - C
FM - C
FM - FM

A)


UC
FM
X
X
FM
CA1 (6.65%)
CCA1 (1.24%)
1
0
−1
−2
−3
−4
−6
−4
−2
0
2

B)


RC
HC
UC
FM
CA1 (6.65%)
CCA1 (1.24%)
1
0
−1
−2
−3
−4
−6
−4
−2
0
2

A)

B)

Differentially abundant OTUs

Coriobacteriaceae_4
Parabacteroides_merdae_4
Parabacteroides_6
Parabacteroides_9
Bacteroides_55
Bacteroides_48
Alistipes_onderdonkii_1
Bacteroides_17
Bacteroides_65
Alistipes_5
Prevotella_copri_1
Bacteroides_uniformis_3
Clostridiales_MG_83
Clostridium_2
Acetivibrio_1
Clostridiales_MG_141
Clostridiales_MG_14
Clostridiales_MG_19
Clostridiales_MF_93
Firmicutes_MG_2
Clostridium_scindens_1
Erysipelatoclostridium_ramosum_1
Blautia_hydrogenotrophica_1
Clostridium_20
Eisenbergiella_tayi_1
Eisenbergiella_massiliensis_1
Clostridiales_MG_4
Clostridiales_5
Hungatella_hathewayi_1
Clostridiales_MF_122
Clostridiales_MG_43
Blautia_4
Ruminococcaceae_35
Intestinimonas_butyriciproducens_1
Blautia_massiliensis_3
Clostridium_32
Clostridium_39
Lachnoclostridium_4
Butyricicoccus_desmolans_1
Flavonifractor_plautii_1
Ruthenibacterium_lactatiformans_1
Lachnospiracea_28
Faecalibacterium_prausnitzii_27
Faecalibacterium_15
Clostridiales_MG_66
Lachnospiraceae_55
Clostridiales_MG_111
Clostridium_12
Blautia_faecis_2
Lachnospiraceae_MG_58
Lachnospiraceae_8
Lachnospiraceae_MG_44
Haemophilus_parainfluenzae_6
Haemophilus_MS_2

A
B
F
P

−6 −4 −2 0 2 4 6

Fold Change (log_2)

Parabacteroides merdae
Clostridium scindens
Erysipelatoclostridium ramosum
Blautia hydrogenotrophica
Eisenbergiella tayi
Eisenbergiella massiliensis
Hungatella hathewayi
Intestinimonas butyriciproducens
Alistipes onderdonkii
Blautia massiliensis
Butyricicoccus desmolans
Flavonifractor plautii
Ruthenibacterium lactatiformans

Fibromyalgia

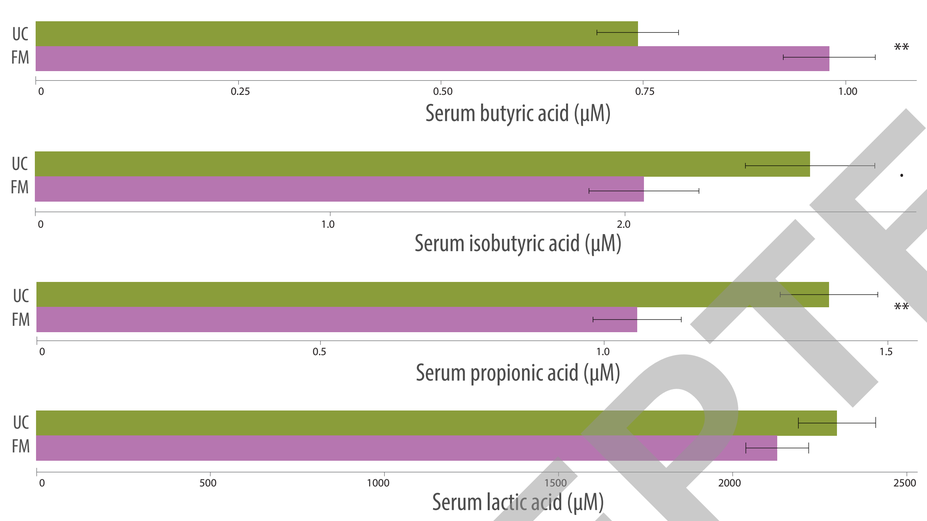
Faecalibacterium prausnitzii
Blautia faecis
Haemophilus parainfluenzae
Prevotella copri
Bacteroides uniformis

Control

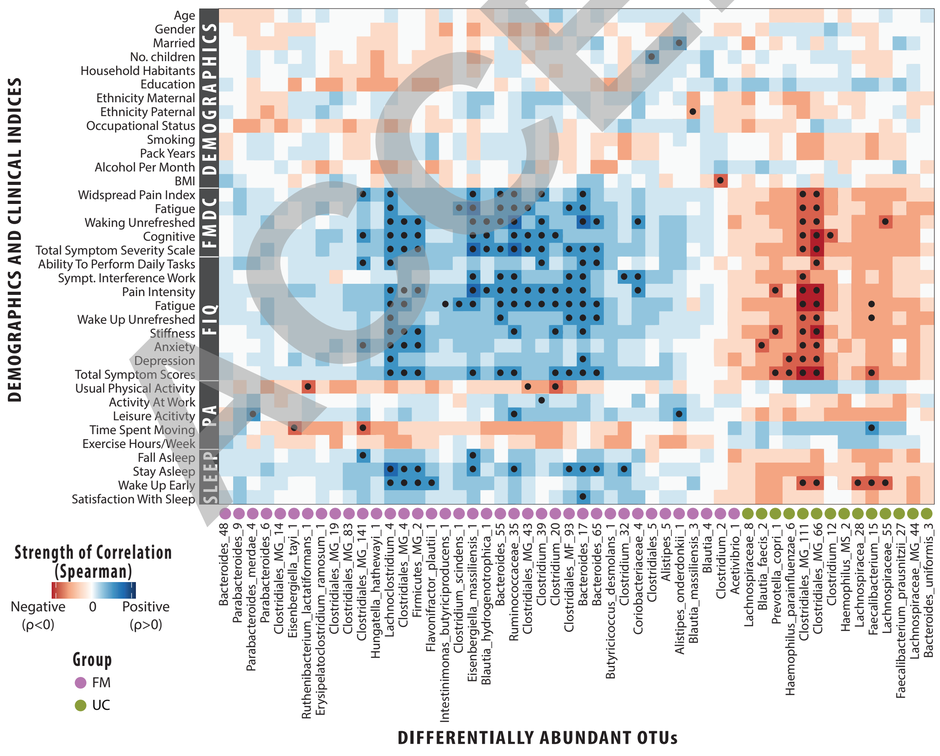
Phylum
A: Actinobacteria
B: Bacteroidetes
F: Firmicutes
P: Proteobacteria

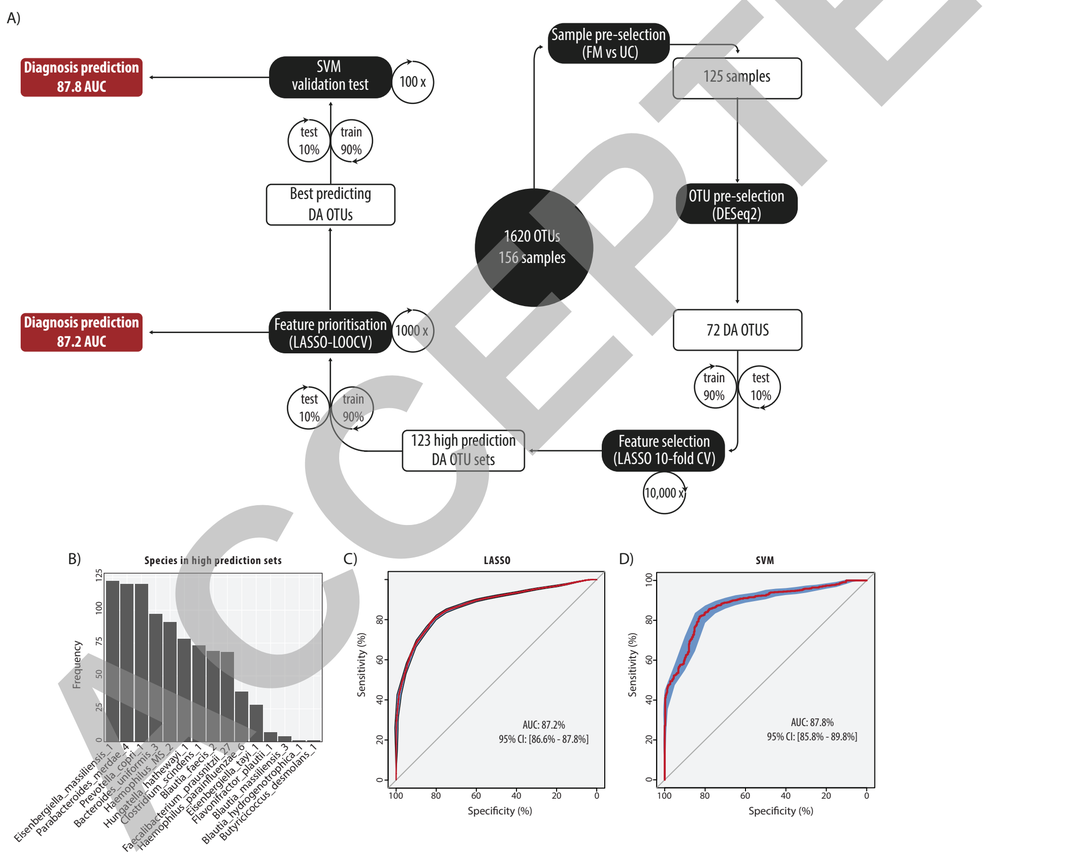
Group
UC
FM

A)


UC
FM
**
0
0.25
0.50
0.75
1.00
Serum butyric acid (μM)
UC
FM
.
0
1.0
2.0
Serum isobutyric acid (μM)
UC
FM
**
0
0.5
1.0
1.5
Serum propionic acid (μM)
UC
FM
0
500
1000
1500
2000
2500
Serum lactic acid (μM)

B)


DEMOGRAPHICS AND CLINICAL INDICES
Age
Gender
Married
No. children
Household Habitants
Education
Ethnicity Maternal
Ethnicity Paternal
Occupational Status
Smoking
Pack Years
Alcohol Per Month
BMI
Widspread Pain Index
Fatigue
Waking Unrefreshed
Cognitive
Total Symptom Severity Scale
Ability To Perform Daily Tasks
Sympt. Interference Work
Pain Intensity
Fatigue
Wake Up Unrefreshed
Stiffness
Anxiety
Depression
Total Symptom Scores
Usual Physical Activity
Activity At Work
Leisure Acitivty
Time Spent Moving
Exercise Hours/Week
Fall Asleep
Stay Asleep
Wake Up Early
Satisfaction With Sleep
DEMOGRAPHICS
FMDC
FIQ
PA
SLEEP
Strength of Correlation (Spearman)
Negative (ρ<0)
0
Positive (ρ>0)
Group
FM
UC
Bacteroides_48
Parabacteroides_9
Parabacteroides_meridae_4
Parabacteroides_6
Clostridiales_MG_14
Eisenbergiella_tayi_1
Ruthenibacterium_lactatiformans_1
Erysipelatoclostridium_ramosum_1
Clostridiales_MG_19
Clostridiales_MG_83
Clostridiales_MG_141
Hungatella_hathewayi_1
Lachnoclostridium_4
Clostridiales_MG_4
Firmicutes_MG_2
Flavonifractor_plautii_1
Intestinimonas_butyriciproducens_1
Clostridium_scindens_1
Eisenbergiella_massiliensis_1
Blautia_hydrogenotrophica_1
Bacteroides_55
Ruminococcaceae_35
Clostridiales_MG_43
Clostridium_39
Clostridium_20
Clostridiales_MF_93
Bacteroides_17
Bacteroides_65
Butyricicoccus_desmolans_1
Clostridium_32
Coriobacteriaceae_4
Clostridiales_5
Alistipes_5
Alistipes_onderdonkii_1
Blautia_massiliensis_3
Blautia_4
Clostridium_2
Acetivibrio_1
Lachnospiraceae_8
Blautia_faecis_2
Prevotella_copri_1
Haemophilus_parainfluenzae_6
Clostridiales_MG_111
Clostridiales_MG_66
Clostridium_12
Haemophilus_MS_2
Lachnospiracea_28
Faecalibacterium_15
Lachnospiraceae_55
Faecalibacterium_prausnitzii_27
Lachnospiraceae_MG_44
Bacteroides_uniformis_3
DIFFERENTIALLY ABUNDANT OTUs


A)
Diagnosis prediction
87.8 AUC
SVM
validation test
100 x
test
10%
train
90%
Best predicting
DA OTUs
Diagnosis prediction
87.2 AUC
Feature prioritisation
(LASSO-LOOCV)
1000 x
test
10%
train
90%
123 high prediction
DA OTU sets
1620 OTUs
156 samples
Sample pre-selection
(FM vs UC)
125 samples
OTU pre-selection
(DESeq2)
72 DA OTUS
train
90%
test
10%
Feature selection
(LASSO 10-fold CV)
10,000 x
B)
Species in high prediction sets
Frequency
0
25
50
75
100
125
Eisenbergiella_massiliensis
Parabacteroides_merdae_4
Prevotella_copri_1
Bacteroides_uniformis_1
Haemophilus_3
Haemophilus_MS_2
Hungatella_hathewayi_1
Blautia_faecis_2
Clostridium_scindens_1
Faecalibacterium_prausnitzii_27
Haemophilus_parainfluenzae_6
Eisenbergiella_tayi_1
Flavonifractor_plautii_1
Blautia_massiliensis_3
Blautia_hydrogenotrophica_1
Butyricicoccus_desmolans_1
C)
LASSO
Sensitivity (%)
Specificity (%)
AUC: 87.2%
95% CI: [86.6% - 87.8%]
D)
SVM
Sensitivity (%)
Specificity (%)
AUC: 87.8%
95% CI: [85.8% - 89.8%]