

RESEARCH

Open Access



Reduced diversity and altered composition of the gut microbiome in individuals with myalgic encephalomyelitis/chronic fatigue syndrome

Ludovic Giloteaux¹, Julia K. Goodrich^{1,2}, William A. Walters^{1,2}, Susan M. Levine³, Ruth E. Ley^{1,2} and Maureen R. Hanson^{1*}

Abstract

Background: Gastrointestinal disturbances are among symptoms commonly reported by individuals diagnosed with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). However, whether ME/CFS is associated with an altered microbiome has remained uncertain. Here, we profiled gut microbial diversity by sequencing 16S ribosomal ribonucleic acid (rRNA) genes from stool as well as inflammatory markers from serum for cases ($n = 48$) and controls ($n = 39$). We also examined a set of inflammatory markers in blood: C-reactive protein (CRP), intestinal fatty acid-binding protein (I-FABP), lipopolysaccharide (LPS), LPS-binding protein (LBP), and soluble CD14 (sCD14).

Results: We observed elevated levels of some blood markers for microbial translocation in ME/CFS patients; levels of LPS, LBP, and sCD14 were elevated in ME/CFS subjects. Levels of LBP correlated with LPS and sCD14 and LPS levels correlated with sCD14. Through deep sequencing of bacterial rRNA markers, we identified differences between the gut microbiomes of healthy individuals and patients with ME/CFS. We observed that bacterial diversity was decreased in the ME/CFS specimens compared to controls, in particular, a reduction in the relative abundance and diversity of members belonging to the Firmicutes phylum. In the patient cohort, we find less diversity as well as increases in specific species often reported to be pro-inflammatory species and reduction in species frequently described as anti-inflammatory. Using a machine learning approach trained on the data obtained from 16S rRNA and inflammatory markers, individuals were classified correctly as ME/CFS with a cross-validation accuracy of 82.93 %.

Conclusions: Our results indicate dysbiosis of the gut microbiota in this disease and further suggest an increased incidence of microbial translocation, which may play a role in inflammatory symptoms in ME/CFS.

Keywords: Myalgic encephalomyelitis, Chronic fatigue syndrome, Inflammation, Lipopolysaccharides, Microbiome, Microbial translocation, Beta-diversity

Background

Myalgic encephalomyelitis (ME), also known as chronic fatigue syndrome (CFS), or ME/CFS, is a debilitating illness of unknown etiology with no widely accepted therapy. Primary symptoms reported by patients are fatigue, muscle and/or joint pain, sore throat, headaches, unrefreshing sleep, and post-exertional malaise and

have been the basis of the widely used Fukuda diagnostic criteria [1]. Many ME/CFS patients also report gastrointestinal (GI) symptoms, including but not limited to irritable bowel syndrome (IBS) [2–6]. Intestinal discomfort is also indicated in a survey of drug use by individuals with CFS compared to controls, which found significantly more use of antacids, H2 blockers, and proton pump inhibitors in the ME/CFS cohort [7].

The prevalence of bowel symptoms has led to attempts to treat the disease by probiotic oral or rectal supplements. Borody et al. [8] reported improvements in a majority of

* Correspondence: mrh5@cornell.edu

¹Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA

Full list of author information is available at the end of the article



patients at 4 weeks following bacteriotherapy comprised of rectal infusion of 13 enteric bacteria, though the number with a sustained response was not well documented. In two small studies, marginal improvement in certain symptoms was reported following oral probiotic therapy [9, 10].

Two reports suggest altered gut microbiota in ME/CFS patients. Using culture-based methods, Sheedy et al. [11] described higher levels of D-lactic acid producing *Enterococcus* and *Streptococcus* spp. in ME/CFS patients vs. controls. More recently, Norwegian ME/CFS patients and healthy controls were found to exhibit differences in gut microbiota composition through a 16S rRNA gene sequencing study [12]. It is well documented that gut microbiota can be significant with respect to pathological intestinal conditions such as ulcerative colitis (UC), Crohn’s disease (CD) [13], and systemic diseases such as diabetes [14]. Because of the frequent occurrence of GI disturbances, as well as these prior reports of abnormalities, we investigated the diversity and composition of the gut microbiota of ME/CFS patients in comparison to healthy individuals.

Along with GI symptoms, individuals with ME/CFS appear to have both immune activation and immune dysfunction. Many of the common symptoms reported by ME/CFS patients are characteristic of inflammatory illnesses [15]. Most reports concerning cytokine levels in ME/CFS patients vs. controls are somewhat limited in scope and discordant, but several recent papers with a 51-plex cytokine assay indicate abnormal immune signatures in plasma and cerebrospinal fluid [16, 17].

Abnormal immune activation can be caused by translocation of microbes from an inflamed gut [18]. A prior report indicated increased IgA and IgM to lipopolysaccharide (LPS) in serum of CFS patients [19]. We therefore assayed plasma levels of LPS and LPS-binding protein, as well as the LPS/LBP receptor sCD14 [20]. We also examined the levels of C-reactive protein, an inflammatory marker, and I-FABP as a marker for gastrointestinal tract integrity [21].

Objective molecular markers for diagnosis of ME/CFS are lacking. We examined the levels of plasma markers and microbiota composition in the diseased vs. healthy subjects in order to determine whether the data, taken together, could predict ME/CFS vs. healthy status.

Results

Study population characteristics

Subjects with ME/CFS were established patients of a ME/CFS specialist, Susan Levine, M.D. and fit the Fukuda diagnostic criteria [1]. This study began before the criteria for systemic exertion intolerance disease (SEID) were established [22], but most, perhaps all, also fit the description of SEID. Of the 48 patients and 39 control participants who

self-reported good health, 34 ME/CFS patients and 7 controls self-reported gastrointestinal disturbances such as constipation, diarrhea, or intestinal discomfort. Many ME/CFS patients are able to identify an acute, often flu-like, illness that immediately preceded the onset of the disease eventually diagnosed as ME/CFS, while others are unaware of an initiating event and consider their onset to be gradual. Among the 48 ME/CFS patients in the study, 19 indicated a gradual and 25 stated a sudden onset. ME/CFS subjects completed the SF-36 form (Additional file 1: Figure S1) and Bell’s Disability scale (Table 1).

In comparison to other studies in which patients diagnosed with ME/CFS also filled out the SF-36 form, our study population fell within the same ranges on the eight subscales of the SF-36 (Additional file 1: Figure S1).

Measurements of levels of microbial translocation markers indicate microbial translocation

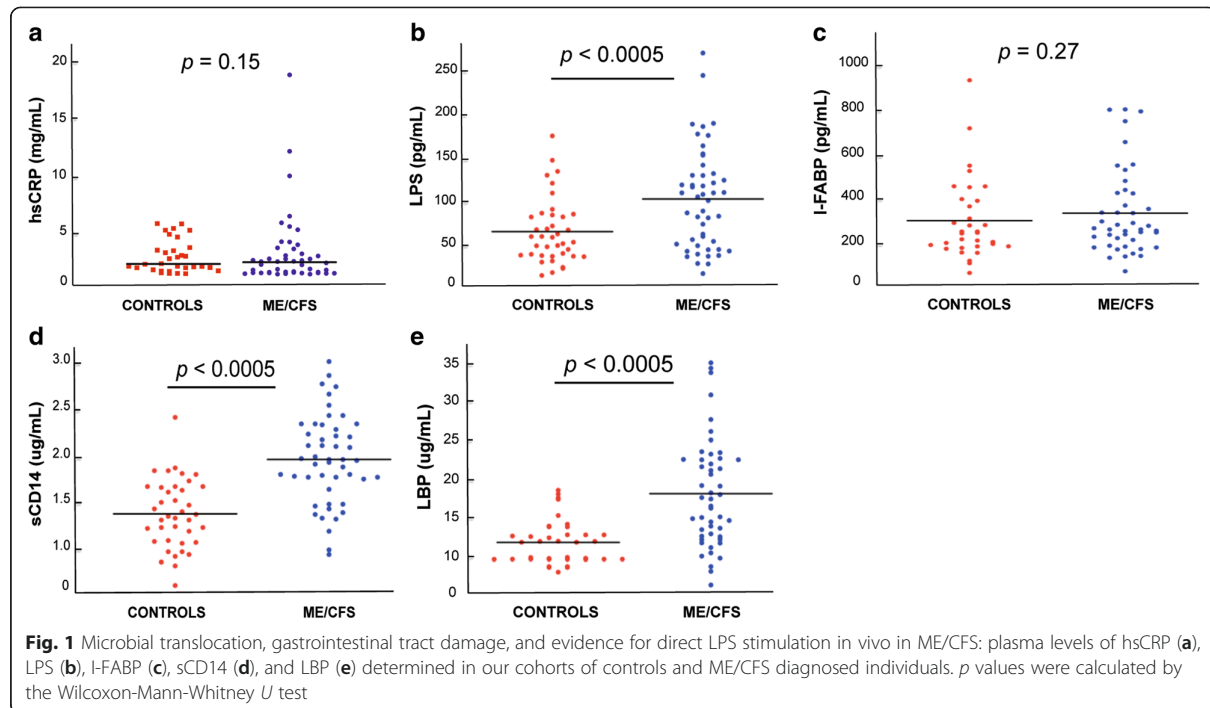
We quantified plasma levels of hsCRP, lipopolysaccharides (LPS) as a marker of microbial translocation (MT) and plasma intestinal fatty acid binding protein (I-FABP) as a marker for gastrointestinal tract damage in both groups. The distribution of plasma hsCRP, LPS and I-FABP is shown in Fig. 1. Levels of hsCRP were higher in the ME/CFS population in comparison to healthy controls (1.38 and 1.21 mg/L, respectively), but the difference was not statistically significant ($P = 0.15$, Fig. 1a, Table 2).

ME/CFS patients had significantly higher plasma LPS levels than healthy individuals (median ME/CFS—119.43 pg/mL vs. controls—74.74 pg/mL, $P < 0.0005$, Fig. 1b and Table 2). The median plasma I-FABP level was 341.9 pg/mL in the ME/CFS group and 301 pg/mL in the healthy group. Though the median I-FABP levels in the ME/CFS group was higher than that of the healthy group, the difference was not statistically significant ($P = 0.27$, Fig. 1c, Table 2).

To obtain further information concerning chronic LPS stimulation in vivo, we also measured plasma sCD14 levels and plasma LBP, which is produced by gastrointestinal and

Table 1 Characteristics of the study population

		Controls (n = 39)	ME/CFS (n = 49)
Gender	Female	30	38
	Male	9	11
Age	Mean ± SD	45.5 ± 9.9	50.2 ± 12.6
	Median (range)	48 (20–61)	51 (19–71)
BMI	Mean ± SD	27.1 ± 6.1	25.5 ± 4.9
	Median (range)	26 (17–47)	24.5 (16–40)
Bell’s disability scale	10–20	NA	15
	30–40		21
	50–60		9
	>60		4



hepatic epithelial cells. Thus, increased LPS in the circulation promote hepatic synthesis of LBP, a plasma protein that increases the binding of LPS to CD14. sCD14 and LBP concentrations in both groups are shown in Fig. 1. For the ME/CFS cohort the median plasma sCD14 concentration was 1.97 ug/mL, and the median LBP plasma concentration was 17.68 ug/mL. These values were significantly different from the plasma sCD14 and LBP concentrations of the healthy volunteers (1.36 ug/mL; $P < 0.0005$ and 12.32 ug/mL; $P < 0.0005$, respectively) (Fig. 1d, e, Table 2).

Next, we analyzed the associations among biomarker measurements in the ME/CFS population. As can be seen in Fig. 2a, b plasma LPS levels correlated positively with levels of sCD14 and LBP ($r = 0.347$, $P < 0.01$ and $r = 0.487$, $P < 0.01$, respectively), consistent with stimulation of sCD14 production by LPS in vivo. In addition, we found a strong significant correlation between plasma sCD14 and hsCRP and sCD14 and LBP; high levels of sCD14 were associated with high levels of hsCRP ($r = 0.507$, $P < 0.01$) and LBP ($r = 0.578$, $P < 0.01$) (Fig. 2c, d). We also analyzed whether enterocyte damage (i.e., I-FABP levels) was associated with the proposed microbial translocation markers LPS, sCD14, and LBP. We found no relationship between I-FABP and LPS levels ($r = -0.125$; $P = 0.278$), I-FABP and sCD14 levels ($r = -0.117$; $P = 0.310$), or I-FABP and LBP levels ($r = -0.08$; $P = 0.488$).

Stool microbiota of ME/CFS patients exhibit reduced diversity and different composition than healthy controls

The hypervariable V4 region of 16S rRNA genes was sequenced from fecal samples of individuals with ME/CFS ($n = 48$) and healthy individuals ($n = 39$). A total of 8,534,117 high-quality and classifiable reads were generated from all samples, with an average of $98,093 \pm 29,231$ reads per sample. Binning sequences using a pairwise identity threshold of 97 %, we obtained an average of 1330 ± 423 operational taxonomic units (OTUs) per sample. The sequence-based rarefaction curves based on the Phylogenetic Diversity (PD) metric were nearly asymptotic and a Wilcoxon rank-sum test demonstrated a statistical difference in the diversity of ME/CFS and healthy individuals ($P = 0.004$, $W = 1268$) (Fig. 3a).

We examined the number of “observed species,” i.e., the number of 97 % ID OTUs observed in 32,223 sequences, the estimators of community evenness (Shannon *H*), and richness (Chao1 and PD) in the two group of samples. ME/CFS samples had a significant overall lower microbial diversity, with lower evenness ($H = 5.33 \pm 0.93$ vs. 5.92 ± 0.93 , $P = 0.004$), and lower richness (observed species, 1204 ± 351 vs. 1486 ± 456 ; Chao1, 2363 ± 704 vs. 2918 ± 885 , $P = 0.002$; PD, 61.6 ± 16.7 vs. 73.4 ± 19.04 , $P = 0.004$) (Fig. 3b).

To evaluate overall differences in beta-diversity between the microbiomes, we applied Principal Component Analysis (PCoA) to weighted and unweighted UniFrac distance

Table 2 Plasma levels of markers of inflammation (hsCRP), microbial translocation (LPS, sCD14, and LBP) and gastrointestinal damage (I-FABP) in ME/CFS and healthy individuals

Analytes	Control	ME/CFS
hsCRP (mg/L)		
<i>n</i>	34	46
Median	1.21	1.38
Quartiles	0.38–2.69	0.70–2.71
Range	0.27–5.09	0.23–19.3
LPS (pg/mL)		
<i>n</i>	39	49
Median	74.74	119.43
Quartiles	54.34–99.54	66.21–144.41
Range	32.21–187.32	34.32–279.30
I-FABP (pg/mL)		
<i>n</i>	33	44
Median	234.40	255.85
Quartiles	159.75–412.90	171.275–450.70
Range	14–1067.9	23.2–909.9
sCD14 (ug/mL)		
<i>n</i>	39	49
Median	1.36	1.97
Quartiles	1.09–1.67	1.71–2.32
Range	0.62–2.42	0.95–3.02
LBP (ug/mL)		
<i>n</i>	37	49
Median	12.32	17.68
Quartiles	10.22–13.73	13.04–22.56
Range	8.65–18.76	7.06–34.52

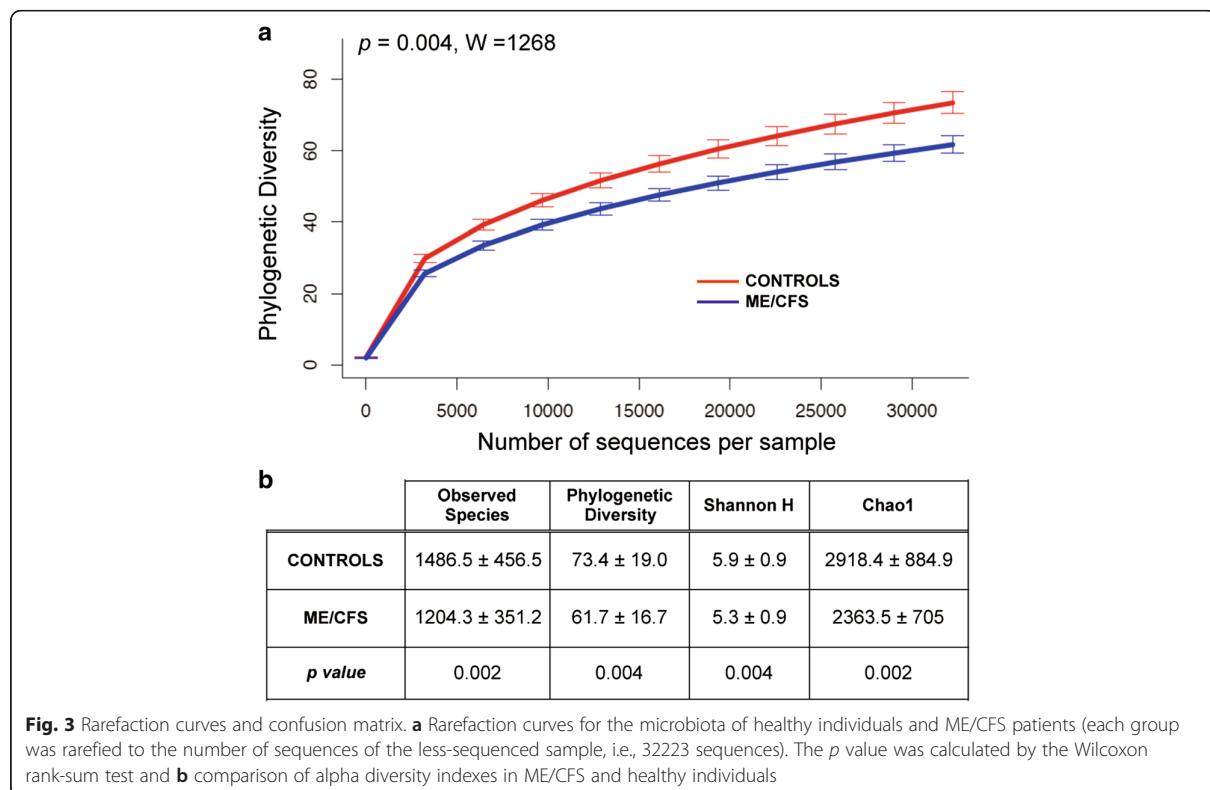
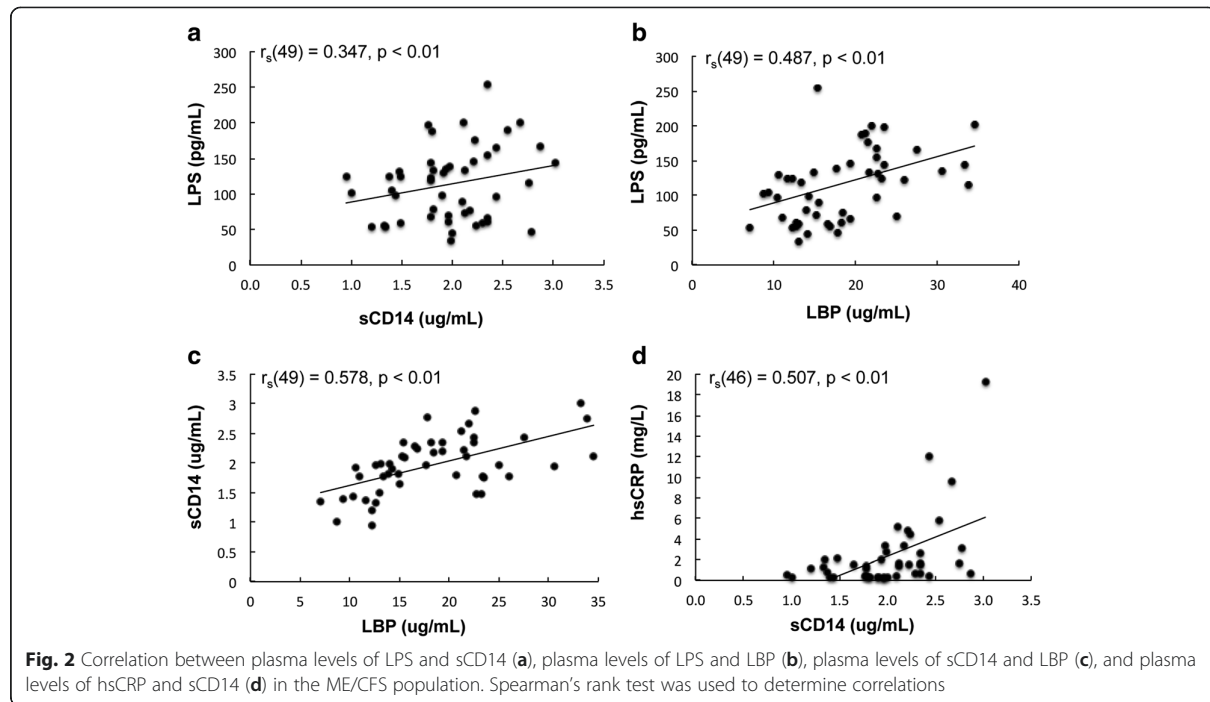
metric matrices generated for the sample set. Within the microbial community cluster, there appears to be no clear difference in beta-diversity between the ME/CFS group and healthy group using both weighted (Additional file 2: Figure S2a) and unweighted (Additional file 2: Figure S2b) UniFrac distance matrices. None of the other parameters tested, i.e., sex, BMI, or clinical data revealed clustering (data not shown). Because beta-diversity clustering as measured by UniFrac shows how dissimilar overall community structure is between samples, the samples may not cluster in a manner that reflects differences detected at the OTU level, or the overall alpha diversity within groups.

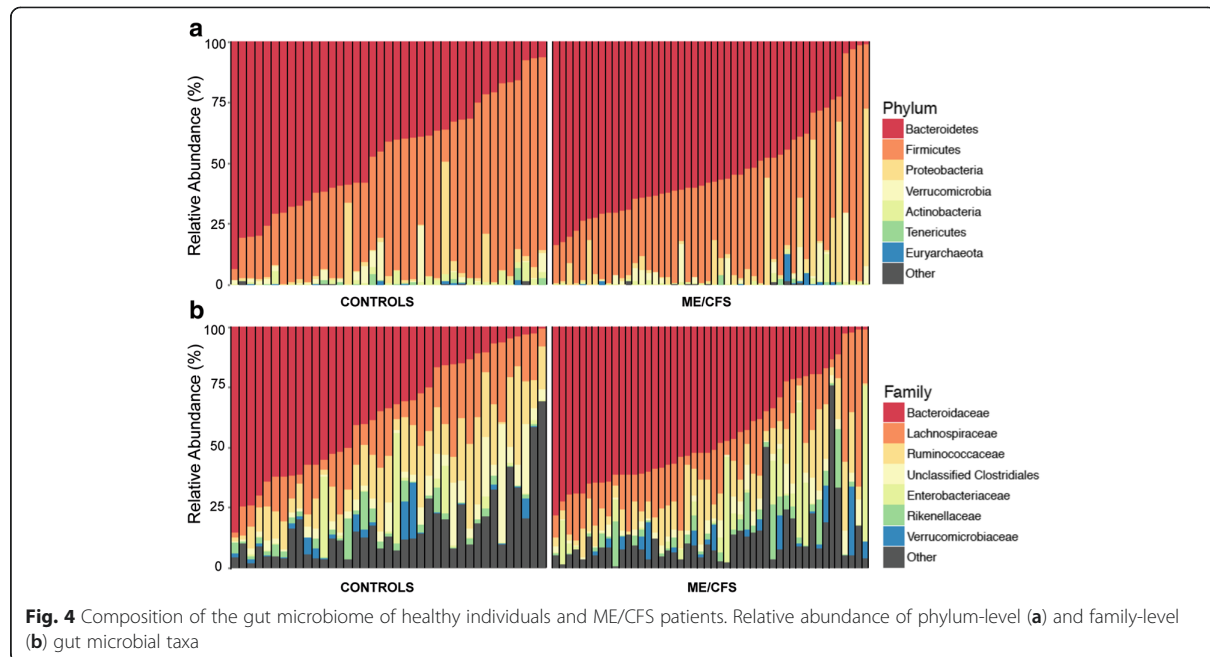
The overall microbial composition for ME/CFS and controls differed at the phylum and family levels (Fig. 4a, b), although none of these differences were statistically significant after multiple test correction. The two largest phyla represented in each dataset of healthy and ME/CFS-afflicted individuals were Firmicutes and Bacteroidetes. In

healthy individuals, this corresponded to 46 and 45 % respectively of the rarified 16S rRNA sequences. Also, Proteobacteria made up the next largest represented phylum (3.6 %), with Verrucomicrobia and Actinobacteria in relatively low relative abundance (2.1 and 1.6 %, respectively). At the phylum level, the abundance of the Bacteroidetes was comparable (52 %) in both datasets (Fig. 4a). ME/CFS samples showed lower relative abundance of Firmicutes (35 %) (Fig. 4a) and higher relative abundance of Proteobacteria (8 %), due almost entirely to a twofold increase in the Proteobacteria family Enterobacteriaceae (6 vs. 3 % for ME/CFS and healthy individuals, respectively) (Fig. 4b). Within the Firmicutes, at the family level, Ruminococcaceae were lower in the ME/CFS samples (16 vs. 11 % in ME/CFS and healthy individuals respectively) (Fig. 4b), whereas Lachnospiraceae were similar among both datasets (16 % for both healthy and ME/CFS samples). Some differences were noted between cases and controls in family members of the Bacteroidetes, i.e., Bacteroidaceae (35 vs. 43 %), Rickenellaceae (3 vs. 4 %), and Prevotellaceae (3.2 vs. 0.7 %). Finally, within the Actinobacteria, Bifidobacteriaceae were lower in the ME/CFS samples (1 vs. 0.5 %).

At the OTU level, 40 OTUs were found to be significantly different between groups after multiple testing correction. The majority of them belonged to the Firmicutes phylum, including members of the Ruminococcaceae family such as *Oscillospira* spp. ($q = 0.016$), *Faecalibacterium prausnitzii* ($q = 0.014$), and *Ruminococcus* spp. ($q = 0.014$) and members of the Lachnospiraceae, i.e., *Coproccoccus* spp. ($q = 0.014$). Other OTUs included members of the Actinobacteria such as *Eggerthella lenta* ($q = 0.014$) and *Collinsella aerofaciens* ($q = 0.014$).

These significant differences were further confirmed by LEfSe analysis, which uses linear discriminant analysis (LDA) coupled with effect size measurements to identify bacterial taxa whose sequences are differentially abundant between ME/CFS and healthy individuals. In addition to detecting significant features, LEfSe also ranks features by effect size, which put features explaining most of the biological difference at top (Segata et al. 2011). LEfSe identified 24 discriminative features (genus level, LDA score >2) whose relative abundance varied significantly among fecal samples taken from the ME/CFS and healthy groups (Fig. 5). ME/CFS microbiota were enriched with an unclassified member of the Desulfohalobacteriaceae and genera from the Firmicutes phylum, i.e., *Oscillospira*, *Lactococcus*, *Anaerotruncus* and *Coprobacillus* and *Eggerthella*, a member of the Actinobacteria phylum ($P < 0.05$, Fig. 5). Eighteen genera were enriched in the control group compared to the ME/CFS group (Fig. 5) with members mainly belonging to the Firmicutes phylum. We observed that members of the Ruminococcaceae and Bifidobacteriaceae, i.e., *Faecalibacterium* and *Bifidobacterium*, respectively, were



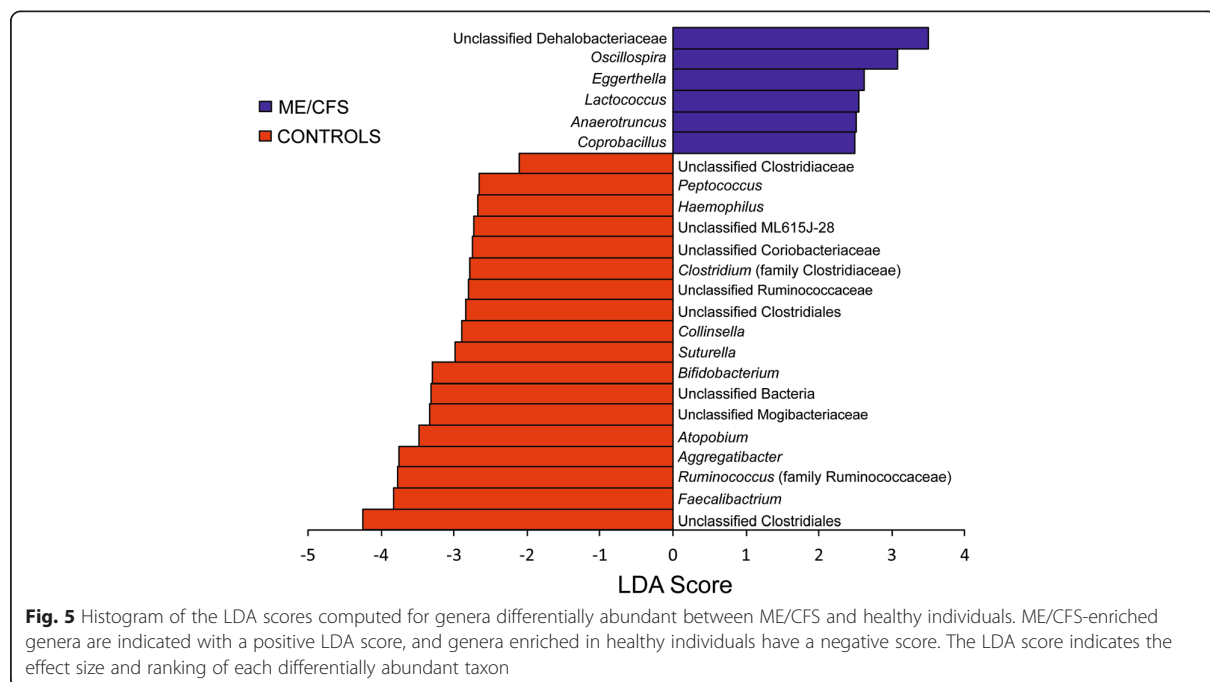


significantly increased in healthy individuals ($P = 0.03$ and 0.04 , respectively).

Classifying subjects into patients vs. controls from inflammatory markers and microbiome data

Using a machine learning approach, samples were mostly successfully classified into healthy and ME/CFS groups,

with the highest proportion of samples correctly classified when genus-level taxa along with data from the inflammatory markers were used in the analysis. With 97 % ID OTUs used in the analysis, 82 % of the samples could be correctly classified (standard deviation of 0.14). With OTUs collapsed at the species level, the average accuracy was 0.80 with a standard deviation of 0.11. Collapsing



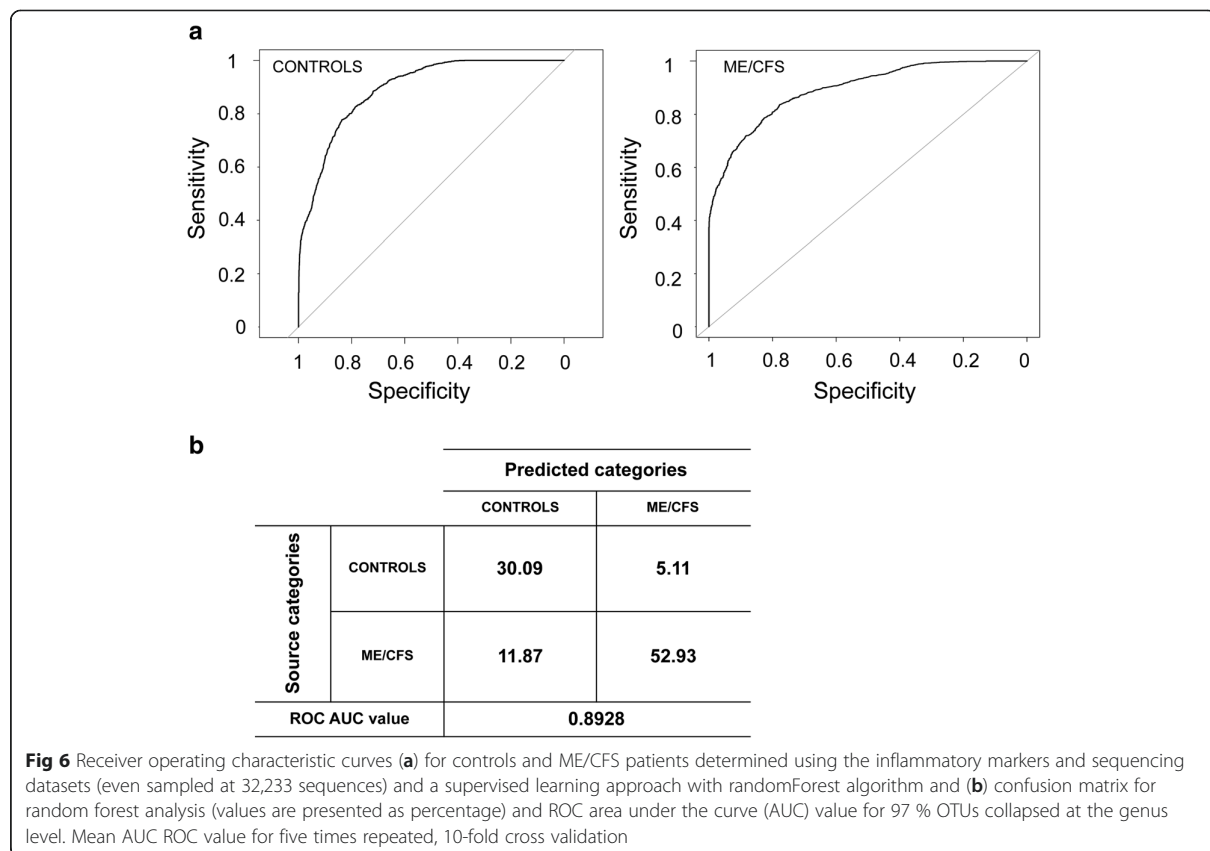
taxonomy to the genus level, individuals with ME/CFS were classified correctly and separately from the healthy group with an average success rate of 0.82 ± 0.12 . The receiver operating curves, the AUC ROC value for the ME/CFS samples (0.89), and the confusion matrix are presented in Fig. 6. The feature importance scores for the genus-level analysis, which shows the relative importance of clinical values and microbial abundances, are available in Additional file 3: Table S1. Additionally, processing microbial sequencing data without including BMI and blood inflammatory marker levels results in 70, 75, and 72 % classification accuracy for genus, species, and OTU-level data respectively (confusion matrices available in Additional file 4: Figure S3).

Discussion

Our analysis of the microbiome in cases suggests that gastrointestinal tract of ME/CFS patients is a pro-inflammatory environment. This environment might cause damage to the intestinal epithelium, thus augmenting microbial translocation (MT) and subsequently triggering an immune response. It has been previously documented that disruption of mucosal barrier function occurs in ME/CFS as demonstrated by the increased serum concentrations of

IgA and IgM to LPS of Gram-negative enterobacteria [19]. Our data supports the hypothesis of increased MT in the ME/CFS group as evidenced by (i) significantly raised levels of plasma LPS and (ii) significantly higher levels of sCD14 and LBP, as indicators of direct LPS stimulation. Increased gut permeability and increased LPS levels have been also described in patients with liver diseases, alcoholic, and nonalcoholic steatohepatitis [23], during chronic HIV infection [24], and in inflammatory bowel disease (IBD) [25, 26] suggesting that an activation of pro-inflammatory and endotoxin-signaling cascades could be important for disease progression in ME/CFS. Consequently, high plasma LPS levels in ME/CFS could result from an increased production of endotoxin upon changes in the gut microbiota. Furthermore, we observed that sCD14 levels positively correlated with levels of LPS, LBP, and hsCRP. If there is damage to the gut mucosa, microbial translocation could increase, altering antimicrobial regulators and dysregulating the innate immune system.

As a marker, LPS is limited to particular microbes, as it is only present in Gram-negative bacteria. sCD14 is produced primarily by macrophages and hepatocytes in response to LPS but is also stimulated by other bacterial and viral agents [27]. LBP functions as a co-factor along



with sCD14 and is constitutively synthesized in hepatocytes to recognize LPS released to the bloodstream but various inflammatory factors such as IL22, IL-6, and TNF- α can induce its expression [28, 29]. Nevertheless, we found significantly more patients with elevated levels of these biomarkers in comparison to the healthy group, suggesting that more MT occurs in people affected by ME/CFS.

Using both aerobic and anaerobic culturing methods, Butt and colleagues were the first to present evidence of altered fecal microbiota in ME/CFS patients compared to healthy individuals [30]. Subsequently, using culture methods and metabolite analysis, Sheedy et al. [11] obtained information concerning the fecal microbiome in patient and health cohorts. Both found that D-lactic acid-producing *Enterococcus* and *Streptococcus* species were strongly over-represented in ME/CFS patients and that among anaerobic bacteria, *Prevotella* was a bacterial genus found to be in excess in subjects with ME/CFS.

Recently, a study used high-throughput 16S rRNA gene sequencing to investigate the presence of specific alterations in the gut microbiota of ME/CFS patients from Belgium and Norway [12]. The authors amplified the V5 and V6 hypervariable 16S rRNA regions and sequenced the amplicons using a Roche FLX 454 sequencer, which resulted in an average of only 6000–7000 reads/sample. In contrast, we amplified the V4 hypervariable region of the 16S rRNA gene, sequenced amplicons using the MiSeq Illumina platform, obtained an average of many more reads/sample (98,000), and compared the resulting sequences to a different database, the Greengenes non-redundant reference database [31]. Our analysis showed that within-sample diversity is lower in the ME/CFS specimens compared to controls. The same indices in the Fremont et al. [12] study did not differ between ME/CFS and healthy subjects [12], likely due to the lower read number they obtained. Lower richness has also been observed in unhealthy or inflammatory states [32, 33] and has been associated with IBD, necrotizing enterocolitis [34], and greater abdominal discomfort levels in patients with food intolerances [35, 36].

Regardless of disease state, bacteria belonging to the Firmicutes, Bacteroidetes, Proteobacteria, and Actinobacteria phyla represented the vast majority of sequences identified. We observed reduced levels of members of the dominant phylum Firmicutes, also noted repeatedly in Crohn's disease patients [13, 37, 38]. Proteobacteria were more abundant in ME/CFS patients than in controls, observed as well in inflammatory bowel disease (IBD) patients [39, 40]. In an inflamed gut, infiltrating macrophages and neutrophils release sulfur- and nitrogen-derived metabolites such as tetrathionate and nitrate [41–44]. Opportunistic members of the Proteobacteria can take advantage of the host inflammatory response by using

these compounds as electron acceptors [43, 44] to generate energy and foster their own growth in the gut. We did not collect information concerning diet of patients and thus do not know whether this factor might have affected the composition and/or metabolism of the colonic microbiota in our cohorts.

We observed significantly lower levels of the genus *Faecalibacterium*, a member of the Ruminococcaceae in the ME/CFS population. For example, *Faecalibacterium prausnitzii*, which produces an anti-inflammatory protein [45], is reduced in ME/CFS cases relative to controls. This genus is also depleted in IBD [13, 38] and ulcerative colitis [46] and has been shown to have anti-inflammatory properties both in vitro and in vivo [37]. *Faecalibacterium* belongs to a group of producers of butyrate, a short chain fatty acid known to have anti-inflammatory properties and to protect the intestine [40]. Individuals with IBD and IBS [47] exhibit a lack of butyrate-producing bacteria and lower levels of butyrate in their gut [48, 49] which modulates different processes including hormone and cytokine secretion (e.g., leptin, IL-10) and activation of immune/inflammatory responses [50–52].

We also found a decrease in *Bifidobacterium*, previously observed in IBS [53–57], IBD [58], and type II diabetes [59]. Bifidobacteria are a group of lactic acid-producing bacteria that are widely used as probiotics and as targets for prebiotics [60]. Treatment with *Bifidobacterium infantis* 35624 was reported to reduce CRP levels in a cohort of ME/CFS patients [61].

We have employed a supervised machine learning approach to help prediction of disease state based on the microbiome sequence datasets [62]. Using this approach, we were able to classify unlabeled samples with some degree of accuracy, as demonstrated by the high AUC ROC value obtained (0.8928) at the genus level. This method has been recently used in several microbiome surveys to accurately place individuals into an IBD/healthy category [63], including ulcerative colitis (ROC AUC = 0.9225) and colonic (ROC AUC = 0.8787) or ileal Crohn's disease (ROC AUC = 0.9699) [64]. Such an approach could therefore serve as a complement to other non-invasive diagnoses of symptoms or as an initial diagnosis to determine if the subject likely has ME/CFS. Because this is a relatively small cohort, to move to a formal diagnostic clinical application, a large cohort of ME/CFS and healthy controls would be needed to verify that classification would retain its accuracy with independent sample handling and sequencing.

Conclusions

Taken together, our results suggest an ongoing damage to the gut mucosa, leading to increased microbial translocation in ME/CFS, which in turn could alter antimicrobial

regulators and dysregulate the innate immune system. Differences between the gut microbiomes of healthy individuals and patients with ME/CFS were identified in terms of relative abundance of specific genera. There is no single precise alteration of the gut microbiota in all ME/CFS patients we examined, but our data converges to support the concept of a less diverse and unstable community of bacteria in the disorder. It highlights the association of specific bacterial taxa with ME/CFS, and the identification of the underlying role of this altered commensal gut microbiota could lead to novel diagnostic and therapeutic strategies that would improve clinical outcome. Future studies may also reveal additional molecular markers that could be combined with gut microbiome information to enhance the sensitivity and specificity of ME/CFS diagnostic assays.

The cause of ME/CFS is unknown, but gut dysbiosis could be contributing to some of the symptoms and their severity. Developing therapeutic interventions aimed at reducing local inflammation, restoring gastrointestinal tract immunity and integrity and modifying the intestinal microbiome may ameliorate ME/CFS symptoms in a number of affected patients.

Methods

Human subjects and sample/data collection

All work involving human subjects was approved by the Cornell University Institutional Review Board. Fecal samples were collected at home by participants in 15-ml conical tubes containing RNAlater (Life Technologies, Grand Island, NY) and refrigerated prior to shipment. Upon arrival at Cornell University, the samples were divided into aliquots and stored at -80°C until processing. Blood samples were drawn into EDTA and heparin tubes from an antecubital vein with subjects in the seated position. Samples were shipped by overnight courier from New York City to Cornell University (Ithaca). Upon receipt, samples were centrifuged at 4000 r.p.m. for 30 min to pellet blood cells, and plasma was stored at -80°C until further analyses. BMI, age, and gender of subjects were recorded. ME/CFS subjects completed the Short Form 36 Health Survey (SF-36.org) and Bell's Disability Scale [65]. Potential controls from the same geographic area as cases were screened by the physician for suitability as healthy controls. As indicated in Table 1, mean and median ages of cases and controls were within 5 years and the female to male ratios were similar.

Plasma level determination of hsCRP, sCD14, LBP, LPS, and I-FABP

High-sensitive C-reactive protein (hsCRP) was measured from unhemolyzed EDTA plasma using a Chemiluminescence immunoassay on an Immulite 2000 (Siemens Medical Solutions Diagnostics, Deerfield, IL). Markers for microbial translocation, sCD14 and LBP, were measured

in plasma samples by commercially available enzyme-linked immunosorbent assays (ELISA). Plasma sCD14 was quantified using the Quantikine Human sCD14 Immunoassay (R&D Systems, Minneapolis, MN), and plasma LBP was measured by LBP soluble ELISA kit (Hycult Biotechnology, Uden, The Netherlands) according to the manufacturers' protocols. Plasma bacterial endotoxin, i.e., LPS, was measured from heparinized blood samples (Brandtzaeg) using the Limulus Amebocyte Lysate (LAL) assay (Lonza Group Ltd, Allendale, NJ). The method uses a chromogenic endpoint assay yielding data as endotoxin units (EU/ml). Briefly, 100 μl of each plasma sample was diluted in 200 μl of β -G-Blocker (Lonza Group Ltd, Allendale, NJ) to eliminate the possibility of false positives. Samples were further diluted with 100 μl of pyrogen-free water to give a final dilution of 1:4. All dilutions were prepared in pyrogen-free tubes. Samples were then placed in a water bath at 85°C for 15 min to inactivate inhibitory plasma proteins. Results of LPS measured were expressed in picograms per milliliter (1 EU/ml = 100 pg/ml). Levels of intestinal fatty acid binding protein (I-FABP), a marker associated with enterocyte damage, were assayed using an ELISA (Hycult Biotechnology, Uden, the Netherlands) according to the manufacturer's instructions. All the samples were run in duplicate.

Statistical analysis of plasma markers levels

We initially performed a Shapiro-Wilk test to check if the data was normally distributed [66]. In case of violation of normality, data was log transformed and checked again for normality. Both parametric independent samples *t* test and a non-parametric Wilcoxon-Mann-Whitney *U* test were used to determine the significance of differences in each subject group. Values of $P < 0.05$ were considered statistically significant. All data from the biomarkers levels determination were processed and analyzed in SPSS Statistics Version 21 (Armonk, NY: IBM Corp).

DNA extraction, 16S rRNA gene sequencing

Metagenomic DNA was isolated from an aliquot of ~ 100 mg from each fecal sample using the PowerSoil-htp DNA isolation kit (MoBio Laboratories Ltd, Carlsbad, CA), which involves both chemical and physical lysis of the cells. We amplified 16S rRNA genes (V4 hypervariable region) from bulk DNA using the 515F and 806R primers as previously described [67] prior to sequencing. Duplicate PCR reactions of samples and extraction blanks consisted of 2.5X HotMasterMix (5-Prime, Inc., Gaithersburg, MD), 10–100 ng DNA template, and 0.05 μM of each primer. DNA amplification of samples and extraction blanks was performed on a 96-well plate with a minicycler PTC 200 (MJ Research) starting with 3-min denaturation at 94°C , followed by 25 cycles consisting of denaturation (45 s at 94°C), annealing (60 s at 50°C), extension (90 s at 72°C),

and a final extension at 72 °C for 10 min. Samples were randomly distributed on the plate with no grouping for sample type. The replicate PCR reactions were combined and purified using a magnetic bead system (Mag-Bind EZPure, Omega Bio-Tek, Norcross, GA). PCR amplicons were quantified using the QuantiT PicoGreen dsDNA Assay Kit (Invitrogen, Carlsbad, CA). Aliquots of amplicons (at equal masses) were combined for a final concentration of approximately 15 ng/μl. Extraction blanks showed no amplification. All amplicons were then sequenced on a single run using the Illumina MiSeq 2x250 bp platform at Cornell Biotechnology Resource Center Genomics Facility.

Quality filtering and analysis of the 16S rRNA gene sequence data were performed with QIIME 1.9.0 as previously described [68]. Briefly, matching paired-end raw sequences (mate-pairs) were merged using the fastq-join command in the ea-utils software package (<http://code.google.com/p/ea-utils>), and merged sequences with less than a 200-bp overlap were filtered out of the dataset. The remaining merged sequences were quality filtered and assigned to samples based on their barcodes using the default parameters of QIIME. Sequences were assigned to 97 % ID OTUs by comparing them to a non-redundant reference database of near-full length sequences [31]. All OTUs that were observed fewer than two times, i.e., singletons, were removed from the analysis. The OTU table was rarefied to the sequence count of the sample with the lowest sequence depth, 32,223 sequences per sample, and used in all subsequent analyses. For statistical comparisons of healthy individuals to those afflicted with ME/CFS, *p* values obtained with the Wilcoxon-Mann-Whitney *U* test were corrected for multiple comparisons using the false discovery rate of Benjamini and Hochberg, implemented in the QIIME pipeline. We used both the weighted and unweighted UniFrac distance metrics as measures of between-sample (beta) diversity and applied principal coordinates analysis (PCoA) to visualize patterns of diversity. Within-samples (alpha) diversity was calculated using three different measures (1) ChaoI index [69]; (2) Shannon Index [70]; and (3) Phylogenetic Diversity [71].

LEfSe analysis and machine learning

Linear discriminant effect size analysis (LEfSe) on filtered datasets [72] was performed at the genus level to find features (genera) differentially represented between healthy and ME/CFS groups. LEfSe combines the standard tests for statistical significance (Kruskal-Wallis test and pairwise Wilcoxon test) with linear discriminate analysis. It ranks features by effect size, which put features that explain most of the biological difference at top. LEfSe analysis was performed under the following conditions: the α value for the factorial Kruskal-Wallis test among classes was 0.05 and the threshold on the logarithmic LDA score for discriminative features was 2.0.

A machine learning approach was used to identify variables discriminating the two groups of samples (feature selection). For these analyses, we used either 97 % OTUs, or taxa abundances based on combining OTUs at the species and genus levels. Classification of samples as healthy controls or ME/CFS was carried out by using a random forest approach with supervised learning [73] and area under the curve (AUC) calculation to optimize feature (e.g., abundance of a particular genus) selection, implemented in the software package R. Scripts, required packages, and instructions for processing data are available on <https://gist.github.com/walterst/2222618976a66b3fc8dd>. In addition to the taxonomic abundance data, levels of inflammatory markers (BMI, sCD14, LBP, LPS, and I-FABP) were included in the analysis. Average accuracies were calculated with five repeats of 10-fold cross validation, which is intended to predict the accuracy of the model, and indicate over-fitting if significantly different than the full dataset results, by subsampling the data and testing this training subsample against the remaining data reference set.

Additional files

Additional file 1: Figure S1. 36-Item Short Form Health Survey (SF-36) profiles from studies reporting SF-36 scores for individuals with a ME/CFS diagnosis. (PDF 2514 kb)

Additional file 2: Figure S2. Principal Coordinate Analysis (PCoA) plot of healthy controls versus subjects with ME/CFS. Distances were calculated with weighted UniFrac (a) and unweighted UniFrac (b). Data were evenly sampled at 32223 sequences per sample. (PDF 2631 kb)

Additional file 3: Table S1. Feature Importance Scores for genus-level supervised learning. The feature importance score is the percentage increase in error rate when the given feature is permuted while other values remain constant. As there are only two categories, the increased error rate is equal for both categories. (XLSX 42 kb)

Additional file 4: Figure S3. Confusion matrices for random forest analysis of microbial sequencing data (values are presented as %) and ROC area under the curve (AUC) values at the genus (a), species (b) and OTU (c) level. (PDF 1170 kb)

Additional file 5: Table S2. Per-sample metadata mapping file used throughout the QIIME pipeline. (XLSX 61 kb)

Abbreviations

hsCRP, high sensitivity C-reactive protein; IBD, inflammatory bowel disease; IBS, irritable bowel syndrome; I-FABP, intestinal fatty acid binding protein; LBP, lipopolysaccharide-binding protein; LPS, lipopolysaccharides; ME/CFS, myalgic encephalomyelitis/chronic fatigue syndrome; MT, microbial translocation; sCD14, soluble CD14

Acknowledgements

We thank the subjects for providing samples and information for the study and Lin Lin for technical assistance.

Funding

This work was supported by grant 1R21AI101614 from NIH NIAID to M.R.H. and R.E.L. The funders had no role in study design, data collection, analysis and interpretation, decision to publish, or preparation of the manuscript.

Availability of data and materials

The sequence data supporting the results of this article are available in the European Bioinformatics Institute Sequence Read Archive under accession number PRJEB13092. The mapping file used for the QIIME pipeline is available on Additional file 5: Table S2.

Authors' contributions

LG designed the experiments, processed the samples, conducted the experiments, and performed the statistical analysis with SPSS. SML recruited, diagnosed, and sampled the blood from the subjects. LG and JKG performed the sequence analysis using QIIME. WAW performed the supervised learning machine analysis. RL and MH contributed to study design. LG, JKG, WAW, RL, and MH performed analysis and writing. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

The institutional review board of Cornell University approved the study, and patients and controls gave written informed consent after the study protocol was fully explained. All consented to blood draw, to provide a stool sample and to the availability of the stored samples for additional bioassays.

Author details

¹Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA. ²Department of Microbiology, Cornell University, Ithaca, NY, USA. ³Private Practice, New York, NY, USA.

Received: 29 February 2016 Accepted: 11 May 2016

Published online: 23 June 2016

References

- Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann Intern Med.* 1994;121(12):953–9.
- Riedl A, Schmidtman M, Stengel A, Goebel M, Wisser AS, Klapp BF, et al. Somatic comorbidities of irritable bowel syndrome: a systematic analysis. *J Psychosom Res.* 2008;64(6):573–82.
- Whitehead WE, Palsson O, Jones KR. Systematic review of the comorbidity of irritable bowel syndrome with other disorders: what are the causes and implications? *Gastroenterology.* 2002;122(4):1140–56.
- Wojczynski MK, North KE, Pedersen NL, Sullivan PF. Irritable bowel syndrome: a co-twin control analysis. *Am J Gastroenterol.* 2007;102(10):2220–9.
- Aaron LA, Burke MM, Buchwald D. Overlapping conditions among patients with chronic fatigue syndrome, fibromyalgia, and temporomandibular disorder. *Arch Intern Med.* 2000;160(2):221–7.
- Komaroff AL, Buchwald D. Symptoms and signs of chronic fatigue syndrome. *Rev Infect Dis.* 1991;13 Suppl 1:S8–11.
- Jones JF, Nisenbaum R, Reeves WC. Medication use by persons with chronic fatigue syndrome: results of a randomized telephone survey in Wichita, Kansas. *Health Qual Life Outcomes.* 2003;1:74.
- Borody TJ, Nowak A, Finlayson S. The GI microbiome and its role in chronic fatigue syndrome: a summary of bacteriotherapy. *ACNEM J.* 2012;31(3):3–8.
- Sullivan A, Nord CE, Evengard B. Effect of supplement with lactic-acid producing bacteria on fatigue and physical activity in patients with chronic fatigue syndrome. *Nutr J.* 2009;8:4.
- Rao AV, Bested AC, Beaulne TM, Katzman MA, Iorio C, Berardi JM, et al. A randomized, double-blind, placebo-controlled pilot study of a probiotic in emotional symptoms of chronic fatigue syndrome. *Gut Pathog.* 2009;1(1):6.
- Sheedy JR, Wettenhall RE, Scanlon D, Gooley PR, Lewis DP, McGregor N, et al. Increased d-lactic acid intestinal bacteria in patients with chronic fatigue syndrome. *In Vivo.* 2009;23(4):621–8.
- Fremont 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;22:50–6.
- Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A.* 2007;104(34):13780–5.
- Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature.* 2012;490(7418):55–60.
- Komaroff AL, Buchwald DS. Chronic fatigue syndrome: an update. *Annu Rev Med.* 1998;49:1–13.
- Hornig M, Gottschalk G, Peterson DL, Knox KK, Schultz AF, Eddy ML, et al. Cytokine network analysis of cerebrospinal fluid in myalgic encephalomyelitis/chronic fatigue syndrome. *Mol Psychiatry.* 2016;21(2):261–9.
- Hornig M, Montoya JG, Klimas NG, Levine S, Felsenstein D, Bateman L, et al. Distinct plasma immune signatures in ME/CFS are present early in the course of illness. *Sci Adv.* 2015;1(1).
- Vyboh K, Jenabian MA, Mehraj V, Routy JP. HIV and the gut microbiota, partners in crime: breaking the vicious cycle to unearth new therapeutic targets. *J Immunol Res.* 2015;2015:614127.
- Maes M, Mihaylova I, Leunis JC. Increased serum IgA and IgM against LPS of enterobacteria in chronic fatigue syndrome (CFS): indication for the involvement of gram-negative enterobacteria in the etiology of CFS and for the presence of an increased gut-intestinal permeability. *J Affect Disord.* 2007;99(1–3):237–40.
- Schumann RR. Function of lipopolysaccharide (LPS)-binding protein (LBP) and CD14, the receptor for LPS/LBP complexes: a short review. *Res Immunol.* 1992;143(1):11–5.
- Pelsters MMAL, Hermens WT, Glatz JFC. Fatty acid-binding proteins as plasma markers of tissue injury. *Clinica Chimica Acta.* 2005;352(1–2):15–35.
- Medicine II. Beyond myalgic encephalomyelitis/chronic fatigue syndrome: redefining an illness. Washington, D.C.: The National Academies Press; 2015.
- Zhu L, Baker SS, Gill C, Liu W, Alkhoury R, Baker RD, et al. Characterization of gut microbiomes in nonalcoholic steatohepatitis (NASH) patients: a connection between endogenous alcohol and NASH. *Hepatology.* 2013;57(2):601–9.
- Openshaw PJ. Crossing barriers: infections of the lung and the gut. *Mucosal Immunol.* 2009;2(2):100–2.
- Wellmann W, Fink PC, Benner F, Schmidt FW. Endotoxaemia in active Crohn's disease. Treatment with whole gut irrigation and 5-aminosalicylic acid. *Gut.* 1986;27(7):814–20.
- Wyatt J, Vogelsang H, Hubl W, Waldhoer T, Lochs H. Intestinal permeability and the prediction of relapse in Crohn's disease. *Lancet.* 1993;341(8858):1437–9.
- Anas AA, Hovius JW, van 't Veer C, van der Poll T, de Vos AF. Role of CD14 in a mouse model of acute lung inflammation induced by different lipopolysaccharide chemotypes. *PLoS One.* 2010;5(4):e10183.
- Wolk K, Witte E, Hoffmann U, Doecke WD, Endesfelder S, Asadullah K, et al. IL-22 induces lipopolysaccharide-binding protein in hepatocytes: a potential systemic role of IL-22 in Crohn's disease. *J Immunol.* 2007;178(9):5973–81.
- Wan Y, Freeswick PD, Khemlani LS, Kispert PH, Wang SC, Su GL, et al. Role of lipopolysaccharide (LPS), interleukin-1, interleukin-6, tumor necrosis factor, and dexamethasone in regulation of LPS-binding protein expression in normal hepatocytes and hepatocytes from LPS-treated rats. *Infect Immun.* 1995;63(7):2435–42.
- Butt HL, Dunstan R, N.R. M, T. R. Bacterial colonisation in patients with persistent fatigue. In Proceedings of the AHMF international clinical and scientific conference Sydney, Australia: AHMF; 2001.
- Brodie EL, Desantis TZ, Joyner DC, Baek SM, Larsen JT, Andersen GL, et al. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl Environ Microbiol.* 2006;72(9):6288–98.
- Chang JY, Antonopoulos DA, Kalra A, Tonelli A, Khalife WT, Schmidt TM, et al. Decreased diversity of the fecal microbiome in recurrent Clostridium difficile-associated diarrhea. *J Infect Dis.* 2008;197(3):435–8.
- Willing BP, Dicksved J, Halfvarson J, Andersson AF, Lucio M, Zheng Z, et al. A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology.* 2010;139(6):1844–54. e1.
- McMurtry VE, Gupta RW, Tran L, Blanchard EE, Penn D, Taylor CM, et al. Bacterial diversity and Clostridia abundance decrease with increasing severity of necrotizing enterocolitis. *Microbiome.* 2015;3:11.
- Fujimoto T, Imaeda H, Takahashi K, Kasumi E, Bamba S, Fujiyama Y, et al. Decreased abundance of Faecalibacterium prausnitzii in the gut microbiota of Crohn's disease. *J Gastroenterol Hepatol.* 2013;28(4):613–9.

36. Hippe B, Remely M, Bartosiewicz N, Riedel M, Nichterl C, Schatz L, et al. Abundance and diversity of GI microbiota rather than IgG4 levels correlate with abdominal inconvenience and gut permeability in consumers claiming food intolerances. *Endocr Metab Immune Disord Drug Targets*. 2014;14(1):67–75.
37. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Gratadoux JJ, et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci U S A*. 2008;105(43):16731–6.
38. Martinez-Medina M, Aldeguer X, Gonzalez-Huix F, Acero D, Garcia-Gil LJ. Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. *Inflamm Bowel Dis*. 2006;12(12):1136–45.
39. Sartor RB. Key questions to guide a better understanding of host-commensal microbiota interactions in intestinal inflammation. *Mucosal Immunol*. 2011;4(2):127–32.
40. Sartor RB. Genetics and environmental interactions shape the intestinal microbiome to promote inflammatory bowel disease versus mucosal homeostasis. *Gastroenterology*. 2010;139(6):1816–9.
41. Davies CA, Rocks SA, O'Shaughnessy MC, Perrett D, Winyard PG. Analysis of nitrite and nitrate in the study of inflammation. *Methods Mol Biol*. 2003;225:305–20.
42. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, et al. Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. *Proc Natl Acad Sci U S A*. 2011;108(42):17480–5.
43. Winter SE, Lopez CA, Baumler AJ. The dynamics of gut-associated microbial communities during inflammation. *EMBO Rep*. 2013;14(4):319–27.
44. Winter SE, Winter MG, Xavier MN, Thiennimitr P, Poon V, Keestra AM, et al. Host-derived nitrate boosts growth of *E. coli* in the inflamed gut. *Science*. 2013;339(6120):708–11.
45. Quevrain E, Maubert MA, Michon C, Chain F, Marquant R, Tailhades J, et al. Identification of an anti-inflammatory protein from Faecalibacterium prausnitzii, a commensal bacterium deficient in Crohn's disease. *Gut*. 2016;65(3):415–25.
46. Machiels K, Joossens M, Sabino J, De Preter V, Arijis I, Eeckhaut V, et al. A decrease of the butyrate-producing species *Roseburia hominis* and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. *Gut*. 2014;63(8):1275–83.
47. Pozuelo M, Panda S, Santiago A, Mendez S, Accarino A, Santos J, et al. Reduction of butyrate- and methane-producing microorganisms in patients with irritable bowel syndrome. *Sci Rep*. 2015;5:12693.
48. Varela E, Manichanh C, Gallart M, Torrejon A, Borrueal N, Casellas F, et al. Colonisation by Faecalibacterium prausnitzii and maintenance of clinical remission in patients with ulcerative colitis. *Aliment Pharmacol Ther*. 2013;38(2):151–61.
49. Sokol H, Seksik P, Furet JP, Firmesse O, Nion-Larmurier I, Beaugerie L, et al. Low counts of Faecalibacterium prausnitzii in colitis microbiota. *Inflamm Bowel Dis*. 2009;15(8):1183–9.
50. Saemann MD, Bohmig GA, Osterreicher CH, Burtscher H, Parolini O, Diakos C, et al. Anti-inflammatory effects of sodium butyrate on human monocytes: potent inhibition of IL-12 and up-regulation of IL-10 production. *FASEB J*. 2000;14(15):2380–2.
51. Segain JP, Raingeard de la Bletiere D, Bourreille A, Leray V, Gervois N, Rosales C, et al. Butyrate inhibits inflammatory responses through NF- κ B inhibition: implications for Crohn's disease. *Gut*. 2000;47(3):397–403.
52. Vinolo MA, Rodrigues HG, Nachbar RT, Curi R. Regulation of inflammation by short chain fatty acids. *Nutrients*. 2011;3(10):858–76.
53. Kassinen A, Krogius-Kurikka L, Makivuokko H, Rinttila T, Paulin L, Corander J, et al. The fecal microbiota of irritable bowel syndrome patients differs significantly from that of healthy subjects. *Gastroenterology*. 2007;133(1):24–33.
54. Tana C, Umesaki Y, Imaoka A, Handa T, Kanazawa M, Fukudo S. Altered profiles of intestinal microbiota and organic acids may be the origin of symptoms in irritable bowel syndrome. *Neurogastroenterol Motil*. 2010;22(5):512–9, e114–5.
55. Malinen E, Rinttila T, Kajander K, Matto J, Kassinen A, Krogius L, et al. Analysis of the fecal microbiota of irritable bowel syndrome patients and healthy controls with real-time PCR. *Am J Gastroenterol*. 2005;100(2):373–82.
56. Matto J, Maunukela L, Kajander K, Palva A, Korpela R, Kassinen A, et al. Composition and temporal stability of gastrointestinal microbiota in irritable bowel syndrome—a longitudinal study in IBS and control subjects. *FEMS Immunol Med Microbiol*. 2005;43(2):213–22.
57. Maukonen J, Satokari R, Matto J, Soderlund H, Mattila-Sandholm T, Saarela M. Prevalence and temporal stability of selected clostridial groups in irritable bowel syndrome in relation to predominant faecal bacteria. *J Med Microbiol*. 2006;55(Pt 5):625–33.
58. Swidsinski A, Ladhoff A, Pernthaler A, Swidsinski S, Loening-Baucke V, Ortner M, et al. Mucosal flora in inflammatory bowel disease. *Gastroenterology*. 2002;122(1):44–54.
59. Wu X, Ma C, Han L, Nawaz M, Gao F, Zhang X, et al. Molecular characterisation of the faecal microbiota in patients with type II diabetes. *Curr Microbiol*. 2010;61(1):69–78.
60. Furrie E, Macfarlane S, Kennedy A, Cummings JH, Walsh SV, O'Neil DA, et al. Synbiotic therapy (Bifidobacterium longum/Synergy 1) initiates resolution of inflammation in patients with active ulcerative colitis: a randomised controlled pilot trial. *Gut*. 2005;54(2):242–9.
61. Groeger D, O'Mahony L, Murphy EF, Bourke JF, Dinan TG, Kiely B, et al. Bifidobacterium infantis 35624 modulates host inflammatory processes beyond the gut. *Gut Microbes*. 2013;4(4):325–39.
62. Knights D, Costello EK, Knight R. Supervised classification of human microbiota. *FEMS Microbiol Rev*. 2011;35(2):343–59.
63. Papa E, Docktor M, Smillie C, Weber S, Preheim SP, Gevers D, et al. Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. *PLoS One*. 2012;7(6):e39242.
64. Walters WA, Xu Z, Knight R. Meta-analyses of human gut microbes associated with obesity and IBD. *FEBS Lett*. 2014;588(22):4223–33.
65. Bell DS. The doctor's guide to chronic fatigue syndrome. Reading, Mass: Addison-Wesley; 1995.
66. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965;52(3–4):591–611.
67. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A*. 2011;108 Suppl 1:4516–22.
68. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhan R, et al. Human genetics shape the gut microbiome. *Cell*. 2014;159(4):789–99.
69. Chao A. Non-parametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*. 1984;11:265–70.
70. Shannon CE. A mathematical theory of communication. *Bell SystTech J*. 1948;27:379–423.
71. Faith DP. Conservation evaluation and phylogenetic diversity. *Biol Conserv*. 1992;61:1–10.
72. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60.
73. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2:18–22.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

