ORIGINAL ARTICLE

# Primary sclerosing cholangitis is characterised by intestinal dysbiosis independent from IBD

João Sabino,[1] Sara Vieira-Silva,[2,3] Kathleen Machiels,[1] Marie Joossens,[2,3,4] Gwen Falony,[2,3] Vera Ballet,[1] Marc Ferrante,[1] Gert Van Assche,[1] Schalk Van der Merwe,[5] Severine Vermeire,[1] Jeroen Raes[2,3]

[1]Translational Research Center for Gastrointestinal Disorders (TARGID), University of Leuven, Leuven, Belgium
[2]Department of Microbiology and Immunology, Laboratory of Molecular Bacteriology, KU Leuven—University of Leuven, Rega Institute for Medical Research, Leuven, Belgium
[3]Center for the Biology of Disease, VIB, Leuven, Belgium
[4]Department of Microbiology, VUB, Brussels, Belgium
[5]Department of Microbiology and Immunology, Center for the Biology of Disease, REGA institute, KU Leuven—VIB, Leuven, Belgium
[6]Department of Hepatology, KU Leuven, Leuven, Belgium

**Correspondence to**
Jeroen Raes, Laboratorium Moleculaire Bacteriologie (Rega Instituut), O&N I Herestraat 49—bus 1028, Leuven 3000, Belgium; jeroen.raes@vib-kuleuven.be

JS, SV-S Joint first co-authorship

SV, JR Joint last co-authorship

Received 30 October 2015
Revised 18 April 2016
Accepted 19 April 2016
Published Online First 20 May 2016

► http://dx.doi.org/10.1136/gutjnl-2016-312137

## ABSTRACT

**Objective** Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease often leading to end-stage liver disease. Its pathogenesis remains largely unknown, although frequent concomitant IBD hints towards common factors underlying gut and bile duct inflammation. Considering the mounting evidence on the involvement of the intestinal microbiota in initiating and determining IBD phenotype, we investigated intestinal microbiota composition in patients with PSC.

**Design** Stool samples were collected from 147 individuals (52 patients with PSC, 52 age, gender and body mass index-matched healthy volunteers, 13 UC and 30 patients with Crohn's disease). An independent validation cohort of 14 PSC and 14 matched controls was recruited. 16S rDNA sequencing of faecal DNA was performed (Illumina MiSeq).

**Results** The microbiota of patients with PSC was characterised by decreased microbiota diversity, and a significant overrepresentation of Enterococcus (p=3.76e-05), Fusobacterium (p=3.76e-05) and Lactobacillus (p=0.0002) genera. This dysbiosis was present in patients with PSC with and without concomitant IBD and was distinct from IBD, and independent of treatment with ursodeoxycholic acid. A decision tree based on abundances of these three genera allowed reliable classification in the validation cohort. In particular, one operational taxonomic unit belonging to the Enterococcus genus was associated with increased levels of serum alkaline phosphatase (p=0.048), a marker of disease severity.

**Conclusions** We here present the first report of PSC-associated faecal dysbiosis, independent from IBD signatures, suggesting the intestinal microbiota could be a contributing factor in PSC pathogenesis. Further studies are needed to confirm these findings and assess causality.

## Significance of this study

### What is already known on this subject?
► Primary sclerosing cholangitis is a cholestatic liver disease strongly associated with IBD.
► Intestinal microbiota play an important role in the pathogenesis of IBD.
► Antibiotic therapy in patients with primary sclerosing cholangitis transiently improves liver function tests without however altering transplant free survival.

### What are the new findings?
► Primary sclerosing cholangitis is associated with alterations in intestinal microbiota, independently of comorbidity with IBD.
► Three genera Enterococcus, Lactobacillus and Fusobacterium are overrepresented in patients with primary sclerosing cholangitis.
► An operational taxonomic unit belonging to the Enterococcus genus is positively correlated with the levels of alkaline phosphatase.

### How might it impact on clinical practice in the foreseeable future?
► Intestinal microbiota modulation through diet, faecal microbiota transplantation, antibiotics or probiotics may be used in the treatment or prevention of primary sclerosing cholangitis.

## INTRODUCTION

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease characterised by the development of multifocal bile duct strictures that can lead to liver fibrosis and subsequent cirrhosis.[1] PSC has an incidence of 1.3 per 100 000 individuals. There is no effective medical treatment for this condition and liver transplantation is offered to patients with PSC with end-stage liver disease, although PSC recurrence occurs in up to 23% of patients after liver transplantation.[2]

The pathogenesis of PSC remains poorly understood, with current evidence suggesting that genetic, immunologic and environmental factors all play a role. Between 60% and 80% of patients with PSC have concomitant IBD, most frequently ulcerative colitis (UC), suggesting that inflammation in the colon is of importance in disease presentation. The intestinal microbiota has also been suggested to play a role in PSC pathogenesis, as translocated bacterial products are more frequently found in explant livers from patients with PSC when compared with patients with other liver disorders.[3] Metronidazole therapy, which alters bacterial microbiota composition, transiently improves liver function tests without however altering transplant free survival.[4] Furthermore, colectomy performed before liver transplantation decreases PSC relapse rate after liver transplantation, indicating that the

colon is instrumental in the initiation of inflammation in the liver.[2] Moreover, a new antigen-dependent mouse model confirmed that immune-mediated cholangitis is caused by T cells primed in the gut-associated lymphoid tissue which further supports the hypothesis that cholangitis is gut triggered and immune mediated.[5] More recently, a Mdr2(−/−) mouse model of PSC was developed, leading to a more severe phenotype of PSC when raised in germ-free conditions, further suggesting a role of the intestinal microbiota in the development of bile duct injury.[6]

The role of the intestinal microbiota in the pathogenesis of IBD is well recognised. Bacteria influence intestinal inflammation through the interplay with the immune system, such as the induction of CD25[+] regulatory T cells, downregulation of proinflammatory and upregulation of anti-inflammatory cytokines.[7] Dysbiosis, the deviation from the normal composition of the human intestinal microbiota, has already been described in IBD. Crohn's disease (CD) dysbiosis is mainly characterised by reduced microbial richness,[7] a decrease in *Faecalibacterium prausnitzii*,[8] *Bifidobacterium adolescentis*, *Dialister invisus* and uncharacterised species of *Clostridium cluster XIVa* and an increase in the mucus-degrading *Ruminococcus gnavus*.[9] On the other hand, patients with UC display normal intestinal microbial richness[10] and UC dysbiosis is characterised by a reduction in *Roseburia hominis* and *F. prausnitzii*, both producers of butyrate, a short-chain fatty acid with known anti-inflammatory properties.[7 11]

Considering the already described IBD-associated dysbiosis and the frequent concomitant development of IBD with PSC, we hypothesised that the intestinal microbiota might be altered in patients with PSC. We analysed the composition of the microbiota in a well-characterised cohort of patients with PSC and compared them to healthy controls and patients with IBD.

## METHODS

### Patients
Patients with PSC, IBD and PSC with concomitant IBD were included. All patients were recruited at the IBD or liver outpatient clinic of the University Hospitals of Leuven (Belgium).

PSC diagnosis followed established guidelines and was based on symptoms and/or signs of chronic cholestatic liver injury, negative anti-mitochondrial autoantibodies and HIV serology, imaging compatible with bile duct injury (magnetic resonance cholangiopancreatography, endoscopic retrograde cholangiopancreatography) and/or liver biopsy with typical findings in the absence of drug use associated with cholestatic liver injury.[12] The diagnosis of liver cirrhosis was established on biopsy or by imaging with MRI or elastography in conjugation with laboratory and clinical findings supporting the diagnosis of cirrhosis (see online supplementary table S1).

Diagnosis of CD or UC was confirmed by a combination of endoscopy, histopathology and radiological and biochemical investigations, according to existing guidelines.[13 14]

Healthy controls were selected from the Flemish Gut Flora Project (FGFP) currently including over 3000 sampled volunteers, to match patients with PSC for age, gender and body mass index (BMI).

All patients and healthy controls signed informed consent before sample collection. The local ethical committee approved the study (reference number: S53684 and S58125).

### Clinical data
Basic demographic data, clinical data and information about possible confounders of microbiota analysis (eg, specific diet, prebiotic and probiotic use, antibiotic treatment in the last 30 days) were collected at the time of inclusion for all patients. Also, complementary clinical information was extracted from the clinical files of the patients. Harvey–Bradshaw index (HBI) and partial Mayo score were collected from patients with CD and UC, respectively. Table 1 summarises study subjects characteristics.

Clinical information from the healthy controls was extracted from the FGFP database, which contains a medical report from the general practitioner and self-reported information about possible confounders for intestinal microbiota analysis (eg, prebiotic and probiotic use).

### Samples and DNA extraction
Fresh faecal samples from patients were collected at the outpatient clinic of the University Hospitals of Leuven and frozen at −80°C within 12 h after sampling. Blood samples were taken during the same clinic visit.

The faecal samples of the healthy controls were collected by the volunteers and immediately frozen at −20°C in their home freezers. They were transported frozen within a week to the laboratory and stored at −80°C.

Faecal calprotectin measurements were performed for all patients with the fCAL ELISA kit (Bühlmann, Schönenbuch, Switzerland). Faecal calprotectin was not quantified in healthy controls and in patients with UC because samples were frozen on collection, which may result in overestimation, as stated by the manufacturer.

Bacterial DNA extraction from faecal samples was performed with the MOBIO PowerMicrobiome RNA isolation kit (MO BIO Laboratories, Carlsbad, California, USA), using an adapted protocol. In short, after mechanical and chemical lysis in Glass Bead Tubes, samples were incubated at 90°C for 10 min. Afterwards, normal protocol was followed with the exclusion of the DNase I step at the end.

Quantification of bacterial DNA for PCR was done with a Qubit 2.0 fluorometer (Life Technologies Grand Island, New York, USA). After PCR amplification, quality control and quantification of the libraries was performed with Fragment Analyzer (Advanced Analytical Technologies, Ames, Iowa, USA).

### Sequencing 16S rRNA gene
The V4 region of the 16S rRNA gene was amplified with primer pairs 515F and 806R, with single multiplex identifier and adaptors as described by Kozich *et al*.[15] Sequencing was performed on Illumina MiSeq sequencer (MiSeq V.2 kit, Illumina, San Diego, California, USA) yielding 250 bp paired-end reads. After demultiplexing, paired-end reads were merged using FLASH software V.1.2.10 (Johns Hopkins University, Baltimore, USA)[16] with default parameters. Combined reads quality threshold was set at minimum 30 quality score over 90% of read length (Fastx tool kit; http://hannonlab.cshl.edu/fastx_toolkit/) and chimeric sequences were filtered out (UCHIME[17]).

### Microbiota analysis
Each sample was downsized to 10 000 reads by random selection of quality-checked reads. Genus and phylum abundance matrices were obtained by mapping to the Ribosomal Database Project (RDP) reference database (RDP classifier[18]). The operational taxonomic unit (OTU)-level abundance matrix was obtained by de novo clustering of reads at 97% identity, corresponding to species-level clustering (USEARCH[19]). For specific OTU taxonomic assignment, we performed a megaBLAST search of the OTU centroid read against NCBI 16S rDNA

**Table 1** Study subjects characteristics from both cohorts (n=175)

| Patient characteristics | PSC only (n=18) | PSC-UC (n=27) | PSC-CD (n=21) | UC (n=13) | CD (n=30) | Healthy controls (n=66) |
|---|---|---|---|---|---|---|
| Male (%) | 10 (55.6) | 20 (74) | 18 (85.7) | 4 (30.8) | 15 (50) | 49 (74) |
| Median (IQR) age (years) | 49 (15.25) | 43 (14) | 49 (17) | 50 (28) | 52 (14.25) | 51.5 (17) |
| Median (IQR) BMI (kg/m²) | 23.45 (8.25) | 23.7 (6.2) | 23.5 (5.2) | 25.6 (4.9) | 25.25 (5.4) | 23.72 (4.9) |
| Current smoker (%) | 1 (5.6) | 1 (3.7) | 7 (33.3) | 2 (15.4) | 11 (36.7) | 4 (7.7) |
| Median (IQR) age (years) at diagnosis of PSC | 37.5 (15) | 32 (8.5) | 35 (21) | NA | NA | NA |
| Liver transplantation (%) | 7 (38.9) | 7 (25.9) | 1 (4.8) | 0 (0) | 0 (0) | 0 (0) |
| Cirrhosis (%) | 1 (5.6) | 7 (25.9) | 5 (23.8) | 0 (0) | 0 (0) | 0 (0) |
| Pouch (%) | 0 (0) | 4 (14.8) | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Calprotectin (μg/g)—median (IQR) | 100 (226.1) | 121 (406) | 339 (1081.95) | NA | 132.5 (214.25) | NA |
| CRP (mg/L)- median (IQR) | 1.4 (3.6) | 2.25 (3.975) | 2.05 (7.15) | 2.2 (3.8) | 2.1 (2.7) | 0.7 (1.3) |
| IBD activity score—median (IQR) | NA | 0 (2) | 1 (4.25) | 2 (3) | 3 (4) | NA |
| Medications | | | | | | |
| UDCA (%) | 13 (72.2) | 18 (66.7) | 16 (76.2) | 0 (0) | 1 (3.3) | 0 (0) |
| 5-Amminosalicylates (%) | 0 (0) | 18 (66.7) | 4 (19) | 12 (92.3) | 2 (6.7) | 0 (0) |
| Corticosteroids (%) | 3 (16.7) | 3 (11.1) | 2 (9.5) | 3 (23) | 1 (3.3) | 0 (0) |
| Immunosuppression (%) | 8 (44.4) | 7 (25.9) | 4 (19) | 4 (30.8) | 8 (26.7) | 0 (0) |
| Anti-TNF α (%) | 0 (0) | 5 (18.5) | 8 (38) | 3 (23) | 16 (53.3) | 0 (0) |
| Antibiotics in the last month (%) | 2 (11.8) | 4 (16.7) | 5 (26.3) | 0 (0) | 0 (0) | 0 (0) |
| Probiotics in the last month (%) | 1 (6.7) | 3 (13) | 3 (18.7) | 0 (0) | 5 (20,8) | 2 (3) |

BMI, body mass index; CD, Crohn's disease; PSC, primary sclerosing cholangitis; UC, ulcerative colitis; UDCA, ursodeoxycholic acid.

sequences. The oligotyping pipeline[20] was used for identification of different oligotypes in the genera of interest.

## Statistical analysis

Statistical analyses were performed in R (V.3.1.3), using the phyloseq[21] and vegan(J Oksanen, FG Blanchet, R Kindt, et al. vegan: Community Ecology Package. R package version 2.3–0, 2015) packages. Continuous variables were tested for normality with the Shapiro–Wilk test. Non-parametric test were applied to analyse microbiome data, with multiple testing correction whenever applicable (adjustment for false discovery rate (FDR)). Adjusted p values <0.05 were considered significant.

Mann–Whitney U (Kruskal–Wallis for more than two groups) was used to test median differences in α-diversity (microbiota species richness) and genera abundances between different groups. Correlation between genera abundances and continuous metadata was performed with Spearman correlation. Principal coordinates analysis (PCoA) on OTU-level community composition (metric: Bray–Curtis dissimilarity) was used to visualise microbiota variation across samples and significance of community differences between groups of patients were tested with Adonis non-parametric test.

Multivariate Association with Linear Models package (MaAsLin R V.0.0.3) was used for deconfounded multivariate assessment of associations between taxa abundances and metadata, using default parameters.[22]

Weka (V.3.6.12, University of Waikato) was used for training (first cohort) and testing (validation cohort) a J48 decision tree classifier for the microbiota signature discriminating PSC from healthy controls.[23] The model's performance in terms of accuracy of prediction was evaluated by the area under the receiver operating characteristic curve (AUC) on the test set (validation cohort).

## RESULTS
### Demographic data

Cohort 1 included 52 patients with PSC (see online supplementary figure S1): 13 had PSC without IBD (hereafter referred to as PSC only). As control groups, 52 age, gender and BMI-matched healthy controls were selected from the FGFP database and 43 additional patients with IBD without PSC were recruited.

A validation cohort of 14 patients with PSC (of whom 9 with concomitant IBD) was recruited as an independent test set. Additionally, 14 age, gender and BMI-matched healthy controls were selected from the FGFP database for the independent test set. These subjects were recruited and sampled after the initial analysis and faecal DNA was extracted and sequenced in an independent run. Patient characteristics from both cohorts are summarised in table 1.

The majority of patients with IBD were in remission, as evidenced by low clinical activity scores (HBI and partial Mayo score) and normal C-reactive protein (CRP) levels.

### Faecal dysbiosis in patients with PSC includes decreased species richness and altered community composition, regardless of concomitant IBD and UDCA treatment

The overall faecal microbiota composition of patients (PSC, CD or UC) was significantly different from that of healthy controls, as evidenced by their separation along the first axis of the PCoA (on OTU-level abundance matrix with the Bray–Curtis dissimilarity; figure 1A; Adonis R²=0.1, p value=0.00099). Focusing on the patients with PSC, samples from the different subgroups (PSC only, PSC-UC and PSC-CD) were not significantly different from one another, such that they cluster together on the PCoA (Adonis p values>0.05; see online supplementary table S2). However, they did cluster apart from patients with CD and UC (Adonis R²=0.0932, p value=0.001; figure 1A). Of note, the microbiota of patients with PSC under ursodeoxycholic acid (UDCA) treatment was not significantly different from patients with PSC without UDCA treatment (see online supplementary table S2). Also, the microbiota of patients with PSC treated who underwent antibiotic treatment in the past month was not significantly different from that of patients with PSC with no antibiotic treatment.

We next set to characterise PSC dysbiosis, starting with global alterations in microbiota richness. The species richness, defined
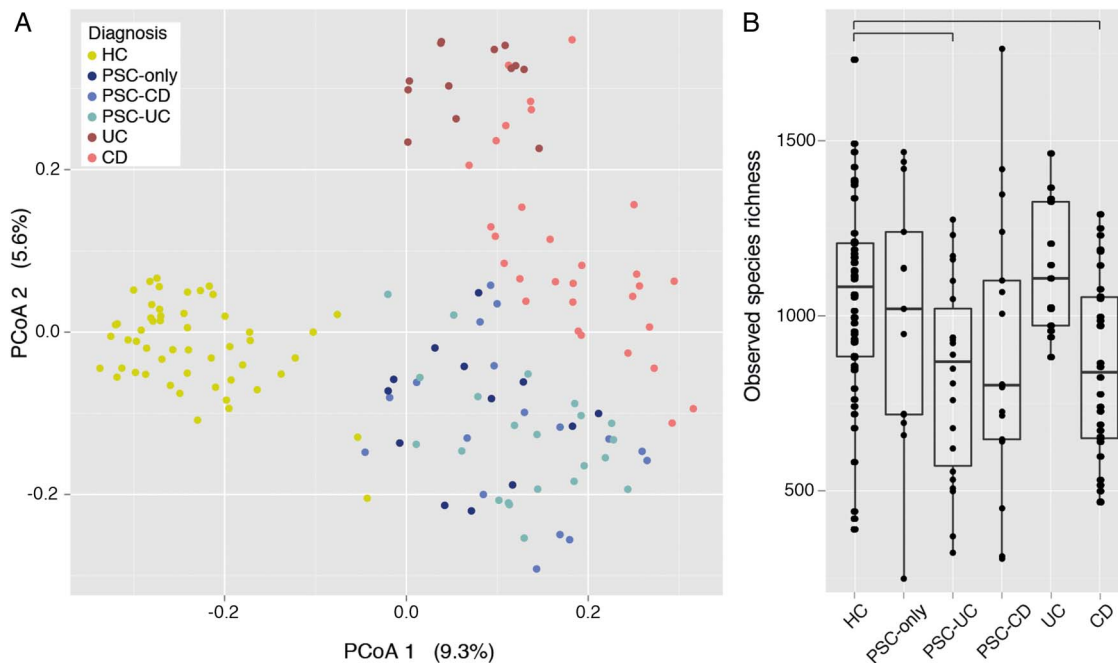
**Figure 1** Faecal microbiota variation across healthy controls (HC) and patients with primary sclerosing cholangitis (PSC) and IBD (n=147). (A) Variation in microbial community composition represented in a principal coordinates analysis (PCoA) of the Bray–Curtis dissimilarity matrix, calculated from the operational taxonomic unit (OTU)-level abundance matrix. Patients with HC, PSC and IBD are significantly different (see online supplementary table S2), while PSC subgroups are not. (B) Box plots representation of microbiota richness (number of observed OTUs per sample) distribution across HC, PSC (PSC only, PSC-ulcerative colitis (UC) and PSC-Crohn's disease (CD)) and IBD (UC and CD). HC have significantly higher species richness than all patients with PSC combined, and than PSC-UC and CD individually. Horizontal bars represent false discovery rate (FDR)-corrected p value <0.05 (Kruskal–Wallis rank sum test). The body of the box plot represents the first and third quartiles of the distribution, and the median. The whiskers extend from the quartiles to the last datapoint within 1.5× IQR, with outliers beyond represented as dots.

as the number of different OTUs observed in the sample, was decreased in patients with PSC compared with healthy controls (p value=0.026; figure 1B). Although this trend was observed in all subgroups of PSC, decreased richness was only significantly lower in PSC-UC compared with healthy controls after splitting the subgroups and multiple testing correction (FDR-adjusted p value=0.019). We also confirmed both the previously reported decreased microbiota richness in patients with CD compared with healthy controls (FDR-adjusted p value=0.0185) and the absence of reduced richness in patients with UC compared with healthy controls.

### Three genera, including *Enterococcus*, are increased in patients with PSC regardless of concomitant IBD and UDCA treatment

Both at phylum-level (figure 2A) and genus-level (figure 2B) composition, significant differences were observed between PSC and healthy controls. At the phylum level, Bacteroidetes were more abundant in patients with PSC, whereas Firmicutes were underrepresented (FDR-adjusted p value=2.097e-03 and 1.636e-06, respectively). Healthy controls had lower abundance of Fusobacteria than patients with PSC (FDR-adjusted p value=1.23e-12).
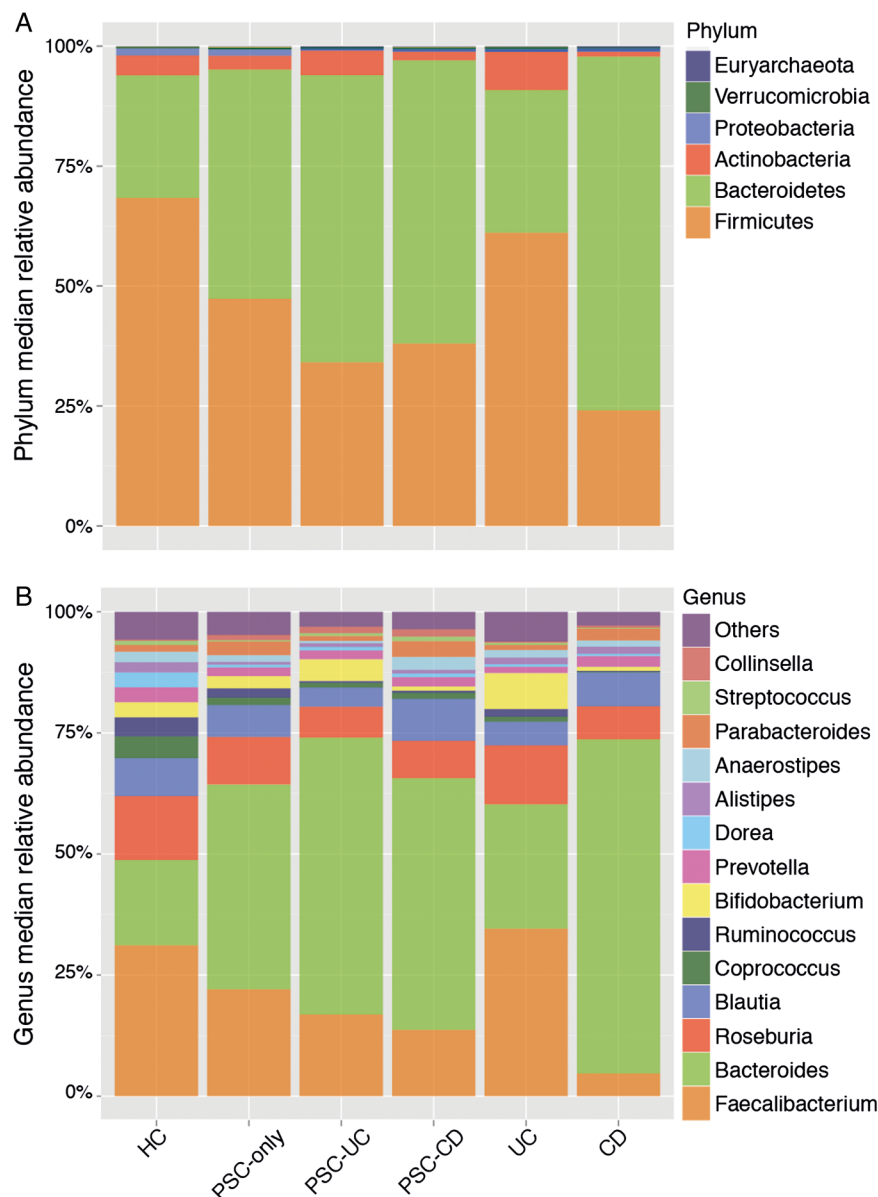
At the genus level, when compared with healthy controls, patients with PSC only displayed increased abundance of *Enterococcus*, *Fusobacterium*, *Lactobacillus*, *Morganella* and *Streptococcus* and decreased abundance of *Anaerostipes* (FDR-adjusted p values=3.76e-05, 3.76e-05, 0.0002, 0.007, 0.046 and 0.046, respectively) (figure 3A). The same signal was found for *Enterococcus*, *Fusobacterium*, *Lactobacillus* and *Streptococcus* when comparing PSC-UC or PSC-CD with

healthy controls (see online supplementary Table S3). There were four genera significantly altered in all the PSC subgroups as compared with healthy controls—*Enterococcus*, *Streptococcus*, *Lactobacillus* and *Fusobacterium*, all but the latter exclusively associated with PSC dysbiosis, and not with UC nor CD dysbiosis (figure 3B). The three bacterial genera remained significantly different between PSC and healthy controls after excluding patients taking UDCA. Most of these changes remained significant after excluding patients with cirrhosis or liver transplantation from the analysis (see online supplementary table S4). After excluding patients with antibiotic use, *Streptococcus* was no longer significantly different between patients with PSC only and healthy controls.

### The three genera microbiota signature discriminates patients with PSC from healthy controls

Eliminating genera with confounding association to IBD (without concomitant PSC) or antibiotic treatment, the microbial signature of PSC described above is composed of three genera: *Enterococcus*, *Lactobacillus* and *Fusobacterium*. All these genera are specifically overrepresented in all subgroups of patients with PSC, after excluding potential confounders (antibiotic treatment, probiotics, treatment with UDCA, liver cirrhosis and liver transplantation). We explored to what extent the observed signature could be used to correctly discriminate PSC samples. We thus used the data from the first cohort to train a J48 decision tree classifier to predict PSC versus healthy controls diagnosis, based on the abundances of the three PSC-associated genera (see online supplementary figure S2). The trained model yielded a correct classification of 95% of the subjects. Testing this model in the validation cohort, which was

**Figure 2** Faecal microbiota composition in healthy controls, patients with primary sclerosing cholangitis (PSC) (PSC-only, PSC-ulcerative colitis (UC) and PSC-Crohn's disease (CD)) and IBD (UC and CD) (n=147). (A) Phylum-level median relative abundances and (B) genus-level median relative abundances. The top 14 more abundant genera are represented and all others are summed into the category 'others'.



## The microbiota signature of PSC is conserved across PSC severity groups, even after controlling for clinical confounders

We then assessed associations between the abundances of genera included in the PSC signature described above and disease severity, both for PSC and IBD. For this purpose, patients with PSC were divided according to disease severity into stable patients, patients with cirrhosis and patients having undergone liver transplantation. Patients with IBD were divided into patients in clinical remission (HBI<4 for CD and partial Mayo<3 for UC), patients with active disease and patients with total colectomy and ileoanal pouch.

We first tested for genera significantly associated with PSC disease severity, while removing the influence of possible confounders, by performing a multivariate analysis (MaAsLin). Sex, age, BMI, smoking status, antibiotic use and UDCA treatment were used as potential clinical confounders. Besides confirming

previously reported associations such as the increase in *Methanobrevibacter* abundance with age, we confirmed that the genera *Fusobacterium*, *Enterococcus* and *Lactobacillus* were significantly decreased in healthy controls compared with all patients with PSC, regardless of disease severity, even when deconfounding for the six other variables (p value <0.001; see online supplementary table S5).

## *Enterococcus* abundance correlates to alkaline phosphatase levels, but this is not confirmed by multivariate analysis

Interestingly, the alterations in the signature genera abundances were more pronounced in patients with cirrhosis and liver transplantation in comparison with patients having stable disease (see online supplementary figure S3A), although these lacked statistical significance, probably due to the reduced sample size after splitting the groups. The same findings were observed in patients with IBD, where the differences at the genus level were more pronounced in patients with active disease or patients with an ileoanal pouch than in patients in clinical remission (see online supplementary figure S3b). Again, these differences were not significant after correction for multiple testing.

**Figure 3** Genus-level microbiota signature of primary sclerosing cholangitis (PSC) (n=147). (A) Venn diagram summarising the overlap between the lists of genera with significantly different abundances in patient groups compared with healthy controls (HC). Genera overrepresented in patients as compared with controls are in red, and underrepresented ones in blue. The lists of genera included in each of the Venn diagram's sets (labelled from A to M) can be found in online supplementary table S3. (B) Box plots showing abundances of *Fusobacterium*, *Enterococcus*, *Lactobacillus* and *Streptococcus* in HC, patients with PSC and IBD. These compose the PSC signature (ie, genera significantly overrepresented or underrepresented in all PSC subgroups compared with HC), with the exception of *Streptococcus*, which association falls below significance after excluding patients under antibiotic treatment. Horizontal bar represents false discovery rate (FDR)-corrected p value <0.05 (Mann–Whitney test). Box plot definition is provided in the legend of figure 1.



A  Genera discriminating patients from HC

L - PSC exclusive signature
 Enterococcus
 Streptococcus
 Lactobacillus
M - IBD or PSC signature
 Fusobacterium
PSC signature

A - PSC-UC signature
 6 genera + 5 genera
B - PSC-CD signature
 Butyricicoccus

B
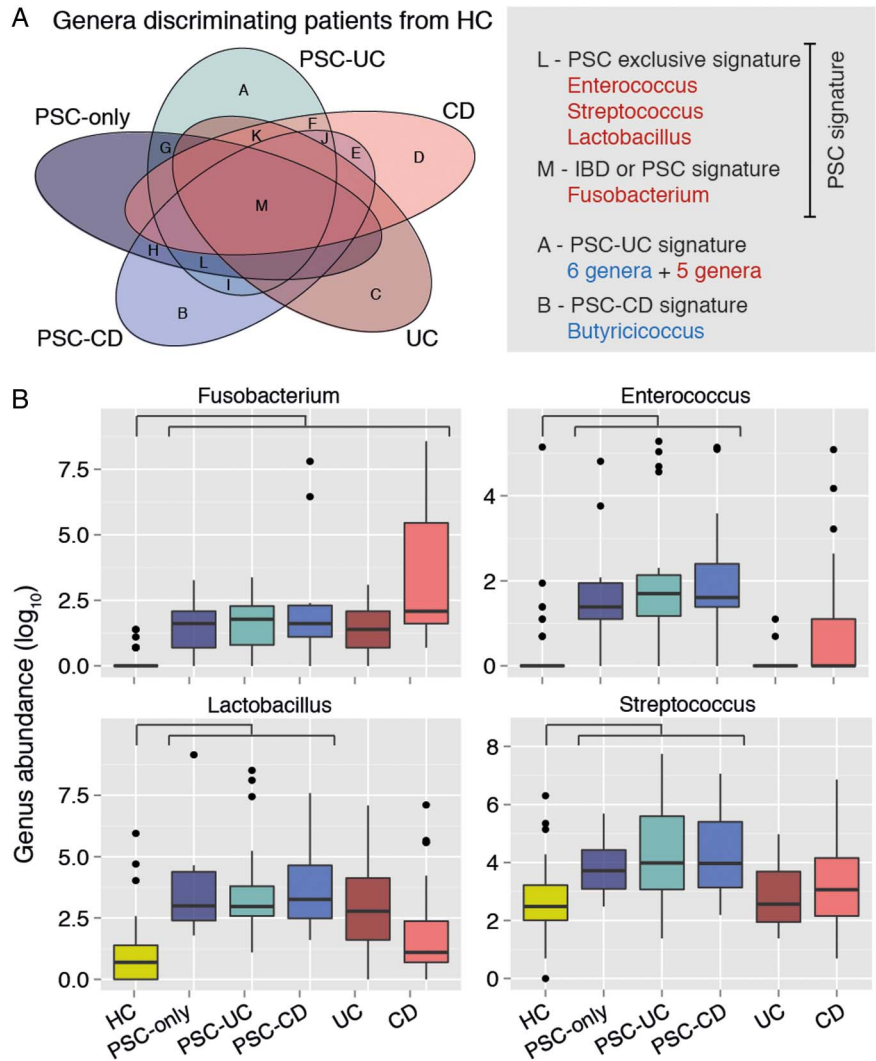
Since disease severity groups provided hints of a possible association between severity and the amplitude of the microbiota signature, we next evaluated the correlation between the three PSC-associated genera and serum γ-glutamyl transpeptidase (GGT) and alkaline phosphatase (ALP), two clinical markers of cholestasis. In fact, a persistent elevation of ALP often triggers the diagnosis of PSC. A significant positive spearman correlation was found between the abundance of *Enterococcus,* and *Lactobacillus* and GGT (see online supplementary table S6). Notably, we found a positive correlation (Spearman ρ=0.29, adjusted p value=0.012199; figure 4) between *Enterococcus* abundance and ALP, although this correlation did not withstand multivariate analysis deconfounding for sex, age, BMI, smoking status and antibiotic use (adjusted p value=0.6).

We also examined the genus *Enterococcus* at OTU (roughly comparable to species) level. There were three OTUs assigned to the genus *Enterococcus*, one of which (OTU1) was significantly different in all subgroups of patients with PSC compared with healthy controls, while showing no significant difference between patients with IBD (without PSC) and healthy controls. Furthermore, the association between OTU1 and PSC diagnosis is significant regardless of disease severity (stable, cirrhosis and liver transplantation), even when deconfounding for sex, age, faecal calprotectin, BMI, smoking, UDCA use, serum ALP levels, antibiotic use and liver transplantation. Comparably to

the genus-level analysis, OTU1 was also positively correlated with GGT, ALP and ALT (Spearman ρ=0.321, 0.258, 0.296, adjusted p value=0.001, 0.048, 0.002, respectively), but could not be confirmed in multivariate analysis. The centroid read of
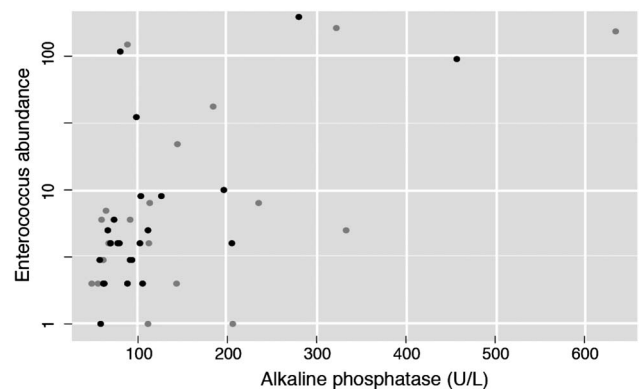


**Figure 4** Positive correlation between *Enterococcus* abundance and alkaline phosphatase (ALP), a clinical marker for cholestasis and primary sclerosing cholangitis (PSC) progression (n=87, Spearman ρ=0.29, adjusted p value=0.012199). Black dots correspond to patients with PSC without cirrhosis or liver transplantation. Grey dots correspond to patients with PSC with either cirrhosis or liver transplantation.

OTU1 had equally high similarity, but below perfect, matches to the species *Enterococcus hirae*, *E. faecium* and *E. faecalis* (see online supplementary table S7). The other two OTUs (OTU2 and OTU3) did not have much impact in the overall results.

The three genera were also decomposed by oligotyping analysis to determine whether a specific oligotype associated with PSC severity. The *Enterococcus*, *Fusobacterium* and *Lactobacillus* genus decomposed in 2, 3 and 13 olygotypes, respectively. One oligotype (OLIG2; see online supplementary figure S4) from the *Enterococcus* genus was positively correlated with serum ALP levels (Spearman ρ=0.34, adjusted p value=0.0452). The other oligotype also correlated positively with ALP serum level, but did not resist multiple testing correction (Spearman r=0.32, non-adjusted p value=0.0332).

## DISCUSSION

In the present study, we analysed the faecal microbiota associated with PSC. Because patients with PSC frequently develop concomitant IBD, we assembled a large cohort of patients with PSC, and compared their microbiota with two control groups, one being patients with IBD without PSC and the other age, gender and BMI-matched healthy volunteers. To confirm our findings, a second smaller validation cohort of patients with PSC and healthy controls was independently recruited, sampled and sequenced. We show that the faecal microbiota in patients with PSC is significantly different from that of healthy volunteers and patients with IBD. Furthermore, samples from patients with both PSC and concomitant IBD clustered together with PSC only and apart from patients with IBD without PSC. Therefore, PSC seems to shape the microbiota composition stronger than IBD, allowing for a clear separation of samples of patients with PSC according to PSC diagnosis and regardless of concomitant IBD.

PSC dysbiosis was characterised by reduced microbiota diversity. We also found a consistent signature of four genera—*Enterococcus*, *Fusobacterium*, *Lactobacillus* and *Streptococcus*—which were overrepresented in all subgroups of PSC (PSC only, PSC-UC and PSC-CD) compared with healthy controls. After taking into account antibiotic use, *Streptococcus* was no longer significantly overrepresented in patients with PSC. The same three genera (*Enterococcus*, *Fusobacterium* and *Lactobacillus*) signature was found when restricting to patients without liver cirrhosis, liver transplantation or associated IBD. We confirmed this signature with multivariate analysis, correcting for possible clinical confounders. This signature allowed discrimination of patients with PSC from healthy controls in the validation cohort with 71% accuracy.

Interestingly, the three genera forming the PSC signature have already been associated with other diseases. *Enterococcus* and *Lactobacillus* have been associated with dysbiosis in liver cirrhosis[24 25] and IBD.[26] These genera represent lactic acid-producing bacteria, with the former encoding a wide range of both intrinsic and acquired resistance to antibiotics.[27] *Fusobacterium* has also been previously associated with liver cirrhosis,[24] IBD[26 28] and colorectal cancer.[29]

It is worth mentioning that a recent study analysed the mucosa-associated microbiota of patients with PSC sampled during colonoscopy and found one sole association to PSC, the underrepresentation of an uncultured Clostridiales II.[30] Expectedly, our results differ, because the transient (faecal) microbiota differs from the adherent (mucosal) microbiota,[26] and because our cohort includes patients with PSC without concomitant IBD and the higher sample size (N=147 vs N=32) allows for higher discovery potential.

Elevated ALP serum levels are used in PSC diagnosis,[12] as well as in follow-up of the disease, as elevated levels reveal cholestasis. Growing evidence suggests that reduced levels of ALP are associated with an improved prognosis in PSC.[31 32] In our cohort, the genus *Enterococcus* positively correlated with ALP in univariate analysis, suggesting a potential link between these bacteria and disease severity. This was further illustrated by the detection of species from the genus *Enterococcus* associated with disease severity, using two different approaches to increase resolution (OTU clustering and oligotyping), although this trend will have to be confirmed in a bigger cohort. Higher ALP serum levels in patients with PSC are associated with more pronounced cholestasis due to increasing severity in biliary strictures. Bile colonisation by enteric bacteria such as *Enterococcus* spp is more common among patients with PSC with dominant strictures.[33] In the study of Pohl *et al*, *E. faecalis* and *E. faecium* were the most frequently isolated species in bile in patients with PSC with dominant strictures. In the present study, we observe an overrepresentation of the genus *Enterococcus* in faecal samples, accounting for up to 2% of the faecal bacteria in patients with PSC. It is extremely unlikely that such a high relative abundance of Enterococcus result solely from colonised bile being discharged in the intestinal tract. Our results therefore suggest that *Enterococcus* population overgrowth in the intestine occurs, independently of the colonisation of strictured bile ducts, in patients with PSC.

In particular, OTU1 of the genus *Enterococcus* was solely responsible for the genus-level positive correlation to disease severity. The centroid sequence of this OTU matches with equally high percentage identity several reference bacterial species, including *E. faecalis* and *E. faecium*. Interestingly, *E. faecalis* has already been associated with impaired intestinal permeability. Gelatinase, a metalloprotease produced by *E. faecalis*, has been shown to alter the epithelial barrier, resulting in higher susceptibility to intestinal inflammation.[34] The impairment of the epithelial barrier might allow bacterial translocation and colonisation of the bile. *Enterococcus* is the most frequently identified genus in bile culture studies, specifically *E. faecium*.[35]

PSC frequently leads to cirrhosis, and in some cases end-stage liver disease. Intestinal dysbiosis has already been described in patients with liver cirrhosis, with an increase in *Veillonella*, *Streptococcus*, *Prevotella*, *Lactobacillus* and *Fusobacterium*, among others.[24] *E. faecalis* has also been associated with liver cirrhosis.[25] In accordance with the literature, the genus *Enterococcus* was overrepresented in patients with cirrhosis from the first cohort when compared with healthy controls. The genus *Veillonella* has recently been associated with PSC.[36] We also found *Streptococcus* and *Veillonella* to be increased in patients with PSC with liver cirrhosis; however, the association to PSC diagnosis was no longer significant when excluding patients with liver cirrhosis from the analysis. At the genus level, patients with PSC with more severe disease seem to have more extreme deviations from healthy controls than patients with a stable course of disease who tend to be more similar to healthy controls. Nonetheless, no significant differences were observed at the genus level between patients with stable disease and patients with either liver cirrhosis or liver transplantation. Furthermore, the increase in *Enterococcus*, *Fusobacterium* and *Lactobacillus* was also found when comparing patients with PSC without liver cirrhosis, liver transplantation or concomitant IBD with healthy controls.

Currently, efficacious medical treatment for PSC is lacking.[12] Notwithstanding, most patients with PSC are being treated with

UDCA[37] due to some evidence pointing towards possible clinical benefits.[31] As such, around 75% of the patients with PSC in our cohort were taking UDCA. Faecal microbiota did not differ between patients with PSC taking UDCA and those not taking the medication. In addition, the three genera microbial signature is also found in patients with PSC that are not taking UDCA. Therefore, our results are not associated with treatment with UDCA.

Intestinal microbiota may play a role in the pathogenesis of IBD and dysbiosis has already been described both in CD and UC.[7] Interestingly, our results are in line with previously published data by showing a decreased α-diversity in CD while no difference was seen in UC, and a decrease of *Faecalibacterium* in CD.[7] [8] [10] The majority of patients with PSC also have IBD, such that PSC is considered to be an extra-intestinal manifestation of IBD.[38] The pathogenesis of PSC has not yet been elucidated; however, mounting evidence suggests that intestinal microbiota might also play a role in its pathogenesis. Genome-wide association studies have already identified 16 risk loci for PSC.[39] Some of these risk loci such *as FUT2* are associated with increased risk of bacterial and viral infections[1] and with microbiome composition.[40] Furthermore, germ-free Mdr2(−/−) mice have a more severe phenotype of PSC.[6]

In conclusion, the faecal microbiota of patients with PSC differs from that of healthy controls. A unique microbial signature of three genera was observed in patients with PSC, irrespective of the presence of IBD and UDCA treatment. A positive correlation was found between the genus *Enterococcus* and ALP serum levels. Our data support the hypothesis that the intestinal microbiota plays an important role in the pathogenesis of this chronic cholestatic liver disease.

## REFERENCES

1. Hirschfield GM, Karlsen TH, Lindor KD, et al. Primary sclerosing cholangitis. *Lancet* 2013;382:1587–99.
2. Alabraba E, Nightingale P, Gunson B, et al. A re-evaluation of the risk factors for the recurrence of primary sclerosing cholangitis in liver allografts. *Liver Transpl* 2009;15:330–40.
3. Olsson R, Björnsson E, Bäckman L, et al. Bile duct bacterial isolates in primary sclerosing cholangitis: a study of explanted livers. *J Hepatol* 1998;28:426–32.
4. Färkkilä M, Karvonen AL, Nurmi H, et al. Metronidazole and ursodeoxycholic acid for primary sclerosing cholangitis: a randomized placebo-controlled trial. *Hepatology* 2004;40:1379–86.
5. Seidel D, Eickmeier I, Kühl AA, et al. CD8T cells primed in the gut-associated lymphoid tissue induce immune-mediated cholangitis in mice. *Hepatology* 2014;59:601–11.
6. Tabibian JH, O'Hara SP, Trussoni CE, et al. Absence of the intestinal microbiota exacerbates hepatobiliary disease in a murine model of primary sclerosing cholangitis. *Hepatology* 2016;63:185–96.
7. Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 2014;146:1489–99.
8. Sokol H, Pigneur B, Watterlot L, et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA* 2008;105:16731–6.
9. Joossens M, Huys G, Cnockaert M, et al. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* 2011;60:631–7.
10. Willing BP, Dicksved J, Halfvarson J, et al. A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology* 2010;139:1844–54.e1.
11. Machiels K, Joossens M, Sabino J, et al. A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. *Gut* 2014;63:1275–83.
12. Lindor KD, Kowdley KV, Harrison ME. ACG clinical guideline: primary sclerosing cholangitis. *Am J Gastroenterol* 2015;110:646–59.
13. Dignass A, Eliakim R, Magro F, et al. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: definitions and diagnosis. *J Crohns Colitis* 2012;6:965–90.
14. Van Assche G, Dignass A, Panes J, et al. The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Definitions and diagnosis. *J Crohns Colitis* 2010;4:7–27.
15. Kozich JJ, Westcott SL, Baxter NT, et al. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 2013;79:5112–20.
16. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;27:2957–63.
17. Edgar RC, Haas BJ, Clemente JC, et al. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;27:2194–200.
18. Wang Q, Garrity GM, Tiedje JM, et al. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007;73:5261–7.
19. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;26:2460–1.
20. Eren AM, Maignien L, Sul WJ, et al. Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* 2013;4:1111–9.
21. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 2013;8:e61217.
22. Tickle TWL, Yiren Lu, Huttenhower C. Multivariate association of microbial communities with rich metadata in high-dimensional studies. (In progress).
23. Hall M, Frank E, Holmes G, et al. The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter* 2009;11:10–18.
24. Qin N, Yang F, Li A, et al. Alterations of the human gut microbiome in liver cirrhosis. *Nature* 2014;513:59–64.
25. Chen Y, Yang F, Lu H, et al. Characterization of fecal microbial communities in patients with liver cirrhosis. *Hepatology* 2011;54:562–72.
26. Gevers D, Kugathasan S, Denson LA, et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 2014;15:382–92.

27  Fisher K, Phillips C. The ecology, epidemiology and virulence of Enterococcus. *Microbiology* 2009;155(Pt 6):1749–57.

28  Strauss J, Kaplan GG, Beck PL, *et al*. Invasive potential of gut mucosa-derived Fusobacterium nucleatum positively correlates with IBD status of the host. *Inflamm Bowel Dis* 2011;17:1971–8.

29  Kostic AD, Gevers D, Pedamallu CS, *et al*. Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res* 2012;22:292–8.

30  Rossen NG, Fuentes S, Boonstra K, *et al*. The mucosa-associated microbiota of PSC patients is characterized by low diversity and low abundance of uncultured Clostridiales II. *J Crohns Colitis* 2015;9:342–8.

31  Al Mamari S, Djordjevic J, Halliday JS, *et al*. Improvement of serum alkaline phosphatase to <1.5 upper limit of normal predicts better outcome and reduced risk of cholangiocarcinoma in primary sclerosing cholangitis. *J Hepatol* 2013;58: 329–34.

32  Lindström L, Hultcrantz R, Boberg KM, *et al*. Association between reduced levels of alkaline phosphatase and survival times of patients with primary sclerosing cholangitis. *Clin Gastroenterol Hepatol* 2013;11:841–6.

33  Pohl J, Ring A, Stremmel W, *et al*. The role of dominant stenoses in bacterial infections of bile ducts in primary sclerosing cholangitis. *Eur J Gastroenterol Hepatol* 2006;18:69–74.

34  Steck N, Hoffmann M, Sava IG, *et al*. Enterococcus faecalis metalloprotease compromises epithelial barrier and contributes to intestinal inflammation. *Gastroenterology* 2011;141:959–71.

35  Kwon W, Jang JY, Kim EC, *et al*. Changing trend in bile microbiology and antibiotic susceptibilities: over 12 years of experience. *Infection* 2013;41:93–102.

36  Kummen M, Holm K, Anmarkrud JA, *et al*. The gut microbial profile in patients with primary sclerosing cholangitis is distinct from patients with ulcerative colitis without biliary disease and healthy controls. *Gut* Published Online First: 17 Feb 2016. doi:10.1136/gutjnl-2015-310500

37  Boonstra K, Weersma RK, van Erpecum KJ, *et al*. Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 2013;58:2045–55.

38  Ott C, Scholmerich J. Extraintestinal manifestations and complications in IBD. *Nat Rev Gastroenterol Hepatol* 2013;10:585–95.

39  Liu JZ, Hov JR, Folseraas T, *et al*. Dense genotyping of immune-related disease regions identifies nine new risk loci for primary sclerosing cholangitis. *Nat Genet* 2013;45:670–5.

40  Wacklin P, Tuimala J, Nikkilä J, *et al*. Faecal microbiota composition in adults is associated with the FUT2 gene determining the secretor status. *PLoS ONE* 2014;9: e94863.