Crohn associated microbial communities associated to colonic mucosal biopsies in patients of the western Mediterranean

Roberto Vidal [a,1], Daniel Ginard [b,1], Sam Khorrami [b], Merit Mora-Ruiz [c], Raul Munoz [c], Marcela Hermoso [a], Sara Díaz [c], Ana Cifuentes [c], Alejandro Orfila [c], Ramon Rosselló-Móra [c,*]

[a] Institute of Biomedical Sciences, Faculty of Medicine, Universidad de Chile, Santiago, CL 8380453, Chile
[b] Department of Gastroenterology and Palma Health Research Institute, Hospital Universitario Son Espases, Palma de Mallorca, Spain
[c] Institut Mediterrani d'Estudis Avançats (IMEDEA; CSIC-UIB), 07190 Illes Balears, Spain

ARTICLE INFO

ABSTRACT

Next generation sequencing approaches allow the retrieval of several orders of magnitude larger numbers of amplified single sequences in 16S rRNA diversity surveys than classical methods. However, the sequences are only partial and thus lack sufficient resolution for a reliable identification. The OPU approach used here, based on a tandem combination of high quality 454 sequences (mean >500 nuc) applying strict OTU thresholds, and phylogenetic inference based on parsimony additions to preexisting trees, seemed to improve the identification yields at the species and genus levels. A total of thirteen biopsies of Crohn-diagnosed patients (CD) and seven healthy controls (HC) were studied. In most of the cases (73%), sequences were affiliated to known species or genera and distinct microbial patterns could be distinguished among the CD subjects, with a common depletion of *Clostridia* and either an increased presence of *Bacteroidetes* (CD1) or an anomalous overrepresentation of *Proteobacteria* (CD2). *Faecalibacterium prausnitzii* presence was undetectable in CD, whereas *Bacteroides vulgatus–B. dorei* characterized HC and some CD groups. Altogether, the results showed that a microbial composition with predominance of *Clostridia* followed by *Bacteroidetes*, with *F. prausnitzii* and *B. vulgatus–B. dorei* as major key bacteria, characterized what could be considered a balanced structure in HC. The depletion of *Clostridia* seemed to be a common trait in CD.

© 2015 Elsevier GmbH. All rights reserved.

Introduction

Microbial diversity measurements based on environmental 16S rRNA genes have permitted the recognition of the vast diversity of as yet uncultured microorganisms in environmental samples [56]. Next generation sequencing (NGS) techniques produce orders of magnitude higher numbers of sequences than the conventional techniques used in microbial molecular ecology, and they have been thoroughly used to reveal diversity in environmental samples (e.g. [4,9,11,31,47,54]). In general, the identification of environmentally occurring discrete 16S rRNA groups (often assumed to be

species) had been based on clustering sequences into operational taxonomic units (OTUs [35]) using an identity threshold of 97% [4,11,31,47] or even lower (95%; [54]). The 97% threshold has been commonly used in microbial molecular ecology to circumscribe putative species based on the observation made by Stackebrandt and Goebel [42]. However, for taxonomic purposes this value was later raised to 98.7% [41], which was considered to be a more adequate minimum threshold for this category [56], although it had not been readily implemented in the molecular ecology of prokaryotes. Embracing sequences at 97% would be too conservative and can lead to underestimation of diversity, since distinct species of the same genus may cluster together. However, the major pitfall of the NGS approach is that the length of the sequences is too short for taxonomic identifications at the species level [56], and the almost complete sequence of the 16S rRNA gene would still be desirable. Generally, 454 pyrosequencing in the past has rendered sequence lengths of <300 nuc (e.g. [9,11,31,47,54]). However, the results obtained with such short sequences seemed to be robust enough to mirror the observations made with classical techniques

* Corresponding author at: Grup de Microbiologia Marina, Department of Ecology and Marine Resources, Institut Mediterrani d'Estudis Avançats (CSIC-UIB), C/Miquel Marqués 21, 07190 Esporles, Illes Balears, Spain. Tel.: +34 971 611 826; fax: +34 971 611 761.
 E-mail address: rossello-mora@uib.es (R. Rosselló-Móra).
[1] Both authors contributed equally.

[11,31]. Lately, improvements in the methodology have allowed much longer amplicon stretches (up to 800 nucleotides with means of >550 nucleotides; [26]), which has led to a more robust identification power due to the larger information content. However, due to the higher costs of 454 pyrosequencing, Illumina is becoming the method of choice for most of the new studies on environmental microbial diversity, despite some evidence that both the method and the different regions studied may not produce robust results comparable to conventional methods [4]. For these reasons, among the different NGS techniques, 454 may still be the best choice for amplicon analyses due to the larger stretches produced with relative low error rates [20].

The human gut microbiota harbors one order of magnitude more cells than that of the somatic and germ cells of the complete body [1], with its complex communities actively participating in host homeostasis [12] and playing important roles in nutrient metabolism and protection [39]. The gut epithelium is covered by a protective mucus layer, which can be damaged in inflammatory bowel diseases (IBD) inducing a disintegration of the web structure [45], finally leading to a modification of the microbial communities with implications for host health. Among the IBD, Crohn disease (CD) is a chronic disorder characterized by patchy inflammation of the gastrointestinal tract. The specific causes remain unclear, but clinical and experimental data consider it to be a consequence of uncontrolled intestinal inflammation in response to elusive environmental enteric microbiota, and immuno-regulatory factors in genetically susceptible individuals [3]. The major consequences are inflammation and ulceration of the intestinal lining [25], with an important reduction of the mucus layer and important changes in the colonic microbiomes [45].

In this work, we present the study of the microbiota attached to the intestinal mucosa of 13 CD biopsies together with 7 healthy control biopsies (HC) by means of the 454 sequencing approach of high quality amplified 16S rRNA gene sequences (mean >550 nuc). In addition, an identification approach was applied based on the recognition of the operational phylogenetic units (OPUs) that are based on phylogenetic inference instead of sequence identity clustering, in order to reflect the diversity observed better [9,26]. The work aimed to evaluate the use of the OPU approach based on 454 technology in order to increase the resolution power at lower taxonomic categories (i.e. genus and species) and compare the associated microbiomes with the colon tissues of HC and CD.

Materials and methods

Patients and samples

Patients with a Crohn disease diagnosis that had to be submitted to colonoscopy for medical revision were recruited for the study between August 2011 and March 2012 (Table 1; control individuals were identified with an S and CD with a C following their collection number). Control subjects were individuals submitted to colonoscopy due to constipation, colorectal cancer screening or anal bleeding, with normal colonoscopy inspection and where inflammatory or other bowel disorders were discarded. None of the subjects were exposed to antibiotics for 1 month prior to colonoscopy, and they all signed an informed consent. Demographics and clinical characteristics (disease localization, inflammatory activity, behavior, medication and surgical history) were collected (Table 1). For practical reasons, most of the CD patients selected had previously been ileal or ileocecal resected as they were colonoscopied for routine analyses, and a few non-resected patients were also included. In all cases, colectomies or ileostomies were exclusion criteria. The CD biopsies were taken from colon intestines with a healthy aspect and avoiding ulceration or lesions in the mucosa.

Subjects received standard bowel cleansing with a polyethylene glycol preparation the evening before colonoscopy. Biopsies of approximately 1 mg and 2 mm^3 were collected from the colon using biopsy forceps during the colonoscopy. Biopsies were immediately placed in sterile tubes and stored at $-80\,°C$ for DNA/RNA extraction. The Balearic Islands' Ethical Committee approved the study.

DNA extraction, PCR amplification and pyrosequencing

Total DNA was extracted using the FastDNA® SPIN Kit (MP Biomedicals), following the manufacturer's indications, and a FastPrep®-24 (MP Biomedicals) instrument for homogenization. DNA was quantified using a NanoDrop® Spectrophotometer (Thermo Fisher Scientific Inc.), then stored at $-20\,°C$ and $20\,\mu g\,\mu L^{-1}$. For amplification, bacterial primers (forward GM3 5′-AGAGTTTGATCMTGGC-3′ and reverse 907RM 5′-CCGTCAATTCMTTTGAGTTT-3′ [27,28] were used in a 30-cycle PCR at $50\,°C$ (Tm). For pyrosequencing, a secondary PCR incorporated tags and linkers into the amplicons using a 1:10 dilution of the original products as template in a 5-cycle PCR with the same conditions. Primers GM3-PS and a variant of 907-PS were used for *Bacteria* (Supplementary Table S1). Products were purified using MSB® Spin PCRapace (INVITEK), following the manufacturer's instructions. The concentration of the barcoded-amplicons was measured with a NanoDrop and, finally, an equimolar mixture of the amplicons was sent to a sequencing company (LifeTechnologies, Valencia, Spain; or Macrogen Inc., Seoul, Korea). The samples were sequenced using 454 GS-FLX+ Titanium technology. The set of sequences has been deposited at the ENA sequence repository under the project accession numbers PRJEB6107 and ERP005574.

Sequence trimming, chimera check, OTU (operational taxonomic unit) clustering

Data was processed using the Mothur pipeline [38]. Briefly, low-quality sequences were removed (<300 nuc with a window size and average quality score of 25, no ambiguities and no mismatches in reads with primers and barcodes were allowed, and a maximum homopolymer of 8 nucleotides). The 10-nuc barcode was used for sample identification. Chimeras were removed with the application Chimera Uchime implemented in Mothur. Sequences were clustered into OTUs (i.e. unique sets of sequences clustered equal to or above a certain identity threshold) at 99% using the UCLUST tool included in QIIME [6]. The most abundant read for each OTU was selected as representative.

Phylogenetic affiliation and OPU (operational phylogenetic unit) design

OTU representatives were added to the non-redundant SILVA REF111 database using the ARB program package [22,33]. Sequences were aligned with SINA using the LTP111 database as a template [32,55]. Alignments were manually inspected and improved, and sequences were added with the ARB parsimony tool to a default tree containing approximately 244,000 sequences of *Bacteria*. The closest relative sequences of an acceptable quality (almost full length with low indetermination and homopolymer occurrence) were selected and merged with the LTP111. The neighbor-joining algorithm was used for the final tree reconstruction. OTU representatives were added to the final tree with the parsimony tool. Sequences were grouped in OPUs based on the manual inspection of the tree [50]. An OPU was considered as the smallest clade containing one or more amplified sequences affiliating together with reference sequences available in the public repositories. When possible, the OPUs should include a type strain sequence present in the LTP database [55], and for identity values

Table 1
Demographic and clinical characteristics.

ID	Group	Birth year	Sex[a]	Current smoker	Diagnosis year	Localization[*]	Behavior[**]	Surgery	Medication (at least 3 months)	Clinical situation[***]	Colonoscopy
2C		1963	M	No	1983	L3	B3	Ileal resection	Immunosuppressive agent	Remission	Normal colon and ileum
3C	CD1	1967	M	No	2003	L1	B2	Ileocolic resection	Immunosuppressive agent	Remission	Normal colon. Inflammatory activity at ileum
10C		1949	M	No	1990	L1	B1	Ileocolic resection	5-aminosalicylates	Clinical activity	Normal colon. Anastomotic stenosis with ulcers
12C		1978	M	No	1997	L1	B2	No	Immunosuppressive agent	Remission	Normal colon. Ileocecal valvular stenosis with ulcers
24C		1989	F	No	2009	L1	B3	Ileocolic resection	Immunosuppressive agent	Remission	Normal colon. Anastomotic stenosis. Inflammatory activity at ileum
4C	CD2	1955	F	No	1990	L1	B2	Ileocolic resection	5-aminosalicylates	Remission	Normal colon. Anastomotic stenosis. Inflammatory activity at ileum.
5C		1992	M	No	2005	L3	B2	Ileocolic resection	Immunosuppressive agent	Clinical activity	Normal colon. Anastomotic stenosis
6C		1958	M	Yes (10 cig)	2011	L2	B2	Ileocolic resection	5-aminosalicylates	Remission	Normal colon and ileum
8C		1974	F	Yes (5 cig)	1997	L3	B3	No	No	Clinical activity	Inflammatory activity at colon
9C		1976	F	Yes (10 cig)	1984	L1	B3	Ileocolic resection	TNF antagonist	Clinical activity	Inflammatory activity at ileum
1C	CD3	1975	M	No	2002	L3	B3	Ileocolic resection	Immunosuppressive agent	Clinical activity	Anastomotic stenosis with ulcers
7C	CD4	1983	F	Yes (5 cig)	2005	L3	B2	No	TNF antagonist *Steroids 1 month before*	Remission	Inflammatory activity at colon and stenosis
11C		1977	F	Yes (10 cig)	1998	L3	B1	No	Budesonide *Immunosuppressive 2 months before*	Clinical activity	Normal colon and ileum
14S		1952	M	No	Control					Constipation	Normal colon and ileum
15S		1965	M	No	Control					Constipation	Normal colon
16S		1945	M	Yes (10 cig)	Control					Hemorrhoidal bleeding	Normal colon
17S	HC	1953	M	No	Control					Hemorrhoidal bleeding	Normal colon and ileum
20S		1936	F	No	Control					CCR screening	Normal colon and ileum
22S		1938	M	No	Control					Hemorrhoidal bleeding	Normal colon
23S		1961	F	No	Control					CCR screening	Normal colon

[a] Sex: M-male; F-female.

[*] Localization: L1-Ileal; L2-Colonic; L3-Ileocolonic.

[**] Behavior: B1-inflammatory; B2-stricture; B3-penetrant disease.

[***] For the control samples the reasons for colonoscopy are indicated.

Table 2
Sample distribution in groups based on clustering analysis, their number of OTUs and OPUs, and the main diversity indices..

Study group	Sample	Nr of sequences	Nr OPUs	Dominance-D index	Shannon–Weiner index	Chao-1	Nr OTUs
	24C	8706	66	0.250	2.209	79.6	178
	2C	7631	84	0.194	2.482	89.1	170
	10C	4801	59	0.142	2.542	62.1	113
CD1	12C	10,414	75	0.132	2.581	78.0	190
	3C	7552	98	0.288	2.326	123.5	235
	Mean	**7820**	**76**	**0.201**	**2.428**	**86.5**	**177**
	SD	2045	15	0.068	0.156	22.9	44
	5C	3048	106	0.044	3.697	107.7	240
	9C	4037	75	0.131	2.865	84.0	192
CD2	6C	6008	86	0.060	3.293	91.6	280
	8C	6952	79	0.050	3.359	80.4	264
	4C	4699	101	0.090	3.204	121.0	249
	Mean	**4949**	**89**	**0.075**	**3.284**	**97.0**	**245**
	SD	1553	14	0.036	0.299	17.1	33
CD3	**1C**	**12,030**	**59**	**0.099**	**2.601**	**74.0**	**239**
	14S	7047	100	0.103	2.953	113.1	183
	15S	2921	79	0.167	2.444	82.9	115
	16S	4617	110	0.060	3.485	121.7	257
	17S	3772	93	0.094	3.173	94.9	184
HC	20S	6182	115	0.124	3.062	117.0	251
	22S	6223	76	0.203	2.447	80.7	172
	23S	5171	129	0.080	3.422	136.8	312
	Mean	**5133**	**100**	**0.120**	**2.990**	**106.7**	**211**
	SD	1470	19	0.047	0.421	21.1	66.0
	7C	8917	105	0.110	2.974	112.8	264
CD4	11C	11,107	85	0.101	2.871	110.5	296
	Mean	**10,012**	**95**	**0.105**	**2.923**	**111.7**	**280**
	SD	1549	14	0.006	0.073	1.6	23

>98.7% with type strain sequences the amplicons were considered to belong to the same species using this conservative threshold, as previously recommended [41]. On the other hand, for identity values <98.7% and >94.5% with the closest relative type strain to the 16S rRNA gene sequence of the same OPU, the amplicons were considered for the same genus [56] but from a different unclassified species.

Statistical analyses

Data was coded as an entry matrix with the number of sequences detected for each OPU and sample. For normalization, absolute numbers were transformed into percentages with respect to the total sequenced stretches for each sample. Rarefaction curves and diversity indices were calculated using the PAST software v1.82b [14]. Clustering analyses and multidimensional scaling (MDS) were performed with the PRIMER 5 program version 5.2.8 (PRIMER-E Ltd., UK). For similarity analyses, data were transformed with Log $(x+1)$ using the Bray–Curtis similarity measure. One-way analysis of variance was performed using STATISTICA version 7.1 software (StatSoft, Inc.). Prior to analyses, data were arcsin(square root(x)) transformed to homogenize group variances, and this ANOVA assumption was verified using Cochran's C test. If ANOVA was significant, post hoc multiple comparisons of group means were performed using the Tukey's honestly significant difference test. Principal component analysis (PCA) was computed with MATLAB (Mathworks).

Results

Subjects, samples, sequences and OPU distribution

All samples rendered bacterial but not archaeal amplicons (using primers and protocols published elsewhere [50]; data not shown). Pyrosequencing generated approximately 1.5×10^5 partial sequences (75% >500 nuc) that had a mean per sample of 6592 ($\pm$2622), with 12,030 (for CD patient 1C) and 2921 (for control patient 15S) being the highest and lowest values (Table 2). After

trimming (at 99.5% quality) and removing chimeras and short sequences (<300 bp) the dataset was reduced to 86.7%. For each patient, the sequences were grouped in OTUs at the 99% sequence identity level, giving a mean of 216 ($\pm$55) OTUs for all samples (Table 2 and Supplementary Tables S2 and S3). The phylogenetic inference using the OTU representatives generated a total of 338 distinct OPUs with a mean of 90 ($\pm$19) for all samples. The OPU approach reduced the diversity observed by OTUs in each sample (Fig. 1), but produced results of higher confidence. In addition, the OPU approach discriminated better the HC samples with higher diversity yields from the CD with reduced diversity. Approximately 10% of the OPUs harbored >1000 sequences, whereas approximately 80% of the OPUs contained >10 sequences. For the current study, all OPUs harboring <10 sequences, as well as those occurring in a single sample, were discarded.

Comparison between community structures among samples

CD showed a mean of 83 ($\pm$16) OPUs, whereas HC samples showed a mean of 101 ($\pm$19). Accordingly, CD rarefaction curves (Fig. 1) saturated faster than HC. Clustering analysis (supplementary Figs. S1 and S2) grouped HC samples homogeneously. On the other hand, CD samples were diverse and four different sample sets could be observed represented by CD1 (2C, 3C, 10C, 12C and 24C), CD2 (4C, 5C, 6C, 8C and 9C), CD3 (only composed of sample 1C), and CD4 (7C and 11C). The latter clustered together with the HC samples. The four CD sample sets exhibited saturation curves according to the diversity indices (Table 2). CD1 and CD3 showed the lowest diversity indices (Shannon–Weiner of approximately 2.5) together with a faster saturation in the rarefaction curves (Fig. 1). CD2 and CD4 showed similar diversity indices (Shannon–Weiner of approximately 3.0) and rarefaction trends as for HC. The OPU coverage was estimated in all cases to be between 80% and 90% of the total expected diversity.

Sequences affiliated mostly with 5 major phyla (*Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, *Firmicutes*, and *Fusobacteria*; Supplementary Table S4), and 11 classes (*Gamma-*, *Beta-*, *Alpha-*, and *Deltaproteobacteria*, *Bacteroidia*, *Flavobacteria*, *Clostridia*,
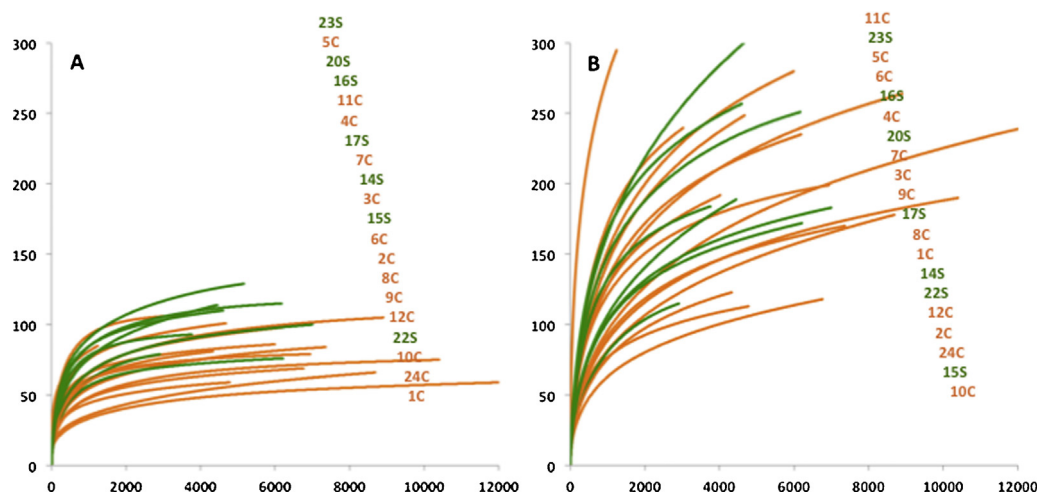
Fig. 1. Reduction in diversity by means of grouping OTUs into OPUs after phylogenetic inference using the ARB parsimony tool [22]. (A) rarefaction curves based on the OPUs detected and their occurrence with respect to the sequence data for each sample. (B) rarefaction curves based on the OTUs detected and their occurrence with respect to the sequence data for each sample.

Fusobacteriia, *Negativicutes*, *Bacilli*, and *Actinobacteria*), as well as in very small numbers with the uncultured clade OD1, *Lentisphaera* and cyanobacteria. In HC and CD4, *Firmicutes* dominated followed by *Bacteroidetes* (Fig. 2 and Supplementary Table S5). CD1 was dominated by *Bacteroidetes*, CD2 by *Proteobacteria*, and CD3 by *Firmicutes*. More precisely, HC and CD4 patients exhibited a dominance of the class *Clostridia*, followed by *Bacteroidia*. CD1 showed an inverted dominance with higher numbers of *Bacteroidia* followed by *Clostridia*. CD2 was strongly dominated by members of *Gamma-* and *Alphaproteobacteria*. The single sample of CD3 showed a clear dominance of *Clostridia* followed by *Fusobacteriia* (Fig. 3). Distinctiveness by multiple ANOVA was taken for values $p < 0.05$ (Supplementary Table S6). Groups CD1 and HC exhibited significant differences for *Proteobacteria* and *Firmicutes* (<0.009295), with special relevance for *Clostridia* (<0.005311), but not for *Actinobacteria* and *Bacteroidetes*. CD2 and HC had different values for four phyla (<0.018473) and classes detected, except for *Bacilli*. *Erysipelotrichia* and *Fusobacteriia* did not show any difference between the samples.

Detailed analysis of the identity of the most relevant OPUs discriminating the samples

PCA (Supplementary Fig. S3) allowed the recognition of the major OPUs influencing the discrimination between the samples. CD4 and HC appeared equally influenced by the same set of OPUs, whereas CD2 and CD3 showed a different pattern. Thirty-three OPUs formed the HC core microbiome (Table 3). Five affiliated with *Proteobacteria*, seven with *Bacteroidetes* and twenty-one with *Firmicutes* (nineteen of them to *Clostridia*). The most relevant OPUs were OPU-086 (*Bacteroides dorei–B. vulgatus*), and OPU-098 (*B. fluxus–B. helcogenes–B. rodentium–B. uniformis*) of the phylum *Bacteroidetes*; OPU-290 (*F. prausnitzii*) and OPU-215 (*Blautia wexlerae*) of the phylum *Firmicutes*; and OPU-001 (*Escherichia–Shigella* sp.) of the *Gammaproteobacteria* (Table 3 and Supplementary Tables S4 and S7).

Despite their presence in all samples, *Proteobacteria* strongly predominated in CD2 (Fig. 2) and the ten most abundant OPUs (supplementary Table S8) together represented almost 44% of the total diversity, contrary to the other samples in which *Proteobacteria* never reached 5% in the HC or 13% in the remaining CD samples. The most important OPU in CD was OPU-001 affiliating with *Escherichia–Shigella* ranging from 5.9% to 7.4% of their total

diversity. Some common OPUs to all samples, present in similar abundances, affiliated with *Sutterella* (OPUs 046-1 and 046-2) and *Cupriavidus* (OPU-028), and to a lesser extent with *Parasutterella* (OPU-047) of *Betaproteobacteria*, and *Haemophilus* (OPU-014) of *Gammaproteobacteria*. OPU-085 and OPU-070-2 (*Reyranella massiliensis* and *Afipia* spp. respectively, both from *Alphaproteobacteria*) were relevant in the CD samples.

Remarkably, *Bacteroidetes* was present in both CD1 and HC (Fig. 2). Especially abundant was the class *Bacteroidia* in which the most relevant OPU was OPU-086 affiliating with the *B. dorei–B. vulgatus* clade of the family *Bacteroidaceae* (Table 3 and supplementary Table S7), although it occurred in minor abundance in CD2 (about 3.4%) and was absent in CD3. In the latter, only a single bacteroidete was detected (OPU-102 affiliating with *Prevotella copri* with values <0.01%). The ten most abundant OPUs in CD1 together represented approximately 52%, whereas in HC and CD4 they ranged between 34% and 37%, respectively, and in CD2 up to 11% (Supplementary Table S8).

Firmicutes was the most relevant group especially in CD3, CD4 and HC (Fig. 2), in which the ten most abundant OPUs (supplementary Table S8) represented at least 32% of the total diversity. The most abundant classes were *Clostridia* and *Fusobacteria*, followed by *Erysipelotrichia* (also considering cluster XVI of *Clostridia* [7,49]), *Negativicutes* and *Bacilli*. OPU-290 (*F. prausnitzii*) was most relevant in both HC and CD4, followed by OPU-215 (*Blautia wexlerae*; supplementary Tables S5 and S8). Both OPUs were either absent or of low relevance in CD1, CD2 and CD3. On the contrary, OPU-142 (*C. ramosum*), OPU-148 (*Fusobacterium canifelium–F. nucleatum*), OPU-152 (*F. ulcerans–F. varium*), OPU-212 (uncultured clostridium (DQ795704)), and OPU-223 (uncultured clostridium (DQ071475)) were relatively abundant in the CD samples, and very low or absent in CD4 and HC samples. Altogether the picture observed indicated that HC and CD4 exhibited a similar content of *Firmicutes* and *Bacteroidetes* (Fig. 2), whereas CD1 to CD3 exhibited modified structures by means of either overrepresentation of *Bacteroidetes* (CD1), *Proteobacteria* (CD2) or *Firmicutes* (CD3).

Discussion

Biopsies of 13 CD and 7 control patients not suffering from CD were studied by a pyrotagging approach. The pyrosequencing study was considered to be high quality as the sequences were far larger (>300 nuc, with a mean of approximately 580 ± 124 nuc,

Table 3
Core microbiome of healthy individuals in comparison with the corresponding abundances in CD patients. The core OPUs are considered to be those present in at least 6 out of 7 individuals. The relative abundances of each OPU are referenced to the total sequences for each sample (the complete matrix is given in supplementary Table S7). The inset colors indicate: blue, zero; yellow to red increasing values above 0.1%. In brackets after the OPU number indicates the affiliation to a class and phylum. Brackets after the species names indicate the sequence entry number.

	CD1					CD2					CD3	CD4		HC						
---	24C	2C	10C	12C	3C	5C	9C	6C	8C	4C	1C	11C	7C	16S	20S	14S	17S	15S	22S	23S
OPU-001 (*Gammapr., Proteobacteria*) *Escherichia coli* (X80725); *Shigella dysenteriae* (X96966)	2.15	8.63	12.30	11.30	0.63	12.52	7.57	7.05	6.74	5.65	5.85	0.47	0.04	0.09	1.20	5.47	0.77	0.14	0.51	0.00
OPU-014 (*Gammapr., Proteobacteria*) *Haemophilus parainfluenzae* (AY362908); *H. sputorum* (JF506642)	0.01	0.66	0.10	2.13	0.00	0.51	1.66	0.43	5.06	0.58	0.10	0.01	0.00	1.72	1.85	0.01	0.51	0.76	0.29	0.00
OPU-028 (*Betapr., Proteobacteria*) *Cupriavidus necator* (AF191737); *Ralstonia pickettii* (AY741342)	0.16	0.21	0.15	0.03	0.15	9.68	2.60	9.74	6.84	4.86	0.01	0.01	0.15	0.35	0.57	0.36	0.37	0.07	0.10	0.04
OPU-067 (*Alphapr., Proteobacteria*) *Mesorhizobium plurifarium* (Y14158)	0.03	0.54	0.46	0.12	0.36	7.63	3.62	7.53	7.68	2.45	0.03	0.02	0.17	0.04	0.25	0.14	0.35	0.00	0.18	0.06
OPU-070-1 (*Alphapr., Proteobacteria*) *Agromonas oligotrophica* (D78366); *Bradyrhizobium japonicum* (U69638)	0.17	0.17	0.35	0.16	0.05	4.34	5.07	4.34	3.16	3.46	0.03	0.02	0.13	0.09	0.35	0.11	0.19	0.07	0.05	0.00
OPU-086 (*Bacteroidia, Bacteroidetes*) *Bacteroides dorei* (AB242142); *B. vulgatus* (AJ867050)	47.89	41.44	29.07	28.01	1.10	0.00	0.00	0.99	0.00	16.40	0.00	21.18	18.03	14.16	29.64	21.99	13.55	24.44	39.78	3.94
OPU-089 (*Bacteroidia, Bacteroidetes*) *Bacteroides faecis* (GQ496624); *B. finegoldii* (AB222699); *B. thetaiotaomicron* (AE015928)	10.22	2.53	0.81	0.22	0.14	0.00	0.00	0.00	0.00	1.07	0.00	2.29	1.39	0.04	0.40	0.86	0.35	0.00	2.43	0.56
OPU-098 (*Bacteroidia, Bacteroidetes*) *Bacteroides fluxus* (AB490802); *B. helcogenes* (AB200227); *B. rodentium* (AB531489); *B. uniformis* (AB050110)	2.29	0.00	0.00	0.18	0.07	0.00	0.00	0.00	0.00	0.92	0.00	0.54	6.81	0.85	0.85	0.94	2.00	0.41	2.84	1.25
OPU-109 (*Bacteroidia, Bacteroidetes*) *Parabacteroides distasonis* (AB238922)	2.49	0.00	2.94	0.90	0.12	0.00	0.00	0.00	0.00	0.34	0.00	0.00	0.00	0.35	0.47	0.13	1.18	0.00	0.21	0.22
OPU-110 (*Bacteroidia, Bacteroidetes*) *Parabacteroides johnsonii* (AB261128); *P. merdae* (AB238928)	0.15	0.00	0.00	0.00	0.22	0.00	0.00	0.00	0.00	0.45	0.00	3.02	1.00	0.98	4.85	0.13	0.00	0.38	0.54	0.20
OPU-111 (*Bacteroidia, Bacteroidetes*) *Barnesiella intestinihominis* (AB370251)	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.35	3.23	0.13	0.08	0.06	0.00	0.10	0.70	0.02
OPU-116 (*Bacteroidia, Bacteroidetes*) *Alistipes finegoldii* (AY643083); *A. onderdonkii* (AY974071); A. *putredinis* (L16497); *A. shahii* (AY974072)	0.01	0.00	0.00	0.00	0.00	1.78	0.16	0.00	0.00	0.00	0.00	0.11	2.46	0.48	1.58	0.07	0.03	0.10	0.11	0.56
OPU-172 (*Negativicutes, Firmicutes*) Uncultured bacterium (AJ583204)	0.05	0.62	0.21	0.06	0.12	5.10	6.45	7.36	5.14	3.35	0.00	0.01	0.03	0.37	0.53	0.24	0.27	0.07	0.02	0.02
OPU-207 (*Clostridia, Firmicutes*) *Dorea longicatena* (AJ132842)	0.01	0.00	0.00	0.60	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.24	0.65	0.98	0.89	0.35	1.20	0.00	0.44
OPU-209 (*Clostridia, Firmicutes*) *Dorea formicigenerans* (L34619)	0.00	0.00	0.00	0.45	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.53	0.48	0.40	1.66	1.34	0.45	0.00	1.23	

Table 3 (*Continued*)

OPU																				
OPU-215 (*Clostridia, Firmicutes*) *Blautia wexlerae* (EF036467)	0.96	0.18	0.96	0.54	0.36	0.00	0.00	0.83	0.00	0.00	0.00	1.14	2.75	0.54	0.63	1.06	1.07	0.38	4.06	5.09
OPU-216 (*Clostridia, Firmicutes*) Uncultured bacterium (AY982765)	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.82	0.67	0.28	1.50	1.96	1.74	0.76	0.47	1.41
OPU-220 (*Clostridia, Firmicutes*) *Blautia luti* (AJ133124)	0.00	0.00	0.00	0.09	0.07	0.00	0.00	0.00	0.00	0.00	0.00	1.82	0.04	0.39	1.08	0.21	0.05	0.21	0.13	0.99
OPU-234 (*Clostridia, Firmicutes*) Uncultured bacterium (AY982766)	0.32	0.00	0.00	0.44	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.88	0.56	0.15	0.43	0.77	0.05	0.28	0.03	2.09
OPU-236 (*Clostridia, Firmicutes*) Uncultured organism (HQ791041)	0.00	0.00	0.00	1.59	0.00	0.00	0.00	0.19	0.00	0.00	0.00	0.00	0.54	1.41	0.22	0.27	0.32	0.24	0.00	0.44
OPU-243 (*Clostridia, Firmicutes*) Uncultured bacterium (EU765024)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.54	0.22	0.07	0.24	1.28	0.03	0.00	0.30
OPU-246 (*Clostridia, Firmicutes*) *Eubacterium hadrum* (FR749932)	0.00	0.00	0.73	0.22	0.12	0.00	0.05	0.00	0.00	0.00	0.00	0.29	0.18	0.00	0.07	0.07	0.27	0.07	3.13	1.35
OPU-248 (*Clostridia, Firmicutes*) *Eubacterium hallii* (L34621)	0.00	0.00	0.00	0.28	0.02	0.00	0.16	0.00	0.00	0.00	0.00	0.64	0.12	0.00	0.12	0.73	0.13	0.03	0.07	0.93
OPU-249 (*Clostridia, Firmicutes*) *Roseburia faecis* (AY305310); *R. hominis* (AJ270482); *R. intestinalis* (AJ312385)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.04	0.04	0.61	0.22	0.62	0.30	0.16	0.14	0.00	0.97
OPU-251 (*Clostridia, Firmicutes*) Uncultured bacterium (EU462328)	0.00	0.00	0.00	0.12	0.08	0.00	0.00	0.00	0.00	0.00	0.00	3.66	2.60	0.52	0.50	0.06	0.56	2.24	0.00	0.30
OPU-277 (*Clostridia, Firmicutes*) *Clostridium lactatifermentans* (AY033434); *C. neopropionicum* (X76746); *C. propionicum* (X77841)	0.01	0.01	0.00	0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.36	1.96	0.27	0.79	1.31	0.45	0.21	0.46
OPU-290 (*Clostridia, Firmicutes*) *Faecalibacterium prausnitzii* (AJ413954)	0.12	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	19.01	24.86	0.00	6.45	8.89	15.61	19.72	20.61	23.38
OPU-296 (*Clostridia, Firmicutes*) *Clostridium leptum* (AJ305238); *C. sporosphaeroides* (X66002)	0.08	0.29	0.00	0.05	0.00	0.00	0.03	0.12	0.00	0.30	0.00	0.22	0.08	0.00	0.12	0.52	0.03	0.21	0.39	0.06
OPU-302 (*Clostridia, Firmicutes*) Uncultured bacterium (EF404944)	0.01	0.00	0.00	0.06	0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.28	0.09	0.15	0.06	0.19	0.03	0.41	0.54
OPU-306 (*Clostridia, Firmicutes*) *Oscillibacter valericigenes* (AB238598)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.07	0.78	0.75	0.17	0.19	0.24	0.13	1.37
OPU-307 (*Clostridia, Firmicutes*) Uncultured bacterium (DQ797232)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	1.20	0.47	0.04	0.67	0.03	0.00	0.50
OPU-314 (*Clostridia, Firmicutes*) *Butyricicoccus pullicaecorum* (EU410376); *Eubacterium desmolans* (L34618)	2.16	0.25	0.00	0.20	0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.75	0.67	0.15	0.53	0.40	0.03	0.03	0.58
OPU-326 (*Bacilli, Firmicutes*) *Streptococcus infantis* (AY485603)	0.00	0.58	0.88	1.10	0.46	0.00	2.76	0.37	3.68	0.02	0.00	0.09	0.07	1.39	0.93	0.00	1.63	0.45	0.11	0.02

and 20% with length >600 nuc) than other previous studies with lengths ±200–300 nuc (e.g. [9,11,31,47,54]). Therefore, there was much higher information content [56], and identification was based on phylogenetic inference (OPUs) [26,50] rather than only identity matches (OTUs) [4,11,31,47,54]. The procedure significantly reduced the diversity yields, but the results showed clearer trends than when only using clustering approaches (Fig. 1). In addition, the procedure allowed the recognition of 338 OPUs that could be assigned to either known species or genera (73%), or to sequences related to as yet uncultured organisms (27%). A similar number of phylotypes had been detected in mucosa and fecal samples of non-CD individuals [8,52,54], which altogether were one order
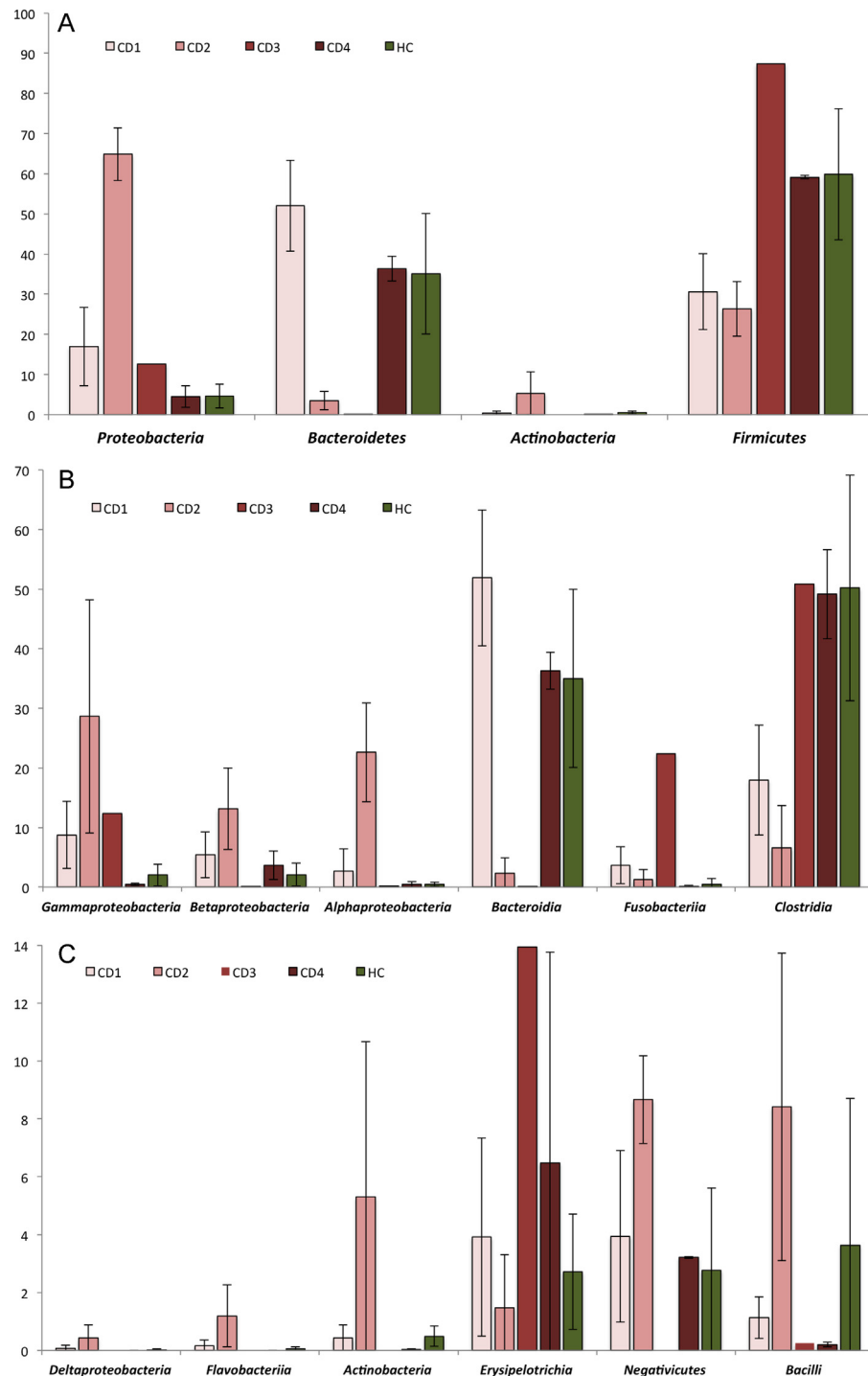
Fig. 2. Mean of the OPU occurrence in major high taxa referred to the percentage of total sequences in each sample: (A) among the four major phyla; (B) among the six classes showing OPU occurrence above 10%; (C) among the six classes exhibiting remarkable occurrence but below 10%. Bars in wine color with different intensities indicate, in increasing order, the values for the groups CD1 to CD4, whereas the values of HC individuals are green. The black bars indicate the standard deviation from the calculated mean. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of magnitude lower than other observations also in feces [46]. However, in our case, the proportions between cultivable and uncultured representatives were inverted in comparison with OTU-based diversity [8,43,46,52], and this may be due to the better quality (sequence length and phylogenetic inference) of the OPU approach used here.

Healthy subjects showed a common major microbial community structure, whereas the 13 CD patients could be divided into two main sets (CD1 and CD2 with five patients each), and three

outlier patients (CD3 and CD4) for which no statistical support could be calculated. CD4 could not be clearly distinguished from the HC patients. HC exhibited consistently higher diversity than CD either in the total composition (lower in CD2), or in the composition and abundance of *Firmicutes* (CD1). Reduced diversities, especially for *Firmicutes*, have already been reported for both fecal-[8,19,24,43], and mucosa-associated microbiota [39,53,54]. HC showed a similar occurrence of *Firmicutes* (±59%) and *Bacteroidetes* (±35%) as the major phyla, similar to that already reported for feces
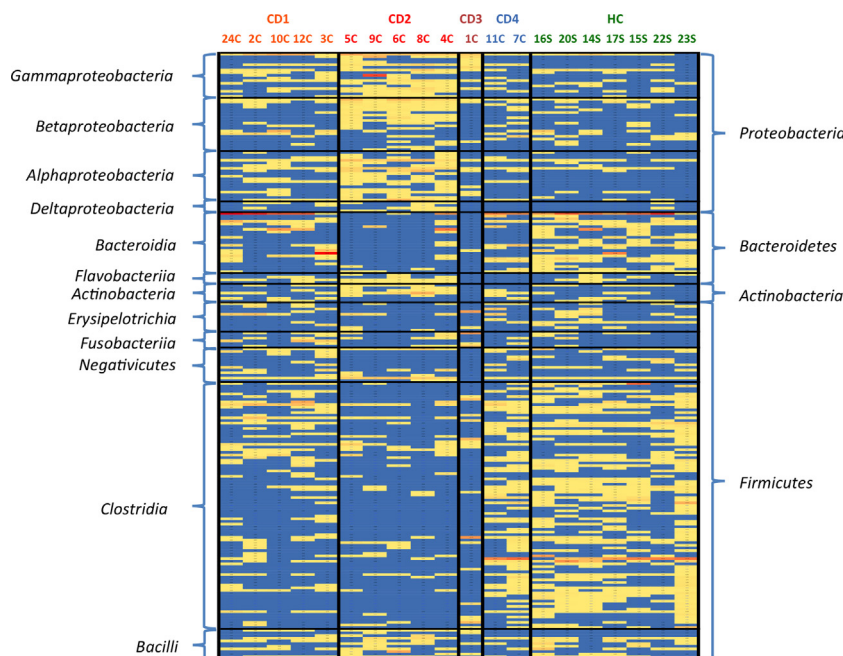
Fig. 3. Color map diagram showing the OPU distributions vertically according to their taxonomy, and horizontally according to the different group sample. The inset colors indicate: blue, zero; yellow to red increasing values above 0.1%. The figure is a reduced feature of supplementary Table S7. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

[8,19,43]. This occurrence, followed by *Proteobacteria*, was also observed in healthy colonic mucosa biopsies [39]. It was remarkable that minor abundances of *Actinobacteria* were detected, as reported by other authors [39], but *Bifidobacterium* members were not detected despite reports of their relative importance in fecal samples [8,18,37,39,48]. Shen et al. [39] recalled the need for the use of high throughput sequencing methods to reveal their occurrence in the adherent microbiota. Our results pointed to a low (or null) relevance in the subjects sampled, since no sequence was detected among the 1.5×10^5 biopsy reads, or among the fecal samples studied (data not shown). However, an amplification bias could not be discarded.

Species and genus affiliations of the major core groups were in accordance with previous observations on colonic mucosa where *F. prausnitzii* was one of the most relevant taxa [54], and *E. rectale* was nearly undetectable (only relevant in 23S and 7C), pointing to the latter being relatively unimportant in the attached microbiota. Despite their presence in almost all samples, no special relevance of other members of the core microbiome reported as abundant [8,43] could be detected, such as *C. leptum* or *B. thetaiotaomicron* (in our study OPU-296 and OPU-089). Finally, and contrary to other studies on fecal microbiota [43], *B. coccoides* (our OPU-221) was nearly absent in the HC samples. However, differences between fecal and adhered microbiomes have already been reported [54,57], and may be mainly due to a selective interaction with the host.

Two main OPUs in the HC core microbiome deserve special attention: OPU-290 (*F. prausnitzii*) and OPU-086 (*B. dorei–B. vulgatus*). *F. prausnitzii* is well known as an indicator of healthy colon microbiomes (e.g. [25,44,54]), which is a major butyrate producer in the gut that helps to maintain its health [21], and has potential therapeutic properties [40]. In all but one (16S) of our HC samples, OPU-290 was the second most relevant given the sequence abundances. In addition, the most abundant group was OPU-086, which affiliated with *B. dorei* and *B. vulgatus*, both isolated and commonly detected in human feces [2,5,13,17,46]. The clade *B. dorei–B. vulgatus* has not previously been reported to be specifically relevant in healthy individuals. However, there are some

studies reporting a reduction in its abundance in patients with IBD or cystic fibrosis [2,29]. Since a partial sequence does not harbor complete genealogic information [56], we could not decide which of the species was really present. Similarly, neither species could be differentiated by MALDI-TOF MS fingerprints [30]. Actually, the reported occurrence of *B. vulgatus* in the literature may also be inexact, since *B. dorei* has been described very recently [2], and the probes and primers designed may not differentiate them (as, for example, the specific primer for *B. vulgatus* [17]) or the identification may have been based on partial sequences [5,29,46]. Therefore, it is hypothesized that they must play an important role in the adhered microbiome in the healthy intestines of the subjects studied here, or at least they should be important members of the intestinal microbiome given their relevance also in the CD1 group.

Contrary to what was expected, CD did not form a clear homogeneous group, since two major CD datasets could be differentiated, and three patients were not very informative. CD1 and CD2, each with five subjects, showed distinct microbiome structures in comparison with HC. CD1 exhibited an increased proportion of *Bacteroidetes*, in agreement with some previous observations [45], whereas CD2 showed an enhanced presence of *Proteobacteria* [16]. Both discrepancies reinforced the hypothesis that distinct disorders, with similar endpoints, may be grouped within the CD diagnosis [36], although the sample size in the case of both groups was too low (5 patients each), and therefore insufficient for drawing any conclusion. However, looking at the clinical data compiled, CD1 and CD2 differed in some traits. For example, 80% of CD1 used immune-suppressors and were non-smokers, whereas only 20% exhibited clinical activity and a Crohn L1 and L3 colon. On the other hand, the data for CD2 were contrary, since 20% used immune-suppressors and 60% were smokers, showed clinical activity and a Crohn L1 and L3 colon. Both coincided in the absence (or very reduced abundances) of *F. prausnitzii* (e.g. 25,44,53,55,56), which was also absent in the studied feces (data not shown).

CD1 was the most similar to HC and had a comparable *Bacteroidetes* composition, but with an important reduction of *Firmicutes*, especially *Clostridia*, as previously reported [24,45]. The

most abundant OPU here also affiliated with the *B. vulgatus–B. dorei* clade. Actually, Swidsinski et al. [45] reported that *B. fragilis* could be a main feature in IBD given its enhanced presence in mucosal microbiota. However, the target of the specific probe S-*-Bfra-0602-a-A-19 differed by only one base with the sequences of *B. dorei–B. vulgatus* (G instead of A at the 5′ end [10]) making it plausible that the abundant cell morphotype detected could be the same as OPU-086. Similarly, Mangin et al. [23] hypothesized that this group may play a role in the Crohn etiology. However, the pathogenic role of these organisms seems unclear, since *B. vulgatus* has been shown to be a protective agent against *E. coli* colonization in gnotobiotic mice [51], but also plays a key role in initiating spontaneous colitis in HLA-B27 transgenic rats [34]. In the light of our findings, *B. vulgatus*-like organisms appeared to be an important (if not the most important) component of our HC and CD1 samples. In addition, CD1 also showed an increased population of *R. gnavus* (OPU-206) in abundances of approximately 5% of the total population. The increase in the presence of this Gram-positive species had been regarded as an indicator of CD together with the reduction of *F. prausnitzii* [15,54].

For CD2, the important increase in sequences of the major proteobacterial classes (*Gamma-*, *Alpha-*, *Beta-* and *Deltaproteobacteria* in decreasing relevance) was the most conspicuous. Members of this group (especially *Enterobacteriaceae*) had been reported to be mucotropic, mainly colonizing the border between mucus and feces, and with increased amounts in patients with an important reduction of the mucus thickness of the gut epithelium [45]. This increase may be common among IBD as it has also been reported for colorectal adenomas [39]. Most relevant were OPU-001 (*Escherichia–Shigella* spp., *Gammaproteobacteria*), OPU-028 (*Cupriavidus–Ralstonia* spp., *Betaproteobacteria*) and OPU-067 (*Mesorhizobium plurifarium*, *Alphaproteobacteria*). However, these three OPUs were not the only ones relevantly enhanced, since there was a complex picture of approximately 30 distinct OPUs belonging to the four major classes of *Proteobacteria*. Thus, due to a balanced presence and the as yet unreported pathogenicity of most of them (apart from OPU-001), it seems plausible to consider opportunistic colonization due to the depletion of *Clostridia* and/or a reduction of the thickness in the mucosa layer [3,45]. The increase of OPU-001 (*Escherichia–Shigella* spp.) has been commonly reported in relation to CD [3,16,25,45,53,54]. The presence of members of this group (generally identified as *E. coli*) correlated in abundance with the severity of inflammation, as well as with the isolation of novel groups of invasive *E. coli* [3].

In summary, the combined use of high quality 454 sequencing and the OPU approach led to an accurate identification of most of the bacterial individuals at the species or genus level. Altogether, the results coincided in that different (at least two in our case) microbiological disorders with similar inflammatory processes may be considered under CD [36,45]. The two major groups observed shared a depletion of the Gram-positive fraction (especially *Clostridia* and, remarkably, *F. prausnitzii*) that either enhances the relevance of the co-colonizing *Bacteroidetes* (CD1), or allows the colonization of an opportunistic community of members of *Proteobacteria* (CD2). From the clinical information of the patients, it was not possible to identify the reasons that led to the disorders, but perhaps genetic or metabolic studies on the patients could reveal the factors involved in the differences of the microbiota composition. One of the most remarkable observations was the relevance of the clade *B. vulgatus–B. dorei*, which, despite the hypothesis that it has no protective role, should have been at least a common and important saprophytic part of the gut epithelial microbiome of the subjects studied here. Simple amplification tests checking the presence or absence of *F. prausnitzii*, OPU-086 and certain *Proteobacteria* (such as OPU-001) could serve for the diagnosis of both major CD groups found here, and may help in the recognition of the genetic or metabolic disorders that lead to the Crohn phenotypes.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.syapm.2015.06.008

References

[1] Bäckhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., Gordon, J.I. (2005) Host-bacterial mutualism in the human intestine. Science 307, 1915–1920.

[2] Bakir, M.A., Sakamoto, M., Kitahara, M., Matsumoto, M., Benno, Y. (2006) *Bacteroides dorei*, sp. nov., isolated from human faeces. Int. J. Syst. Evol. Microbiol. 56, 1639–1643.

[3] Baumgart, M., Dogan, B., Rishniw, M., Weitzman, G., Bosworth, B., Yantiss, R., Orsi, R.H., Wiedmann, M., McDonough, P., Kim, S.G., Berg, D., Scukken, Y., Scherl, E., Simpson, K.W. (2007) Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of *Clostridiales* in Crohn's disease involving the ileum. ISME J. 1, 403–418.

[4] Birtel, J., Walser, J.-C., Pichon, S., Bürgmann, H., Matthews, B. (2015) Estimating bacterial diversity for ecological studies: methods, metrics and assumptions. PLOS ONE 10, e0125356.

[5] Bruzzese, E., Callegari, M.L., Raia, V., Viscovo, S., Scotto, R., Ferrari, S., Morelli, L., Buccigrossi, V., Lo Vecchio, A., Ruberto, E., Guarino, A. (2014) Disrupted intestinal microbiota and intestinal inflammation in children with cystic fibrosis and its restoration with *Lactobacillus* GG: a randomised clinical trial. PLOS ONE 9, e87796.

[6] Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Buschman, F.D., Costello, E.K., Fierer, N., González Peña, A., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R. (2010) QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.

[7] Collins, M.D., Lawson, P.A., Willems, A., Cordoba, J.J., Fernández-Garayzabal, J., García, P., Cai, J., Hippe, H., Farrow, J.A. (1994) The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. Int. J. Syst. Bacteriol. 44, 812–826.

[8] Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., Relman, D.A. (2005) Diversity of the human intestinal microbial flora. Science 308, 1635–1638.

[9] França, L., López-López, A., Rosselló-Móra, R., Da Costa, M. (2014) Microbial diversity and dynamics of a groundwater and still bottled natural mineral water. Environ. Microbiol. 17, 577–593.

[10] Franks, A.H., Harmsen, H.J.M., Raangs, G.C., Jansen, G.J., Schut, F., Welling, G.W. (1998) Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Appl. Environ. Microbiol. 64, 3336–3345.

[11] Gobet, A., Boetius, A., Ramette, A. (2014) Ecological coherence of diversity patterns derived from classical fingerprinting and next generation sequencing techniques. Environ. Microbiol. 16, 2672–2681.

[12] Guarner, F., Malagelada, J.R. (2003) Gut flora in health and disease. Lancet 306, 512–519.

[13] Gustafsson, R.J., Ohlsson, B., Benoni, C., Jeppsson, B., Olsson, C. (2012) Mucosa-associated bacteria in two middle-aged women diagnosed with collagenous colitis. World J. Gastroenterol. 18, 1628–1634.

[14] Hammer, Ø., Ryan, P., Harper, D. (2001) PAST: paleontological statistics software package for education and data analysis. Palaeontol. Electron. 4, 4–9, http://palaeo-electronica.org/2001_1/past/issue1_01.htm

[15] Joossens, M., Huys, G., Cnockaert, M., De Preter, V., Verbeke, K., Rutgeerts, P., Vandamme, P., Vermeire, S. (2011) Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut 60, 631–637.

[16] Kleessen, B., Kroesen, A.J., Buhr, H.J., Blaut, M. (2002) Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. Scand. J. Gastroenterol. 37, 1034–1041.

[17] Kreader, C.A. (1995) Design and evaluation of *Bacteroides* DNA probes for the specific detection of human fecal pollution. Appl. Environ. Microbiol. 61, 1171–1179.

[18] Lay, C., Rigottier-Gois, L., Holmstrøm, K., Rajilic, M., Vaughan, E.E., de Vos, W.M., Collins, M.D., Thiel, R., Namsolleck, P., Blaut, M., Doré, J. (2005) Colonic microbiota signatures across five northern European countries. Appl. Environ. Microbiol. 71, 4153–4155.

[19] Ley, R.E., Turnbaugh, P.J., Klein, S., Gordon, J.I. (2006) Human gut microbes associated with obesity. Nature 444, 1022–1023.

[20] Liu, L., Li, Y., Li, S., Hu, H., He, Y., Pong, R., Lin, D., Lu, L., Law, M. (2012) Comparison of next-generation sequencing systems. J. Biomed. Biotechnol. 11, http://dx.doi.org/10.1155/2012/251364, Article ID 251364.

[21] Louis, P., Flint, H. (2009) Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. FEMS Microbiol. Lett. 294, 1–8.

[22] Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A.W., Gross, O., Grumann, S., Hermann, S., Jost, R., König, A., Liss, T., Lussmann, T., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., Schleifer, K.H. (2004) ARB: a software environment for sequence data. Nucleic Acids Res. 32, 1363–1371.

[23] Mangin, I., Bonnet, R., Seksik, P., Rigottier-Gois, L., Sutren, M., Bouhnik, Y., Neut, C., Collins, M.D., Colombel, J.-F., Marteau, P., Doré, J. (2004) Molecular inventory of faecal microflora in patients with Crohn's disease. FEMS Microbiol. Ecol. 50, 25–36.

[24] Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R., Jarrin, C., Chardon, P., Marteau, P., Roca, J., Dore, J. (2006) Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut 55, 205–211.

[25] Martínez-Medina, M., Aldeguer, X., González-Huix, F., Acero, D., Garcia-Gil, J. (2006) 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. 12, 1136–1145.

[26] Mora-Ruiz, M.R., Font-Verdera, F., Díaz-Gil, C., Urdiain, M., Rodríguez-Valdecantos, G., González, B., Orfila, A., Rosselló-Móra, R. (2015) Moderate halophilic bacteria colonizing the phylloplane of halophytes of the subfamily *Salicornioideae* (*Amaranthaceae*). Syst. Appl. Microbiol, in press.

[27] Muyzer, G., Brinkhoff, T.T., Nübel, U., Santegoeds, C., Schäfer, H., Waver, C. (2004) Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. In: Kowalchuk, G.A., de Bruijn, F.J., Head, I.M., Akkermans, A.D.L., van Elsas, J.D. (Eds.), Molecular Microbial Ecology Manual, 2nd ed., Kluwer Academic Publishers, Dordecht, The Netherlands, pp. 743–770.

[28] Muyzer, G., Teske, A., Wirsen, C.O., Jannasch, H.W. (1995) Phylogenetic relationship of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient electrophoresis of 16S rDNA fragments. Arch. Microbiol. 164, 165–172.

[29] Noor, S.O., Ridgway, K., Scovell, L., Kemsley, E.K., Lund, E.K., Jamieson, C., Johnson, T., Narbad, A. (2010) Ulcerative colitis and irritable bowel patients exhibit distinct abnormalities of the gut microbiota. BMC Gastroenterol. 10, 134.

[30] Pedersen, R., Marmolin, E.S., Justesen, U.S. (2013) Species differentiation of *Bacteroides dorei* from *Bacteroides vulgatus* and *Bacteroides ovatus* from *Bacteroides xylanisolvens* – back to basics. Anaerobe 24, 1–3.

[31] Pilloni, G., Granitsiotis, M.S., Engel, M., Lueders, T. (2012) Testing the limits of 454 pyrotag sequencing: reproducibility, quantitative assessment and comparison to T-RFLP fingerprinting of aquifer microbes. PLoS ONE 7, e40467.

[32] Pruesse, E., Peplies, J., Glockner, F.O. (2012) SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28, 1823–1829.

[33] Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596.

[34] Rath, H.C., Wilson, K.H., Sartor, R.B. (1999) Differential induction of colitis and gastritis in HLA-B27 transgenic rats selectively colonized with *Bacteroides vulgatus* or *Escherichia coli*. Infect. Immun. 67, 2969–2974.

[35] Rosselló-Móra, R., López-López, A. (2008) The least common denominator: species or operational taxonomic units? In: Zengler, K. (Ed.), Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back, ASM Press, Washington, DC, pp. 117–130.

[36] Sartor, R.B. (2006) Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis Nat. Clin. Pract. Gastroenterol. Hepatol. 3, 390–407.

[37] Scanlan, P.D., Shanahan, F., O'Mahony, C., Marchesi, J.R. (2006) Culture-independent analyses of temporal variation of the dominant fecal microbiota

[38] and targeted bacterial subgroups in Crohn's disease. J. Clin. Microbiol. 44, 390–3988.

[38] Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.

[39] Shen, X.J., Rawls, J.F., Randall, T., Burcal, L., Mpande, C.N., Jenkins, N., Jovov, B., Abdo, Z., Sandler, R.S., Keku, T.O. (2010) Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. Gut Microbes 1, 138–147.

[40] Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L.G., Gratadoux, J.-J., Blugeon, S., Bridonneau, C., Furet, J.-P., Corthier, G., Grangette, C., Vasquez, N., Pochard, P., Trugnan, G., Thomas, G., Blottière, H.M., Doré, J., Marteau, P., Seksik, P., Langella, P. (2008) *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc. Natl. Acad. Sci. U.S.A. 105, 16731–16736.

[41] Stackebrandt, E., Ebers, J. (2006) Taxonomic parameters revisited: tarnished gold standards. Microbiol. Today 33, 152–155.

[42] Stackebrandt, E., Goebel, B.M. (1994) Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int. J. Syst. Bacteriol. 44, 846–849.

[43] Suau, A., Bonnet, R., Sutren, M., Godon, J.-J., Gibson, G.R., Collins, M.D., Doré, J. (1999) Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol. 65, 4799–4807.

[44] Swidsinski, A., Loening-Baucke, V., Vaneechoutte, M., Doerffel, Y. (2008) Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. Inflamm. Bowel Dis. 14, 147–161.

[45] Swidsinski, A., Weber, J., Loening-Bauke, V., Hale, L.P., Lochs, H. (2005) Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. J. Clin. Microbiol. 43, 3380–3389.

[46] Tap, J., Mondot, S., Levenez, F., Pelletier, E., Caron, C., Furet, J.-P., Ugarte, E., Muñoz-Tamayo, R., Paslier, D.L.E., Nalin, R., Dore, J., Leclerc, M. (2009) Towards the human intestinal microbiota phylogenetic core. Environ. Microbiol. 11, 2574–2584.

[47] Turnbaugh, P.J., Ridaura, V.K., Faith, J.J., Rey, F.E., Knight, R., Gordon, J.I. (2009) The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci. Transl. Med. 1, 6ra14.

[48] Turroni, F., Marchesi, J.R., Foroni, E., Gueimonde, M., Shanahan, F., Margolles, A., van Sinderen, D., Ventura, M. (2009) Microbiomic analysis of the bifidobacterial population in the human distal gut. ISME J. 3, 745–751.

[49] Verbarg, S., Rheims, H., Emus, S., Frühling, A., Kroppenstedt, R.M., Stackebrandt, E., Schumann, P. (2004) *Erysipelothrix inopinata* sp. nov., isolated in the course of sterile filtration of vegetable peptone broth, and description of the *Erysipelotrichaceae* fam. nov. Int. J. Syst. Evol. Microbiol. 54, 221–225.

[50] Viver, T., Cifuentes, A., Díaz, S., Rodríguez-Valdecantos, G., González, B., Antón, J., Rosselló-Móra, R. (2015) Diversity of extremely halophilic cultivable prokaryotes in Mediterranean, Atlantic and Pacific solar salterns: evidence that unexplored sites constitute sources of cultivable novelty. Syst. Appl. Microbiol. 38, 266–275.

[51] Waidmann, M., Bechtold, O., Frick, J.-S., Lehr, H.-A., Schubert, S., Dobrindt, U., Löffler, J., Bohn, E., Autenrieth, I.B. (2003) *Bacteroides vulgatus* protects against *Escherichia coli*-induced colitis in gnotobiotic interleuking-2-deficient mice. Gastroenterology 125, 162–177.

[52] Walker, A.W., Ince, J., Duncan, S.H., Webster, L.M., Holtrop, G., Ze, X., Brown, D., Stares, M.D., Scott, P., Bergerat, A., Louis, P., McIntosh, F., Johnstone, A.M., Lobley, G.E., Parkhill, J., Flint, H.J. (2011) Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J. 5, 220–230.

[53] Willing, B., Halfvarson, J., Dicksved, J., Rosenquist, M., Järnerot, G., Engstrand, L., Tysk, C., Jansson, J. (2009) Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. Inflamm. Bowel Dis. 15, 653–660.

[54] Willing, B.P., Dicksved, J., Halfvarson, J., Andersson, A.F., Lucio, M., Zheng, Z., Järnerot, G., Tysk, C., Jansson, J.K., Engstrand, L. (2010) A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary inflammatory bowel disease phenotypes. Gastroenterology 139, 1844–1854.

[55] Yarza, P., Ludwig, W., Euzéby, J., Amann, R., Schleifer, K., Glöckner, F.O., Rosselló-Móra, R. (2010) Update of the all-species living tree project based on 16S and 23S rRNA sequence analyses. Syst. Appl. Microbiol. 33, 291–299.

[56] Yarza, P., Yilmaz, W., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzéby, J., Amann, R., Rosselló-Móra, R. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12, 635–645.

[57] Zoetendal, E.G., Wright, A., Vilpponen-Salmela, T., Ben-Amor, K., Akkermans, A.D.L., de Vos, W. (2002) Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl. Environ. Microbiol. 68, 3401–3407.