Reduced Diversity of Fecal Microbiota in Crohn's Disease revealed by a Metagenomic Approach

C Manichanh[1], L Rigottier-Gois[1], E Bonnaud[1], K Gloux[1], E Pelletier[2], L Frangeul[3], R Nalin[4], C Jarrin[4], P Chardon[5], P Marteau[6], J Roca[7] and J Dore[1]

[1]Unité d'Ecologie et de Physiologie du système Digestif, INRA Jouy-en-Josas, France. [2]Genoscope, Centre National de Séquençage, CP5706 91057, Evry cedex, France, [3]Genopole, Institut Pasteur, 28 rue du Dr Roux, 75724 Paris 15, Paris, France. [4]LibraGen, Bat. Canal Biotech I, 3 rue des Satellites, 31400 Toulouse, France. [5]UMR INRA CEA radiobiologie et étude du génome, INRA Domaine de Vilvert, 78352 Jouy-en-Josas, cedex France. [6]Gastroenterology Department, European Hospital Georges Pompidou, 20 rue Leblanc 75908 Paris cedex 15 France. [7]Department of Molecular Biology, IBMB-CSIC, Jordi Girona 18-26, Barcelona 08034, Spain

keywords : Crohn's disease, Metagenome, fecal microbiota, phylogeny, Firmicutes

Corresponding author:
Dr. Joel Dore, INRA - UEPSD, 78350 Jouy-en-Josas, France.
E-mail: joel.dore@jouy.inra.fr.

Abbreviations: CD, Crohn's disease; CDAI; IBD, inflammatory bowel disease; FISH, fluorescent in situ hybridization; OTU, operational taxonomic unit; HSL, Healthy subject library; CPL, Crohn patient library.

Abstract

Background and Aim: A role of the intestinal microbial community (microbiota) in the onset and chronicity of Crohn's disease (CD) is strongly suspected. However, investigation of such a complex ecosystem is difficult, even with culture-independent molecular approaches.

Methods: We used, for the first time, a comprehensive metagenomic approach to investigate the full range of intestinal microbial diversity. We used a fosmid vector to construct two libraries of genomic DNA isolated directly from fecal samples of six healthy donors and six patients with Crohn's disease. Bacterial diversity was analyzed by screening the two DNA libraries, each composed of 25,000 clones, for the 16S rRNA gene by DNA hybridization.

Results: Among 1190 selected clones, we identified 125 non-redundant ribotypes mainly represented by the phyla Bacteroidetes and Firmicutes. Among the Firmicutes, 43 distinct ribotypes were identified in the healthy microbiota, compared to only 13 in CD ($p<0.025$). Fluorescence *in situ* hybridization directly targeting 16S rRNA in fecal samples analysed individually ($n=12$) confirmed the significant reduction in the proportion of bacteria belonging to this phylum in CD patients ($p<0.02$).

Conclusion: Metagenomic approach allowed us to detect a reduced complexity of the bacterial phylum Firmicutes as a signature of the fecal microbiota in patients with Crohn's disease. It also indicated the presence of new bacterial species.

Introduction

Under certain circumstances, the indigenous gut microbes will be determinant in the onset and maintenance of inflammatory bowel disease, especially Crohn's disease (CD), to which genetic susceptibility and disorders in mucosal immunity could also be associated.[1] Mutations in the CARD15/NOD2 gene have been associated to a higher risk of CD.[2] Similarly, dysbiosis has recently been coined as a trait linked to CD.[3]

Since most intestinal commensals cannot be cultured,[4][5] genomic strategies have therefore been developed to overcome this limitation. Using classical techniques, such as temporal temperature gradient gel electrophoresis (TTGE) [6] and single strand conformation polymorphism (SSCP),[7] a few studies have shown an alteration in the diversity of the microbiota in Crohn's disease, but the precise microbial species or metabolites involved remain unknown.

Metagenomic approach, in which the microbial community is studied as a single dynamic entity, has been used to investigate complex environments such as water and soil,[8][9] but has not yet been applied to the human intestinal microbiota. In metagenomics, microbial genomes are fragmented and cloned in order to obtain a single library representing the microbiome (genomes of a complex community). Whereas former researches focussed on comparisons of PCR derived 16S rDNA libraries, metagenomic libraries similarly permit comparison of species diversity based on a PCR independent approach but also a complete description of genes based on genomic sequence analysis and thereby a full description of functionalities of the whole ecosystem.

Here we present the first metagenomic approach to investigate the fecal bacterial composition in CD. We isolated genomic DNA directly from fecal samples of six healthy donors and six CD patients, and used the pooled DNA to construct one library for each group of subjects. The libraries were then screened for 16S rRNA genes by DNA hybridization, and the microbial diversity was determined by 16S rRNA gene sequencing and phylogenetic analysis. The metagenomic results were confirmed by using a fluorescence *in situ* hybridization method applied to individual fecal samples from the same subjects.

Materials and Methods

Methodology

We extracted genomic DNA from fecal samples of six healthy individuals and six patients with Crohn's disease. We used the pooled DNA to construct two metagenomic libraries for each group of subjects. Each library contained 25.000 clones. Two macroarrays were built after fosmid extraction and DNA spotting on nylon membrane for each library. In order to analyse microbial diversity, 16S rRNA genes were screened by DNA hybridization and sequenced, and phylogenetic analysis was performed (fig 1).

Patients and samples.

None of the patients and healthy subjects had received antibiotics for at least 3 months before sampling. The 6 patients with CD were 5 women and 1 man, aged 18-44 years. The patients were in remission (with a Crohn's disease activity index < 150). All of them had ileal disease; four patients had associated lesions in the duodenum, caecum, left colon and anus (one case for each associated lesion). Two patients received no treatment at the time of sampling; two received 5-ASA (2 or 3 g/d) and two received azathioprine (150 mg/d). The 6 healthy volunteers were 1 woman and 5 men aged 29-43 years.

Fecal samples were collected from the twelve subjects, and were processed in less than three hours in order to preserve strictly anaerobic bacteria from oxygen. Bacterial cells were recovered by Nycodenz density

gradient centrifugation,[10] then washed twice with PBS (10 min at 8000 *g*). The cell pellets were then incorporated in agarose before gentle bacterial lysis[10] and DNA digestion by the enzyme *Sau 3AI*.[11]

Library construction.

High-molecular-weight bacterial DNA fragments (38 to 48 kbp) were isolated as previously described.[10] The metagenomic DNA was then cloned into fosmid, a bacterially-propagated phagemid vector system suitable for cloning genomic inserts approximately 40 kilobases in size, by using the EpiFos library production kit (Epicentre Technologies) as recommended by the manufacturer (fig 1).

Macroarray preparation.

Screening for 16S rRNA genes in the fosmid libraries involved extraction of the recombinant fosmids in order to remove the 16S rRNA gene of *E. coli* used in the cloning system. Extraction was performed with a NucleoSpin® 96 Flash kit as recommended by the manufacturer (Macherey-Nagel, Hoerdt, France). Between 2 and 4 ng of each extracted fosmid was spotted on nylon membranes measuring 22 x 22 cm (Amersham Biosciences, Orsay, France) in a 5 x 5 format, using the Qbot robot (Genetix, Saint-Marcel, France); about 50,000 spots can be deposited on each membrane.

Macroarray probe preparation.

In order to screen for 16S rRNA genes using the most universal probe, we used whole DNA directly extracted from fecal samples [12] as a template to generate a mixed probe by PCR amplification. The primers we used for amplification (ACM 008 F and EUB 338, table 1) generated PCR products of about 330 bp in length, representative of the diversity of the original ecosystem. The fecal samples were collected from the six healthy individuals used to construct the metagenomic library, and from two other healthy individuals. PCR was run with a Peltier PTC-100TM Thermal Cycler (MJ ResearchTM, Inc., Massasuchetts, USA) at 94°C for 10 min, 9 cycles: 92°C for 1 min; 58°C for 1 min; 72°C for 1 min; and 72°C for 10 min. The forward and reverse primers were ACM 008 F and EUB 338, respectively (table 1). The concentrations of PCR products were evaluated on the basis of band intensities after gel electrophoresis. PCR products (30 ng) were labeled with [α-^{33}P]dATP by using the Megaprime™ DNA labeling System (Amersham Biosciences, Uppsala, Sweden).

Table 1. Primers and probes used in this study.

Primer or Probe	5' to 3' sequence	Experiments
ACM 008 F*	AGAGTTTGATCCTGGCTCAG	Sequencing Macroarray-hybridization
ADM 330 F*	ACTCCTACGGGAGGCAGC	sequencing
SSM 1100 R*	GGGCCTTGTACACACCG	sequencing
TTM 1517 R*	GYAACGAGCGCAACCC	sequencing
NON 338**	ACATCCTACGGGAGGC	FISH
EUB 338**	GCTGCCTCCCGTAGGAGT	FISH Macroarray-hybridization
Clep 866**	GGTGGATWACTTATTGTG	FISH
Clep 866 competitor 1**	GGTGGAWACTTATTGTG	FISH
Clep 866 competitor 2**	GGTGGATWACTTATTGCG	FISH
Bac 303**	CCAATGTGGGGACCTT	FISH
Erec 482**	GCTTCTTAGTCARGTACCG	FISH

Probes Clep 866, Erec482 and Bac 303 were developed to detect members of the *Clostridium leptum*, *Clostridium coccoides* and *Bacteroides* groups, respectively. ACM 008 F and EUB 338 were used as primers to develop a probe for 16S rDNA library hybridization after PCR amplification, using DNA extracted from fecal samples and [α-^{33}P]dATP labeling. *: personnal

communication of D. Le Paslier, **: Lay et al, 2005. F: Forward, R: Reverse.

Macroarray hybridization

The spotted membranes were first incubated in pre-hybridization solution (6x SSC, 1% SDS, 10x Denhardt, 50% formamide) at 42°C for 2 hours. Then a ^{33}P-labelled probe (10^6 cpm/ml) and sonicated salmon sperm DNA (100 µg/ml, denatured for 10 min at 100°C) were added to the pre-hybridization solution and the incubation was continued at 42°C overnight. After hybridization the membranes were washed once in 2x SSC, 0.1% SDS at 42°C for 10 min, once in 2x SSC, 0.1% SDS at 60°C for 10 min, once in 0.2x SSC, 0.1% SDS at 60°C for 10 min, and once in 0.1x SSC, 0.1% SDS at 65°C for 10 min. Phosphor screens were exposed to the membranes at room temperature for 5 h and then scanned with the Molecular Dynamics Storm system and analyzed with Xdigitise software (http://www.molgen.mpg.de/~xdigitise/).

Sequences.

All sequences were determined by Genoscope (Evry, France) on an ABI 3730 DNA sequencer using four primers (ACM 008 F, ADM 330 F, SSM 1100 R and TTM 1517 R) specific for bacteria (table 1). The 16S rRNA gene was sequenced in four fragments which were then quality-controlled using the PHRED v0.020425.c program and assembled with the PHRAP v0.990319 program. Assembled sequences covering approximately positions 200 to 1300 were selected for subsequent analysis. For each library, sequences were aligned using the ClustalW v1.83 program.[13] Highly variable regions of the multiple alignment were removed, and conserved regions were selected with Gblocks,[14] using parameters optimized for rRNA alignments.[15] The use of this computerized method avoids the manual refinement of multiple alignments. Distance matrices were computed with DNADIST v3.6[16] and trees were constructed with NEIGHBOR v3.6.[17] Unrooted trees were drawn with Treeview.[18] We defined an operational taxonomic unit (OTU) as a cluster of 16S rDNA sequences sharing at least 98% of similarity.[19] A single representative fecal clone was selected for each OTU or ribotype. This clone was used as a reference sequence for calculating phylogenetic distances from other aligned sequences. Chimeric sequences were checked using the RDP CHECK_CHIMERA program (RDP-II release 8.1). The stability of branches was assessed by the bootstrap method with 1000 replicates. The term "uncultured OTUs" refers to species recovered by molecular methods only. Unidentified species (or totally novel species) correspond to microorganisms that match no available sequences in public databases (NCBI and RDP-II release 9).[20] We used phylum- and group-level nomenclatures as defined in Berguey's Manual of Systematic Bacteriology (second edition) and the RDP Naive Bayesian classifier, respectively. The representative nature of our libraries was estimated by using Good's coverage formula.[21]

Good's coverage was calculated as $[1-(n/N)] \times 100$, where n is the number of single-clone OTU and N is the total number of sequences for the analyzed sample.

Fluorescence *In Situ* Hybridization (FISH).

FISH combined with flow cytometry was used to analyze the composition of fecal samples, targeting the 16S rRNA of three major groups of bacteria (Lay *et al.*, 2005). The 16S rRNA-specific oligonucleotide probes used in this study are listed in table 1. Probes EUB 338 and NON 338 were used as positive and negative controls, respectively. Probes Bac 303, Erec 482 and Clep 866 were used to detect the *Bacteroides*, *Clostridium coccoides* and *Clostridium leptum* groups, respectively. The specificity of probe Clep 866 was optimized by using unlabelled mismatched oligonucleotides as competitors (table 1). Fecal samples stored at -70°C were fixed in

paraformaldehyde and hybridized with specific probes, and positive signals were detected by flow cytometry as previously described (Lay *et al*., 2005).

Statistical

The data were analyzed with the χ^2 test or Student's t test for two independent sets of samples. P values <0.05 were considered to denote significant differences.

Results
Metagenomic libraries

Bacterial fractions from the six healthy individuals and the six CD patients were pooled to build a healthy-subject DNA library (HSL) and a Crohn's-patient DNA library (CPL). Each metagenomic DNA library contained 25,000 fosmid clones. The cumulative size of the inserts spanned 2 Gbp, corresponding to about 500 times the size of the *E. coli* genome (4.1×10^6 bp). The very low redundancy of the cloned fragments, as determined by randomly searching for cloned inserts having similar end sequences (data not shown), confirmed that major biases can be avoided by using this method. The DNA macroarrays were constructed after fosmid DNA extraction from the 50,000 recombinant clones and spotting on two nylon membranes, one for HSL and one for CPL (fig 1).

16S rRNA screening and sequencing, and bioinformatics analysis

Among the 50,000 clones, 1520 were positive for 16S rRNA sequences (650 from HSL and 870 from CPL). Each clone was sequenced with a set of four universal bacterial primers spanning the 1500 bp of the 16S rRNA gene sequence. After quality control and assembly, 1190 clones yielded sequences of more than 1000 informative base pairs suitable for phylogenetic analysis (536 for HSL and 654 for CPL).

Composition of the healthy and CD microbiota

Analysis of the 1190 clones identified 125 non-redundant operational taxonomic units (OTUs or ribotypes, GenBank accession numbers AY850400 to AY850541). According to public nucleotide databases (NCBI and RDP-II), 95 OTUs corresponded to species that are not contained in current culture collections (fig 2, fig 3). All the molecular species thus recovered fell into four phylogenetic phyla: Bacteroidetes (53 OTUs ; Gram negative bacteria), Firmicutes (54 OTUs ; low GC Gram positive bacteria), Actinobacteria (9 OTUs ; high GC Gram positive bacteria) and Proteobacteria (9 OTUs ; Gram negative bacteria). Sixty-six OTUs appeared completely novel, being unrelated to previously cultured microorganisms or to cloned rRNA sequences. The 89% coverage provided by the 1190 clones indicated that any new clone sequenced only had an 11% chance of corresponding to an unknown species. Hence, our metagenomic libraries covered nearly 90% of the dominant species that could be cloned from the fecal ecosystem.

Comparison of the CD and healthy libraries

The coverage values of the two libraries were not significantly different (90% and 87% for HSL and CPL, respectively). The same four dominant phyla were found in the two groups of subjects (fig 4). Bacteroidetes and Firmicutes accounted for the largest number of ribotypes, as shown by the 16S rRNA phylogenetic trees (fig 2, fig 3). The total number of Bacteroidetes OTUs was similar in the two groups, while differences in the *Prevotella* and *Bacteroides fragilis* subgroups were found. In addition, one "unclassified Porphyromonadaceae" species was observed only in the CD library and represented as much as 8.7% of the clones. The most striking

difference was a global loss of microbial diversity in CD (88 ribotypes in healthy subjects, 54 in CD patients). This was essentially due to far fewer Firmicute ribotypes in the CD patients (fig 3). Indeed, 43 Firmicute OTUs were found in the healthy subjects, compared to only 13 in the CD patients ($p<0.005$) (fig 4). The number of ribotypes shared by the two libraries was also significantly lower in this phylum (2/13, 15%) than in the Bacteroidetes group (13/33, 39%).

Confirmation by FISH results

The next question to be addressed was whether this reduction of diversity in the Firmicutes could be due to only one patient or it would be found in every patient of the study. To address this question, we applied a FISH method to each of the 12 fecal samples, individually. FISH combined with flow cytometry was used to count viable bacteria with ribosomal RNA transcription activity. Specific probes against the three major bacterial groups (table 1) revealed a significant relative reduction in the *Clostridium leptum* group in the CD patients compared to the healthy subjects ($P<0.02$), whereas the proportions of the *Clostridium coccoides* and *Bacteroides* groups were similar in the two groups of subjects (fig 5). The results obtained by FISH and the metagenomic approach were highly consistent, showing a marked reduction in the number of active cells and in the diversity of the *Clostridium leptum* group in patients with Crohn's disease.

Discussion

This study suggests that the fecal microbiota of patients with Crohn's disease contains a markedly reduced diversity of Firmicutes. In particular, the *Clostridium leptum* phylogenetic group was significantly less abundant in CD patients than in healthy subjects. This observation among patients in remission suggests that it could correspond to a primary modification, i.e. a modification of the fecal microbiota that may exist before the onset of the disease and/or in absence of major perturbations caused by the inflammation process. We focussed on patients in quiescence because during the active phase of the disease, the microbiota is remodelled [6] and alterations in its composition could be a consequence more than a cause of inflammation.

A previous analysis of amplified and cloned rRNA gene libraries from healthy volunteers (n=3) and CD patients (n=3)[22] also indicated that the Firmicutes (*Clostridium leptum* and *Clostridium coccoides* groups) had significantly lesser complexity in CD (25 versus 73 OTUs) (personnal commnunication of Irène Mangin). These results were obtained whether the subjects were considered individually or pooled. As in the present study, the number of ribotypes in the phylum Bacteroidetes was similar in the two libraries.

Although the cause of inflammatory bowel disease remains unknown, the indigenous intestinal microbiota is considered as a major if not the main trigger of inflammation, both in animal models and in humans.[23] A genetically determined abnormal response to indigenous bacteria is suspected as well as a dysbiosis.[3][23] However the precise microbial species or metabolites involved remain unknown. We share the current concept that the onset of Crohn’s disease could be due to an altered microbiota and that a dysbiosis could enhance the risk of disease. In that respect, the reduction in proportion of one microbial group is expected to be compensated by a greater representation of others. In the present study, the less diverse and less represented microbial species in CD patients are Gram positive anaerobic bacteria that usually account for a major fraction of the fecal microbiota of healthy subjects. These species may provide CpG DNAs with immunomodulatory activities.[24] A reduced proportion of Firmicutes may also be compensated by an increased representation of Gram

negative bacteria which are known to express more pro-inflammatory molecules such as lipopolisaccharide (LPS). In agreement with the latter point, we observed a specific association of Gram negative species of the Porphyromonadaceae family in CD patients.

In both our libraries, the Firmicutes comprised mainly the *Clostridium leptum* and *Clostridium coccoides* groups, along with a number of unidentified species. These two groups have been described as essential components of the human indigenous intestinal microbiota.[25] They contain all known microorganisms producing large amounts of butyrate, which is not only the main energy source for colonic epithelial cells,[26] but also inhibits pro-inflammatory cytokine mRNA expression in the mucosa, by NFκB activation and IκB degradation.[27] The loss of butyrate producers observed here could upset the dialogue between host epithelial cells and resident microorganisms, hence contributing to the development of CD-associated ulcerations.

Our results emphasize the advantages of molecular methods over culture-based approaches for comprehensive description of complex microbial communities.[28] Indeed, we found 69% totally novel OTUs in CD patients, compared to 34% in healthy subjects, further supporting the existence of microbial dysbiosis in CD[1] [22] (fig 1 and 2, respectively). Known OTUs identified in this study (corresponding to previously cultured microorganisms) corresponded to species known to reside in the human[28] and pig[29] intestinal tract. These results are in good agreement with previous studies of human feces using molecular approaches.[22] Together, 16S rRNA sequence-based studies of the human intestinal microbiota have led to the recognition of over 50% totally novel species,[30][31] suggesting that this community may be far more complex than previously thought.[6]

Our use of a DNA macroarray-based strategy to analyze two large metagenomic libraries yielded evidence that CD may be characterized by a reduction in normal anaerobic bacterial diversity, especially among the Firmicutes, and did not support the role of a specific pathogen.

Although our study only involved a limited number of subjects (6 healthy volunteers and 6 patients), our results lead to the same conclusions whether samples were pooled in the metagenomic approach or taken individually using the FISH method. Our observations point to formerly unrecognized specificities of the fecal microbiota of CD patients. Nevertheless, the number of subjects analysed was low since the metagenomic approach cannot be applied to a large set of samples. Our results should be confirmed by an epidemiological investigation focussing on the reduction in Firmicutes as well as over-representation of species such as the "uncultured Porphyromonadaceae" species. Specific diagnostic tools could be designed to detect these. In addition, these would allow a more targeted use of antimicrobials or probiotics.

Genetic exploration of metagenomic libraries will further provide an access to the identification of genes and functionalities specific of Crohn's disease. This is currently under investigation.

Acknowledgments.

This research was supported by the French Ministry of Research. We thank C. Dossat, J. Castresana, P. Seksik, P. Robe, D, M. Aguilera and the members of the CRB-GADIE INRA platform for their excellent technical assistance. We are also grateful to I. Mangin, A. Sghir and D. Le Paslier for fruitful scientific discussions.

Footnotes.

Conflict of interest: None declared.

References

1. Guarner F, Malagelada JR. Gut flora in health and disease. *Lancet* 2003;**360**:512-9.
2. Hugot JP, Chamaillard M, Zouali. H, *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;**411**:599-603.
3. Tamboli CP, Neut C, Desreumaux P, *et al*. Dysbiosis in inflammatory bowel disease. *Gut* 2004;**51**:1-4.
4. Moore WE and Holdeman LV. Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Appl Microbiol* 1974;**27**:961-79.
5. Hooper L V, Gordon J I. Commensal host-bacterial Relationships in the gut. *Science* 2001;**292**:1115-8.
6. Seksik P, Rigottier-Gois L, Gramet G, *et al*. Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. *Gut* 2003;**52**:237-42.
7. Ott SJ, Musfeldt M, Wenderoth DF, *et al*. Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut* 2004;**53**:685-93.
8. Venter JC, Remington K, Heidelberg JF, *et al*. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 2004;**304**:66-74.
9. Rondon MR, August PR, Bettermann AD, *et al*. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol* 2000;**66**:2541-7.
10. Courtois S, Cappellano CM, Ball M, *et al*. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl Environ Microbiol* 2003;**69**:49-55.
11. Stein JL, Marsh TL, Ying Wu K, Shizuya H, *et al*. Characterization of Uncultured Prokaryotes: Isolation and Analysis of a 40-Kilobase-Pair Genome Fragment from a Planktonic Marine Archaeon. *J. Bacteriology* 1996;**178**:591-9.
12. Godon JJ, Zumstein E, Dabert P, *et al*. Molecular Microbial Diversity of an Anaerobic Digestor as Determined by Small-Subunit rDNA Sequence Analysis. *Appl Environ Microbiol* 1997;**63**:2802-13.
13. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;**22**:4673-80.
14. Castresana J. Selection of Conserved blocks from multiple alignments for their in phylogenetic analysis. *Mol Biol Evol* 2000;**17**:540-52.
15. Massana R, Castresana J, Balague V, *et al*. Phylogenetic and Ecological analysis of novel Marine Stramenopiles. *Appl Environ Microbiol* 2004;**64**:171-3.
16. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980;**16**:111-20.
17. Felsenstein J. PHYLIP Phylogeny Inference Package version 3.5c. Department of Genetics, University of Washington, Seattle 1993.
18. Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 1996;**12**:357-8.
19. Rossello-Mora R, Amann R. The species concept for prokaryotes. *FEMS Microbiol Rev* 2001;**25**:39-67.
20. Cole JR, Chai B, Marsh TL, *et al*. Ribosomal Database Project. The Ribosomal Database Project RDP-II: previewing a new autoaligner that allows regular updates and the prokaryotic taxonomy. *Nucleic Acids Res* 2003;**31**:442-3.
21. Good IJ. The population frequencies of species and the estimation of population parameters. *Biometrika* 1953;**40**:237-64.

22. Mangin I, Bonnet R, Seksik P, *et al*. Molecular inventory of faecal microflora in patients with Crohn's disease. *FEMS Microbiol Ecol* 2004;**50**:25-36.
23. Sartor RB. Therapeutic manipulation of the enteric microflora in inflammatory bowel diseases: antibiotics, probiotics, and prebiotics. *Gastroenterology* 2004;**126**:1620-33.
24. Sartor RB. Targeting enteric bacteria in treatment of inflammatory bowel diseases: why, how, and when. *Curr Opin Gastroenterol* 2003;**19**:358-65.
25. Duncan SH, Barcenilla A, Stewart CS, *et al*. Acetate utilization and butyryl coenzyme A CoA:acetate-CoA transferase in butyrate-producing bacteria from the human large intestine. *Appl Environ Microbiol* 2002;**68**:5186-90.
26. Barcenilla A, Pryde SE, Martin JC, *et al*. Phylogenetic Relationships of Butyrate-Producing Bacteria from the Human Gut. *Appl Environ Microbiol* 2000;**66**:1654-61.
27. Segain JP, Raingeard de la Bletiere D, Bourreille A, *et al*. Butyrate inhibits inflammatory responses through NFκB inhibition: implications for Crohn's disease. *Gut* 2000;**47**:397-403.
28. Suau A, Bonnet R, Sutren M, *et al*. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl Environ Microbiol* 1999;**65**:4799-807.
29. Leser TD, Amenuvor JZ, Jensen TK, *et al*. Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. *Appl Environ Microbiol* 2002;**68**:673-90.
30. Hold GL, Pryde SE, Russel VJ, *et al*. Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. *FEMS Microbiol Ecol* 2002;**39**:33-39.
31. Hayashi H, Sakamoto M, Benno Y. Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods. *Microbiol Immunol* 2002;**46**:535-48.

Figure legends

Figure 1 Construction and screening of Metagnomic libraries. Bacterial cells were recovered by Nycodenz density gradient centrifugation from fecal samples of 6 healthy subjects and 6 patients with CD in remission. High-molecular-weight bacterial DNA fragments (38 to 48 kbp) were isolated. The metagenomic DNA was then cloned into fosmids. Screening for 16S rRNA genes in the fosmid libraries was performed by DNA hybridization on membrane nylon. All sequences were determined using four primers (ACM 008 F, ADM 330 F, SSM 1100 R and TTM 1517 R) specific for bacteria (table 1). Assembled sequences covering approximately positions 200 to 1300 were selected for subsequent phylogenetic analysis. For each library, sequences were aligned using the ClustalW v1.83 program. Distance matrices were computed with DNADIST v3.6 and trees were constructed with NEIGHBOR v3.6.

Figure 2 Phylogenetic relationships among the dominant OTUs identified in the fecal microbiome of healthy subjects (library based on 16S rRNA sequencing of 536 clones). The identification number (x) corresponds to the name of the metagenomic clone, which has been submitted to GenBank under the name Adhufec-x-ml for the healthy-subject library (accession numbers AY850400 to AY850487). The identity of culturable bacteria and clone names are available from Genbank or RDP. When they are indicated side by side with clone identification number (x), they correspond to the same OTU. Number of clones for each OTU is inserted in brackets. The scale bar represents 10% sequence divergence. Bootstrap values are based on 1000 replications.

Figure 3 Phylogenetic relationships among the dominant OTUs identified in the fecal microbiome of Crohn's disease patients (library based on 16S rRNA sequencing of 654 clones). The identification number (x) corresponds to the name of the metagenomic clone, which has been submitted to GenBank under the name Cadhufec-x-ml for the CD library (accession numbers AY850488 to AY850541). The identity of culturable bacteria and clone names are available from Genbank or RDP. When they are indicated side by side with clone numbers (x), they correspond to the same OTU. Number of clones for each OTU is inserted in brackets. The scale bar represents 10% sequence divergence. Bootstrap values are based on 1000 replications.

Figure 4 Comparison of the libraries derived from healthy subjects and patients with Crohn's disease. The number of ribotypes belonging to each group and subgroup were compared between the two libraries. The data were analyzed with the χ^2 test. Stars indicate significant differences between the healthy subjects and CD patients ($p<0.025$).

Figure 5 Fluorescence in situ hybridization (FISH) of fecal samples,coupled to flow cytometry. The 16S rRNA of the three major groups of bacteria was targeted by FISH. Fecal samples from six healthy donors and six CD patients were analyzed independently. The mean bacterial percentages are reported in the histogram (black columns for healthy donors and white columns for CD patients). The healthy subjects and CD patients were compared by using Student's t test for two independent sets of samples. The asterisk indicates a significant difference between healthy subjects and CD patients for the *Clostridium leptum* group ($P < 0.02$).

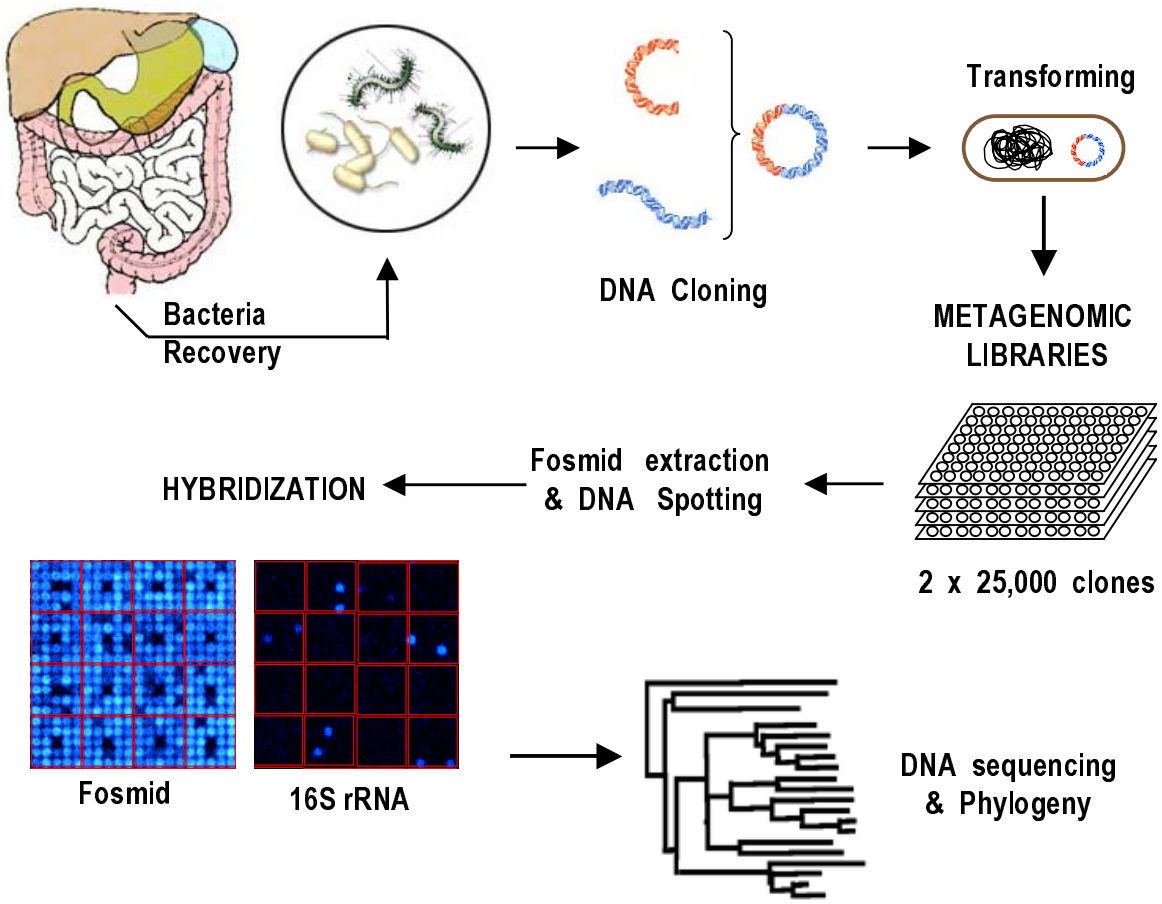

Figure 1

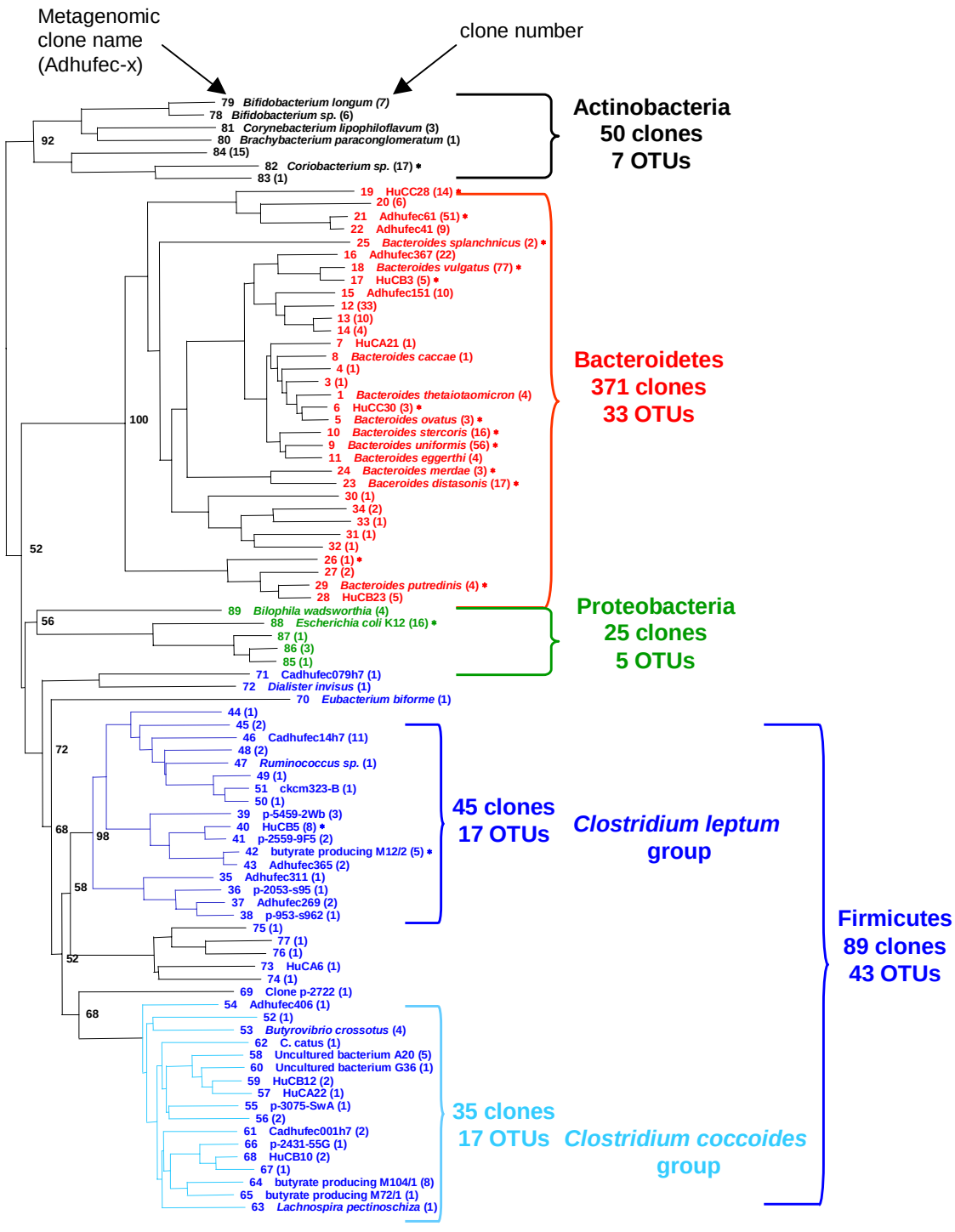

Figure 2

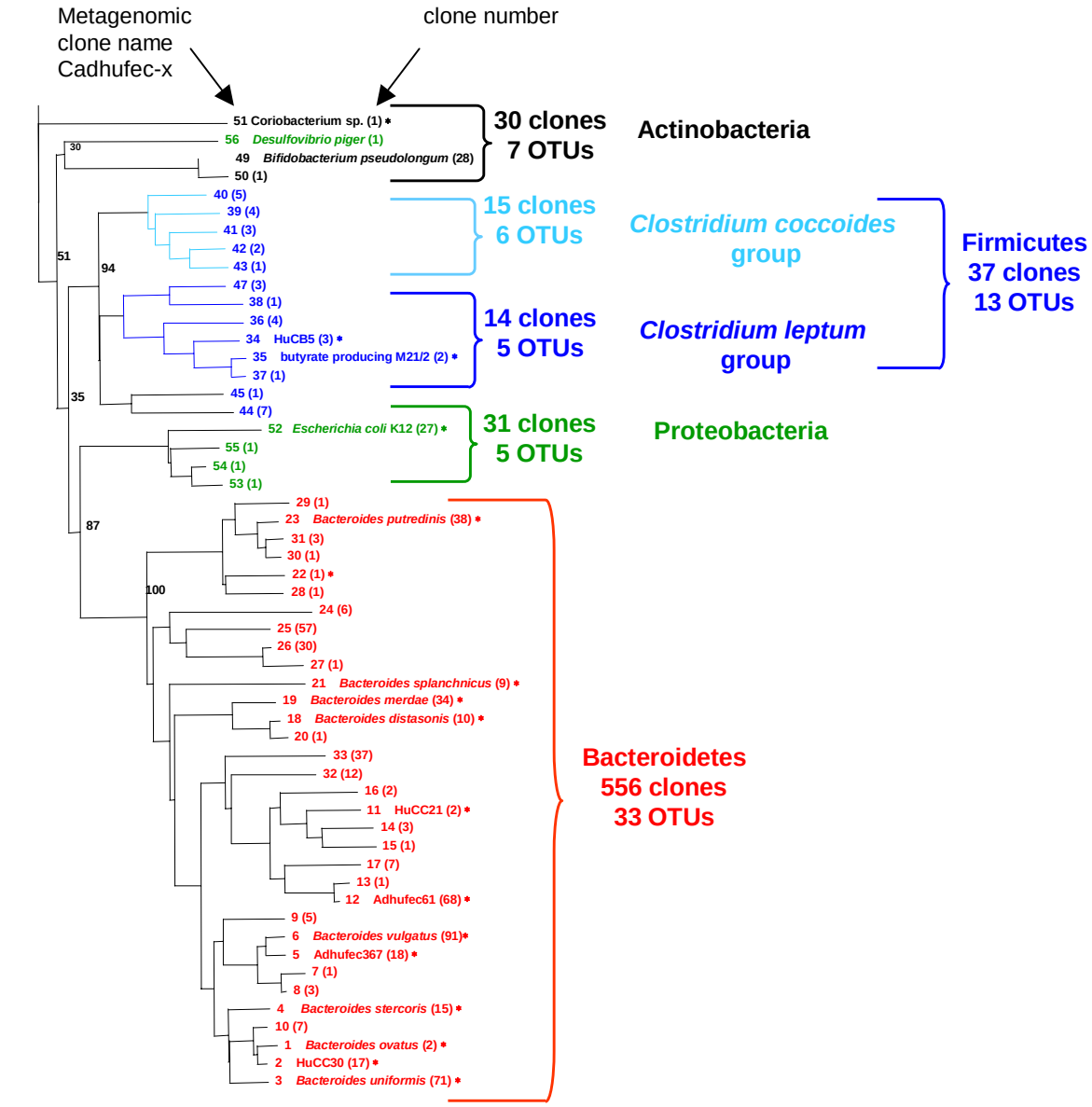

Figure 3

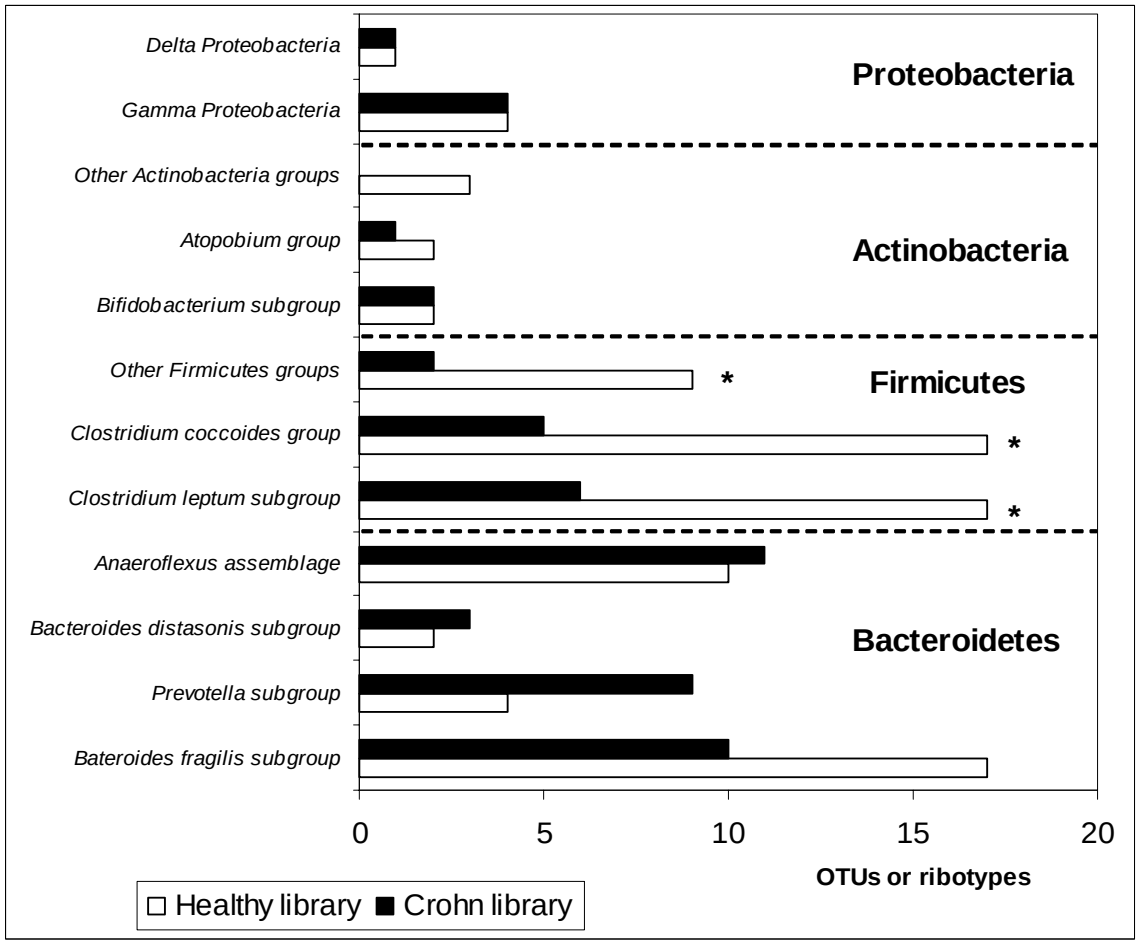

Figure 4

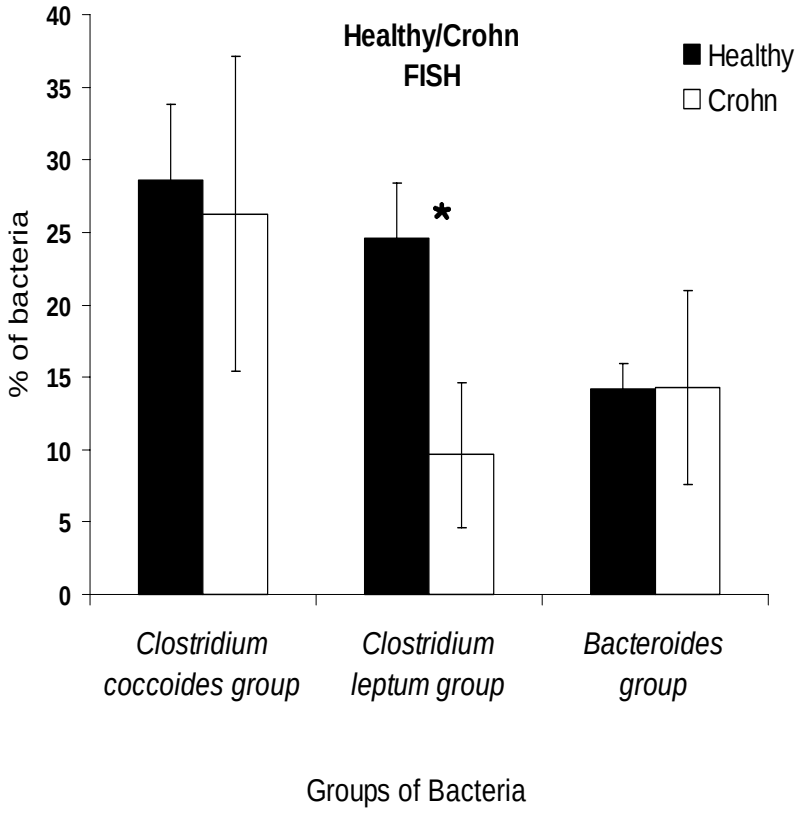

Figure 5