Molecular diversity and phylogeny of human colonic bacteria

A. Sghir[1], J. Doré[1], R.I. Mackie[2]

[1]Unité d'Ecologie et Physiologie du Système Digestif, INRA, 78352 Jouy-en-Josas, France
[2]Animal Sciences Department, University of Illinois at Urbana-Champaign, USA

ABSTRACT

The comparison of microscopic and culture counts clearly indicates the existence of an important, yet uncultured fraction within the human gut microbial community. We herein illustrate the application of modern molecular microbial ecology techniques based on sequence comparisons of 16S ribosomal DNA for the specific and sensitive evaluation of the species diversity and the composition of the human digestive microbiota. In our studies, comparative sequencing of 110 cloned rDNAs obtained using PCR technology from a human fecal DNA extract indicated the presence of 3 major phylogenetic lineages within the human fecal microbiota: the *Bacteroides* group and the sub-groups *Clostridium leptum* and *Clostridium coccoides*. Of the 42 « molecular species » identified based on a sequence similarity cutoff of 97%, only a quarter were related to previously characterized microbial species confirming the occurrence of a large fraction of yet uncultured microorganisms. Quantitative dot-blot hybridization with total RNA extracts of fecal samples from 10 human volunteers, using a panel of rRNA-targeted oligonucleotide probes of increasing specificity, indicated that within the Bacterial domain, which was predominant, the *Bacteroides* group and the subgroups *C. leptum* and *C. coccoides* represented 35.5 $\pm$ 4.5%; 30.4 $\pm$ 6.5% and 19.7 $\pm$ 2.2% of total Bacterial 16S rRNA, respectively. Adding probes specific for the Enterics and *Lactobacillus* groups and the genus *Bifidobacterium*, 92.4% of all Bacterial rRNA was accounted for. Results from previous studies described in the literature will be reviewed and compared to our findings.

Description of the ecosystem

The entire length of the adult human colon is about 150cm and consists of the ascending colon, transverse colon and descending colon. The human colon can be considered as a large continuous culture fermentor with an approximate volume of 500ml and an average of 220g of digestive contents [4] containing a complex microbiota interacting with living host cells. An average of 10^{11} bacterial cells per gram of content belonging to hundreds of different species have been reported [15]. Complex fermentations carried out by the microbiota in the colon produce a variety of metabolites which can be beneficial or detrimental to the health and nutritional status of the host. Because of the inaccessibility of the organ, difficulties are encountered in trying to sample from the colon and much of the pioneer work used fecal rather than colonic samples [23]. Early work analyzing the colonic microbiota has focused on enumeration and identification of microbial isolates. Despite the use of various media, microbiologists have found a remarkable difference in the number and morphology of organisms when comparing natural samples with those obtained from

Microbial Biosystems: New Frontiers
Proceedings of the 8th International Symposium on Microbial Ecology
Bell CR, Brylinsky M, Johnson-Green P (ed)
Atlantic Canada Society for Microbial Ecology, Halifax, Canada, 1999.

enrichment cultures and isolates from habitat-simulating or selective media. Furthermore, these approaches only examined strains of the major bacterial groups that were dominant in this complex and diverse ecosystem. These observations lead to the postulation of the "plate count anomaly" with the major concept being that a large fraction of bacterial species observed in the natural environment fail to grow under synthetic laboratory conditions.

Application of modern molecular approaches, such as sequence analysis of conserved genes, has allowed the unambiguous affiliation of organisms to their nearest phylogenetic neighbor and a better description of the complexity of the gastrointestinal ecosystem. This will continue, in combination with other molecular techniques, to help microbial ecologists determine the role microorganisms play in their natural environment.

Species composition of the colonic microbiota

Knowledge concerning numbers and species of bacteria in the human colon is important because it provides an index of the metabolic potential of the colonic biota. A considerable amount of effort has been devoted to the isolation and identification of the numerically predominant species of the colonic flora. Over 90% of the bacteria are obligate anaerobes. The numerically predominant genera are presented in Table 1. Among the numerically predominant species are at least five species of *Bacteroides*, one species of *Fusobacterium*, at least four species of *Eubacterium*, one species of *Peptostreptococcus* and four species of *Bifidobacterium* [15,16]. The species composition of different individuals is similar but differs proportionately and quantitatively between individuals. Although Moore and Holdeman [16] were able to account for 90% of the fecal organisms that could be seen on gram-stained slides most estimates are much lower than this. Finegold *et al*. [8] estimated that only about 10% of the human colonic organisms were cultivated while Wilson *et al*. [26] recovered only 58% of the microscopic count. A summary of the cultivated fraction of the colonic biota described by different authors is presented in Table 2. Thus, it is likely that some major groups of colonic bacteria have not yet been identified because they have not been cultivated thus far.

Table1. Culturable human intestinal microbiota.

Percent of total anaerobes		
Genus	**ref [15]**	**ref [8]**
Bacteroides	30	56
Eubacterium	26	14
Bifidobacterium	11	4
Peptostreptococcus	9	4
Fusobacterium	8	0.1
Ruminococcus	4	9
Clostridium	2	2
Lactobacillus	2	1
Streptococcus	2	6
others	3	1

Table 2. Cultivated fraction of the human fecal microbiota

Reference	Cultivated fraction (%)
[15]	88%
[8]	24%
[12]	14%
[26]	58%
Doré et al. unpublished	20 -30%

Modern molecular microbial ecology techniques based on sequence comparisons of nucleic acids provide a culture-independent classification scheme which predicts natural evolutionary relationships and can be applied to specific and sensitive evaluation of all members of complex environmental ecosystems, including the digestive tract of animals and humans. An example of the power of these modern molecular ecology techniques is provided by the analysis of 16S rRNA sequences [18,13]. The highly conserved regions of the rRNA molecule can serve as primer binding sites for *in vitro* amplification by PCR. The more conserved regions are also useful as targets for universal probes that react with all living organisms or for discriminating between broad phylogenetic groups such as the domains Archaea, Bacteria and Eucarya. The more variable sequence regions are appropriate for the design of genus, species and sometimes even strain-specific hybridization probes. Ribosomal DNA amplification, cloning and comparative sequencing allow a complete assessment of microbial diversity within complex ecosystems. Approaches independent of growth conditions and media confirm the extreme diversity of the human fecal microbiota and confirm the occurrence of numerous bacterial species that have so far escaped cultivation and characterization by comparative rRNA sequencing.

Inventory of microbial species diversity within a human fecal sample: 16S rDNA gene amplification, cloning and sequencing

Over 450 16S rDNA clones were obtained from bulk DNA extracted from a human fecal sample. Of these, 110 rDNA clones were partially sequenced (600 bases at 5' end) and their sequences aligned with previously determined sequences. With the exception of 3 chimeras and one sequence distantly related to any of the known sequences and that could not be aligned properly, 42 Operational Taxonomic Units (OTUs) were defined on the basis of a degree of sequence similarity equal to or greater than 97%. Of these 42 OTUs, 3/4 corresponded to microorganisms with no known relatives in the Ribosomal Database Project. The vast majority of rDNA clones characterized belonged to 3 monophyletic clusters : the *Bacteroides* group (40 clones), the *Clostridium coccoides* subgroup (39) and the *Clostridium leptum* subgroup (24) (Table 3). Approximately half of the OTUs were represented by a single clone, suggesting that microbial diversity of the dominant populations of the ecosystem was still incompletely described. Also, no rDNA sequences were found to represent the genus *Bifidobacterium*. This could be attributed to a PCR bias since one of the primers used had 2 mismatches with bifidobacterial rDNA sequences.

Indeed, PCR primers described by Lane [11], which have been used by many other workers [19,20], consistently failed to work with some "difficult" samples. DNA samples

Table 3. Phylogenetic analysis of the diversity of human intestinal microbiota

Phylogenetic groups	*Bacteroides*	*C. leptum*	*C. coccoides*	Others
Clones	40	24	39	4
TU	14	10	14	4
NTU	10	8	9	4

TU = taxonomic unit > 97% sequence similarity
NTU = new taxonomic unit < 97% sequence similarity

and culture media confirm the extreme diversity of the human fecal microbiota, and indicate the occurrence of numerous species that have so far escaped cultivation and characterization. They extracted from various sources, including deep sea sediments, oral bacteria, and bacteria isolated from epilithon (biofilms associated with stones in lotic habitats), were found to be poor templates for amplification of the 16S rRNA gene with amplimers such as 27f and either 1492R or 1392R [14]. The wide adoption of amplimers 27f, 1392R, and 1492R has an empirical basis, and although their utility for investigating the molecular ecology of natural microbial communities is often assumed, they have not been systematically tested. Marchesi et al. [14] described a new set of universal primers (63f and 1387R) for the domain Bacteria and demonstrated that these primers were more successful than the broadly used 27F and 1392R. However, these primers have many mismatches with their target in the 16S rDNA sequences of some predominant intestinal bacteria such as *Enterococcus* (4 mismatches), *Bifidobacterium, Clostridium, Ruminococcus*, and *Eubacterium* (3 mismatches) as well as *Lactobacillus* (2 mismatches). This emphasizes the need for designing and developing new universal primers to cover the whole bacterial domain. A new primer was designed and is being tested in our laboratory (Sghir et al. unpublished).

Molecular ecology of the human fecal microbiota

A panel of 10 probes (Table 4) was designed and used to quantify the relative proportion of rRNAs of selected microbial groups within fecal samples from 10 healthy human adults using quantitative dot-blot hybridizations as described previously [5]. Using a universal probe as reference and domain specific probes for Bacteria, Archaea and Eucarya, the domain Bacteria represented by far the greatest fraction (78.2 $\pm$ 4.3%) of total rRNA. Archaea and Eucarya only represented 0.28 and 2.02% of total rRNA, respectively. Within the domain Bacteria, using group specific probes together with a domain probe as reference, ribosomal RNAs of the groups *Bacteroides*, Enterics, *Lactobacillus, C. leptum, C. coccoides* and the genus *Bifidobacterium* represented 35.5 $\pm$ 4.5%; 2.7 $\pm$ 0.7%; < 1.0%; 30.4 $\pm$ 6.5%; 19.7 $\pm$ 2.2% and 4.1 $\pm$ 0.7% of total Bacterial rRNA, respectively. Thus, using a panel of six group specific additive probes, 92.4% of all Bacterial rRNA was accounted for. The comparison of relative cultural counts and relative rRNA proportions, which was possible for *Bacteroides, Bifidobacterium* and *E. coli*, indicated that specific

Table 4. 16S rRNA Oligonucleotide probes used in the cited studies

Target group	Probe/Primer sequence (5'-3')	Name (OPD Nomenclature)	Reference
Universal	GACGGGCGGTGTGTACAA	S-*-Univ-1390-a-A-18	[28]
Bacteria	GCTGCCTCCCGTAGGAGT	S-D-Bact-0338-a-A-18	[2]
Archaea	GTGCTCCCCCGCCAATTCCT	S-D-Arch-0915-a-A-20	[2]
Eucarya	ACCAGACTTGCCCTCC	S-D-Euca-0502-a-A-16	[2]
Bacteroides-Prevotella Porphyromonas group	GCACTTAAGCCGACACCT	S-*-Bacto-1080-a-A-18	[5]
Enteric group	CTTTTGCAACCCACT	S-G-Enter-1432-a-A-15	*B*
Bifidobacterium	CCGGTTTTMAGGGATCC	S-G-Bif-1278-b-A-17	[5]
Lactobacillus group	YCACCGCTACACATGRAGTTCCACT	S-G-Lacb-0722-a-A-25	[21]
C. leptum sub-group	GTTTTATCAACGGCAGTC	S-G-Clept-1089-a-A-18	*C*
Bacteroides fragilis group	AACACTTTCAAACGCCGAG	S-*-Bfra-0602-a-A-19	[9]
C. coccoides-E rectale group	GCTTCTTAGTCARGTACCG	S-*-Erec-0482-a-A-19	[9]

[a]Oligonucleotide Probe Database. Alm *et al.* (1996).
[b]Pochart et al., unpublished
[c]Dore et al., unpublished

also emphasize the importance of two monophyletic sub-groups within the *Clostridium* and relatives, namely the *C. leptum* subgroup and *C. coccoides* phylogenetic group.

Recently Franks et al. [9] have developed and applied six 16S rRNA probes for major species and groups of anaerobic intestinal bacteria to enumerate bacterial populations in the feces of healthy human volunteers using the fluorescent *in situ* hybridization (FISH) technique. Using a *Bacteroides fragilis* group specific probe and a *Bacteroides distasonis* species-specific probe in combination accounted for 20% of the fecal flora, while the *Clostridium coccoides-Eubacterium rectale* group counted with Erec482 probe accounted for 29% of the biota. The proportion of *Bifidobacterium* enumerated (3%) was comparable to the culturable counts [12]. Franks et al. [9] also used a probe called Lowgc2P to detect 12% of the total human fecal biota. This agreed with the 1 to 10% proportion of the total human fecal 16S rRNA detected with Lowgc2P in the study of Wilson and Blitchington [26], who indicated that the organisms phylogenetically related to rDNA clones for whom the probe was designed grew very poorly on culture plates. Two other microorganisms detected by rDNA analysis were not previously known to colonize the human intestine. *Lachnospira pectinoschiza* is readily culturable from the rumen [3] but has never been detected in the numerous studies of bacterial colonies isolated from the human colonic biota.

Results from Wang et al. [25] showed differences between the results of PCR and the culture method when trying to amplify the 16S rDNA of some fecal bacteria. *C. clostridiiforme* had a high PCR titer for all humans and animal fecal samples, but this species is not often listed as a predominant species by culture methods [7,15,17]. One

reason for this difference is that PCR methods detect bacteria *in situ*, whereas culture methods detect bacteria after enrichment.

Sharp and MacFarlane [22] reported, using 16S rRNA analysis of chemostat enrichments fermenting resistant starch, novel species of clostridia in colonic bacterial populations, and detected *C.butyricum*-related species producing butyric acid from this substrate. This bacterium was not detectable and could not be identified using phenotypic techniques. Kimura et al. [10] analyzed fecal populations of bifidobacteria and lactobacilli collected from 10 subjects. Their results showed that when using ribotyping and pulsed field gel electrophoresis each subject harbored a unique collection of bifidobacteria and lactobacilli. Our results regarding *Bifidobacterium* species isolated from a chemostat enrichment using FOS as the sole source of carbon and energy showed a high diversity profile of this genus when the isolates were analyzed by ribotyping, since we detected 9 different ribotypes among 15 isolates [21]. Zoetendal et al. [29] using temperature gel gradient electrophoresis profiling of human fecal samples showed that each healthy human has their own unique fecal microbiota and that the dominant active flora of a given individual is stable over time. The reasons for this unique individuality are likely to be found in host factors.

Conclusion

Applying this wealth of varied molecular techniques has helped microbial ecologists reveal new information on the composition of the normal fecal microbiota, and to more accurately describe natural variation in bacterial populations. Although the contribution of molecular techniques to microbial ecology in general must be significant, one molecule alone cannot completely cover all facets of microbial ecology [27]. Thus the function of an ecosystem must be deduced from analysis of genes expressed through rRNA, mRNA and proteins. The network of broad physiological interactions must also be verified by *in vitro* reconstitution of isolates. In addition, all these strategies must be performed in conjunction with a thorough physical and chemical analysis of the natural samples [24].

Acknowledgements

This work was funded in part by grants from the Bureau des Ressources Génétiques and the EU FLAIR program under project N° CT-97-3035.

References

1. Alm EW, Oerther DE, Larsen N, Stahl DA, and Raskin L (1996) The oligonucleotide probe database. Appl Environ Microbiol 10:3557-3559
2. Amann RI, Krumholz L, and Stahl DA (1990) Fluorescent oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. Journal of Bacteriology 172:762-770
3. Bryant MP, and Small N (1956) Characteristics of two new genera of anaerobic curved rods isolated from rumen of the cattle. J Bacteriol 72:22-26
4. Cummings JH, Banwell JG, Englyst HN, Coleman N, Segal I, Bersohn, D (1990) The amount and composition of large bowel contents. Gastroenterol 98:A408
5. Doré J, Sghir A, Hannequart-Gramet G, Corthier G, Pochart P (1998) Design and evaluation of a 16S rRNA-targeted oligonucleotide probe for specific detection and quantification of human faecal Bacteroides populations System Appl Microbiol 21:65-71

6. Doré J, Gramet G, Goderel I, Pochart P (1998) Culture independent characterization of human faecal flora using rRNA-targeted hybridization probes. Gen select evol 30 suppl.1:287-296
7. Drasar B, Robert K (1990) Control of the large bowel microflora. In: M. J. Hill and P. D. Marsh (ed), Human microbial ecology. CRC Press, Inc., Boca Raton, Fla
8. Finegold SM, Sutter VL, Mathisen GE (1983) Normal indigenous intestinal flora. In: Human Intestinal Microflora in Health and Disease, Henteges DJ (ed), Academic Press, New York
9. Franks H, Harmsen HJM, Raangs GC, Jansen GJ, Schut F, Welling G.W (1998) Variations of bacterial pouplations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Appl Environ Microbiol 64:3336-3345
10. Kimura K, McCartney AL, McConnell MA Tannock G. W (1997) Analysis of fecal populations of bifidobacteria and lactobacilli and investigation of the immunological responses of their human hosts to the predominant strains. Appl Environ Microbiol 63:3394-8
11. Lane DJ (1991) 16S/23S rRNA sequencing. In: E Stackebrandt and M Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Wiley and Sons, Chichester, United kingdom
12. Langendijk P, Schut F, Jansen GJ, Raangs GC, Camphuis GR, Wilkinson MF, Welling GW (1995) Quantitative fluorescence in situ hybridization of Bifidobacterium spp. with genus-specific 16S rRNA-targeted probes and its application in fecal samples. Appl Environ Microbiol 61:3069-3075
13. Mackie RI, Doré J. (1997) Molecular ecology of gastrointestinal ecosystems. In: Progress in Microbial Ecology SBM/ICOME Sao Paulo, Brazil
14. Marchesi JR, SatoT, Weightman AJ, Tracey AM, Fry JC, Hiom SJ, Wade WG (1998) Design and evaluation of useful Bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. Appl and Environ Microbiol 64:795-799.
15. Moore WEC, Holdeman LV (1974) Human fecal flora : the normal flora of 20 Japanese- Hawaiiens. Appl Environ Microbiol 27:961-979
16. Moore WE, Cato EP, Holdeman LV (1978) Some current concepts in intestinal bacteriology. Amer J Clin Nutr 31:S33-42
17. Moore WEC, Moore LH (1995) Intestinal flora of populations that have a high risk of colon cancer. Appl Environ Microbiol 61:3202-3207
18. Olsen GJ, Woese (1993) CR Ribosomal RNA: a key to phylogeny. FASEB J 7:113-123.
19. Parks RJ, Cragg BA, Fry JC, Herbert RA, Wimpenny JWT (1990) Bacterial biomass and activity in deep sediments layers from the Peru margin. Philos Trans R Soc Lond B 133:139-153
20. Reysenbach AL, Wikham GS, Pace NR (1994) Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. Appl Environ Microbiol 60:2113-2119
21. Sghir A, Chow JM, Mackie RI (1998) Phenotypic and molecular characterization of 15 Bifidobacterium isolates from human intestinal contents and their potential use as probiotics

for humans. abstract 258, 5ième congrès de la Société Française de Microbiologie 27-29 Avril, Lille. SFM, Institut Pasteur, Paris, France.
22. Sharp R, MacFarlane GT (1997) 16S rRNA analysis of resistant starch chemostat enrichments reveals novel species of clostridia in colonic bacterial populations. Bioscience Microflora 16 Supplement
23. Simon GL, Gorbach SL. Intestinal microflora (1982) Med Clin North Am May 66(3):557-74
24. Stackebrandt E (1999) Biodiversity of Prokaryotes: Current updates on bacterial taxonomy. 1st Arab environmental biotechnology conference March 5-8th 1999, Abu Dhabi, UAE
25. Wang RF, Cao WW, Cerniglia CE (1996) PCR detection and quantitation of predominant anaerobic bacteria in human and animal fecal samples. Appl Environ Microbiol 62:1242-1247
26. Wilson KH, Blitchington RB (1996) Human colonic biota studied by ribosomal DNA sequence analysis. Appl Environ Microbiol 62:2273-2278
27. Von Wintzingerode F, Gobel U, Stackebrandt E (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol Rev 21:213-229.
28. Zheng D, Alm EW, Stahl DA, Raskin L (1996) Characterization of universal small-subunit rRNA hybridization probes for quantitative molecular microbial ecology studies. Appl Environ Microbiol 62:4504-4513
29. Zoetendal EG, Akkermans AD, WM De Vos (1998) Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. Appl Environ Microbiol. 64:3854-3859