

Minireview

Diversity of the human gastrointestinal tract microbiota revisited

Mirjana Rajilić-Stojanović,* Hauke Smidt and Willem M. de Vos

Laboratory of Microbiology, Wageningen University, Hesselink van Suchtelenweg 4, 6703 CT Wageningen, the Netherlands.

Summary

Since the early days of microbiology, more than a century ago, representatives of over 400 different microbial species have been isolated and fully characterized from human gastrointestinal samples. However, during the past decade molecular ecological studies based on ribosomal RNA (rRNA) sequences have revealed that cultivation has been able only to access a small fraction of the microbial diversity within the gastrointestinal tract. The increasing number of deposited rRNA sequences calls for the setting up a curated database that allows handling of the excessive degree of redundancy that threatens the usability of public databases. The integration of data from cultivation-based studies and molecular inventories of small subunit (SSU) rRNA diversity, presented here for the first time, provides a systematic framework of the microbial diversity in the human gastrointestinal tract of more than 1000 different species-level phylogenetic types (phylotypes). Such knowledge is essential for the design of high-throughput approaches such as phylogenetic DNA microarrays for the comprehensive analysis of gastrointestinal tract microbiota at multiple levels of taxonomic resolution. Development of such approaches is likely to be pivotal to generating novel insights in microbiota functionality in health and disease.

Introduction

The human gastrointestinal microbiota represents a complex ecosystem that consists of bacteria, archaea,

yeasts and filamentous fungi (Finegold *et al.*, 1974; Miller and Wolin, 1986). While Archaea and Eukarya are represented by members of a single phylum each, bacterial community within the human gastrointestinal tract is exceptionally diverse. Members of nine bacterial phyla were found to inhabit the human gastrointestinal tract of which Firmicutes, Bacteroidetes and Actinobacteria are dominant (Bäckhed *et al.*, 2005). Members of Proteobacteria are also common and diverse, but they are usually secondary to the above. Recent reports indicated that viruses represent another important constituent, as more than 1200 viral genotypes were identified in human faeces with a density of up to 10^9 virions per g of dry material (Breitbart *et al.*, 2003; Zhang *et al.*, 2006).

Many studies have attempted to describe the normal gastrointestinal microbiota in terms of microbial species that inhabit healthy humans. However, individual differences in the microbiota composition make the definition of this concept very challenging. Because only a limited number of individuals has been subjected to the analysis of intestinal microbial diversity, the description of the normal composition remains incomplete. Both molecular, small subunit (SSU) (16S and 18S) ribosomal RNA (rRNA)-based, and traditional cultivation studies show that the mere selection of some different subjects is sufficient for the discovery of novel intestinal inhabitants (Holmstrom *et al.*, 2004; Eckburg *et al.*, 2005; Kitahara *et al.*, 2005; Wang *et al.*, 2005).

Even though it is clear that the present view of the microbial diversity within the human gastrointestinal tract is incomplete, it is difficult to estimate to what degree. Most work has been performed in an uncoordinated way over more than a century, by many different groups, and by using a wide variety of approaches. As a coherent meta-analysis of the findings of those numerous studies is still lacking, there is no general agreement even concerning the number of different bacterial species that can inhabit the healthy human intestine. Still, the vast majority of authors refer to an expected diversity of about 400–500 species, but in some of the recent reports the number of 1000 gastrointestinal species appears (Noverr and Huffnagle, 2004; Nicholson *et al.*, 2005; Phillips, 2006). Recently, Bäckhed and coauthors determined that 800

Received 20 November, 2006; accepted 4 May, 2007. *For correspondence. E-mail mirjana.rajilic@wur.nl; Tel. (+31) 317 483118; Fax (+31) 317 483829.

Table 1. Number of subjects whose gastrointestinal microbiota composition was described by SSU rRNA gene sequencing.

Subject	Fraction of microbiota	Number of subjects	Reference
Healthy adult	Total	11	Zoetendal <i>et al.</i> , 1998; Suau <i>et al.</i> , 1999; Bonnet, 2002; Bonnet <i>et al.</i> , 2002; Hayashi <i>et al.</i> , 2002a,b; Eckburg <i>et al.</i> , 2005; Wang <i>et al.</i> , 2005
Healthy infant		4	Favier <i>et al.</i> , 2002; Wang <i>et al.</i> , 2004
Healthy elderly		13	Hold <i>et al.</i> , 2002; Hayashi <i>et al.</i> , 2003; Wang <i>et al.</i> , 2003
Crohn's disease patient		4	Mangin <i>et al.</i> , 2004
Ulcerative colitis patient		2	R. A. Hutson (unpublished)
Healthy adult	<i>Lactobacilli</i>	12	Heilig <i>et al.</i> (2002)
Healthy infant		1	
Healthy adult	<i>Bifidobacteria</i>	5	Satokari <i>et al.</i> (2001)
Healthy adult	Bacteria enriched on mucin-based medium	6	Derrien <i>et al.</i> (2004)
Healthy adult	Butyrate producing bacteria	3	Barcenilla <i>et al.</i> , 2000; Duncan <i>et al.</i> , 2004
Healthy infant		1	

Inventory period from 1998 until the beginning of 2006.

distinct 16S rRNA gene sequences present in the GenBank were derived from the human gastrointestinal samples (Bäckhed *et al.*, 2005). However, a comprehensive picture of the diversity of the microbiota in the gastrointestinal tract is still lacking. This is because the results from cultivation-based studies have never been curated, analysed and integrated with molecular data. Hence, we provide a systematic overview of the gastrointestinal microbiota diversity, with specific attention on the community structure, based on a large body of the critical and phylogenetic analysis of the literature and public databases entries.

The present standpoint of the diversity of the human gastrointestinal microbiota

The first identified bacterial species recovered from a human gastrointestinal sample was *Escherichia coli* isolated in 1885 from children's diarrhoeal faeces (Anonymous, 1981). The description of the human gastrointestinal inhabitants continued throughout the previous century and resulted in the full characterization of over 400 cultivated species (Rajilić-Stojanović, 2007). However, it was the application of molecular techniques that has revolutionized the view of this ecosystem (Furrie, 2006). Application of molecular techniques revealed that the microbiota was significantly more complex than previously anticipated, as only a fraction of the bacteria living in the human intestine can currently be cultured (Suau *et al.*, 1999). The proportion of reported cultivable bacteria varies between studies from 20% for the mucosal samples of three healthy individuals (Eckburg *et al.*, 2005) to 46% for the three elderly Japanese subjects (Hayashi *et al.*, 2003). The other major outcome of the molecular ecology revolution was that the composition of the microbiota is subject-specific, and dominated by yet uncharac-

terized phylotypes (Barcenilla *et al.*, 2000; Hayashi *et al.*, 2002a,b; 2003; Hold *et al.*, 2002; Mangin *et al.*, 2004; Eckburg *et al.*, 2005). This observation is in contrast to the often cited (e.g. Kleessen, 2000; Wang *et al.*, 2002), but likely incorrect, statement that the human gastrointestinal microbiota is diverse but dominated by a very limited number of bacterial species (Drasar and Barrow, 1985).

The first comprehensive study aiming at the characterization of the microbial community of the human intestine based on the SSU rRNA gene sequence analysis, was published in 1999 (Suau *et al.*, 1999). However, prior to that date a preliminary insight into the diversity of the intestinal ecosystem based on SSU rRNA gene sequences was reported in a study that provided a rapid method for profiling the intestinal microbiota (Zoetendal *et al.*, 1998). Since then and until the beginning of 2006, the total gastrointestinal microbiota of 34 subjects and specific groups of 28 subjects have been studied based on the SSU rRNA sequencing approach (Table 1).

Although relatively new, molecular studies have been very productive, and until the beginning of 2006 about 15 000 SSU rRNA gene sequences, which were obtained from human gastrointestinal samples, could be retrieved from public databases. However, most of those sequences originate from a single study, in which each cloned 16S rRNA gene insert was sequenced and deposited, resulting in a public catalogue of 13 335 sequences (Eckburg *et al.*, 2005). This study of Eckburg and colleagues represents the first application of high throughput sequencing applied on the samples of the human gastrointestinal origin, which have enabled better insight into the diversity and abundance of 395 detected phylotypes along the gastrointestinal tract of three individuals. Several subsequent studies confirmed that high throughput sequencing is a powerful approach of analysing a complex microbial ecosystem, such as the human gas-

gastrointestinal microbiota (Bik *et al.*, 2006; Ley *et al.*, 2006). An undesired consequence of such massive sequencing approach is the deposition of large number of identical sequences that has caused extensive redundancy in public databases. This jeopardizes the usefulness of public databases, as it limits, for instance, the recognition of the closest cultivated relative during the BLAST search (Altschul *et al.*, 1990). Such extensive redundancy can be omitted by the deposition of only one representative sequence per identified phylotype. Several strategies can be used for this purpose and they include analysing the uniqueness of the SSU rRNA gene sequence, for instance by restriction analysis of the amplicons (Suau *et al.*, 1999; Heilig *et al.*, 2002), phylogenetic analysis of partial SSU rRNA (Hayashi *et al.*, 2002a,b; Wang *et al.*, 2005), or by simple selection of distinct sequences based on *in silico* identification of unique phylotypes. For the purpose of the analysis of the human gastrointestinal diversity presented in this article all redundant sequences have been excluded.

Phylogenetic analysis of the gastrointestinal microbiota diversity

To assess the diversity of the human gastrointestinal microbiota, we integrated the results obtained in the reported molecular studies of the human gastrointestinal diversity by a careful phylogenetic analysis of the SSU rRNA gene sequences. This step was indispensable because each clone of the SSU rRNA gene is given a distinct name, even when it contains only another copy of the very same gene. Although most studies include phylogenetic analysis of the obtained sequences, the results are not preserved in a publicly available format. Consequently, integration of the results from different studies is impossible without performing extensive phylogenetic analysis. Repetition of this effort leads to an unreasonable loss of energy and time. Therefore, there is an urgent need for establishing a framework for the curation of fast-growing databases of SSU rRNA gene sequences as it is attempted here for those obtained from human gastrointestinal samples.

In addition to the high level of redundancy within public databases, phylogenetic analysis of the SSU rRNA gene sequences obtained in different studies is hampered by the fact that sequences might differ in size and the region of the SSU rRNA gene that was sequenced. A perfectly appropriate tool for the analysis of such data set is not available, but the ARB software seems to be the most suitable (Ludwig *et al.*, 2004). For the purpose of this review, the ARB database release from 2002 was updated with SSU rRNA gene sequences obtained from human gastrointestinal tract samples that were not present in this release but could be recovered in other SSU rRNA gene

databases, such as RDP (Cole *et al.*, 2003), or GeneBank (<http://www.ncbi.nlm.nih.gov/Genbank>). For the selected sequences a distance matrix was calculated and a threshold of 2% sequence divergence was used to define distinct phylotypes. Calculation of similarity (distance) matrices as implemented in ARB has a major advantage compared with the routinely performed BLAST search (Altschul *et al.*, 1990), as an alignment with an ambiguous nucleotide is not considered as a mismatch. As a result, sequences of low quality can have high similarity scores. This is particularly relevant, as for some cultivated organisms SSU rRNA gene sequences are of low quality, and contain a relevant number of ambiguous sequences. However, the quality of the ARB similarity matrix is highly dependent on the quality of the alignment, which is edited by the operator and thus more error-prone than the automatically generated alignment in BLAST.

A similar approach, based on the ARB sequence similarity matrix, was previously used by Bäckhed and colleagues (2005). However, in that study only the number of different phylotypes was determined. In this review, we have identified the distinct phylotypes, and for each monophyletic group with $\geq 98\%$ sequence identity, a representative type sequence was assigned. Because sequences of gastrointestinal isolates were combined with those retrieved in cultivation-independent studies, the representative phylotype sequence was chosen as high quality and the most complete SSU rRNA sequence of the monophyletic group, if possible from the cultured representative. Based on the sequence information and using Parsimony procedures, as implemented in ARB, a phylogenetic tree was constructed (Fig. S1). The sequences which were not present in the ARB database release 2002 were added to the tree with the use of appropriate filters fitting the sequence phylogeny and length.

The analysis of the human gastrointestinal SSU rRNA gene sequences showed that, until the beginning of 2006, 898 distinct bacterial phylotypes were reported in cultivation-independent studies alone. The vast majority of the recovered phylotypes (622; representing 69%) appeared to be subject-specific, which is due to the high level of interindividual variation of the microbiota composition, but also influenced by the fact that the microbiota of only a limited number of individuals was studied. The majority of the phylotypes represents organisms that have not yet been cultured. Among 898 phylotypes, only 158, representing 18% of the total number of phylotypes, correspond to fully characterized isolates. Integration of data from different studies revealed that the contribution of the cultivated organisms in the total diversity assessed by sequencing of the SSU rRNA gene is lower than within individual studies. This is mainly due to the fact that each attempt to describe the gastrointestinal microbiota by

SSU rRNA gene sequencing renders a proportion of novel phylotypes that may be as high as 62% (Eckburg *et al.*, 2005). This has been leading to an ever increasing number of novel phylotypes that are only represented by sequences, rather than by cultured isolates. Moreover, a large proportion of the SSU rRNA sequences corresponding to cultivated species is repeatedly detected (59%; 93 out of 158) when compared with uncultivated repeatedly detected uncultured phylotypes (25%; 183 out of 740).

Cultured versus uncultured

Our current knowledge of the human gastrointestinal microbiota diversity originates from cultivation-based and molecular studies. Thus, generating the most complete view of the known gastrointestinal diversity is possible only through the integration of the results from both types of studies. As SSU rRNA gene sequences are in most cases the only available information for the phylotypes reported in molecular analyses, the integration is possible only on the basis of these sequences.

However, for 46 bacterial and eight eukaryotic human gastrointestinal isolates the SSU rRNA gene has not yet been sequenced (Rajilić-Stojanović, 2007). Consequently, these organisms could not be included in the SSU rRNA based phylogenetic analysis. They were considered as distinct species, which were not recovered in the cultivation-independent studies, although it has to be noted that it can not be excluded that isolates with lacking SSU rRNA sequence match with molecular phylotypes.

The integration of data sets from cultivation-dependent and -independent studies showed that 1183 unique bacterial, three archaeal, and 17 eukaryotic phylotypes are reported as inhabitants of the human gastrointestinal tract. The fully opened phylogenetic tree of the hitherto known diversity of the human gastrointestinal tract is available as a supplementary material (Fig. S1). The compressed version of this phylogenetic tree is depicted in Fig. 1 and summarized distribution of the phylotypes is given in Table 2.

Only 138 species, 12% of the total species richness, were recovered by application of both molecular and cultivation-based approaches (Fig. 2). Remarkably, another 20 phylotypes detected by the cultivation-independent studies matched cultivated and fully described bacterial species, which were not previously isolated from the human gastrointestinal tract (Table 3). Examples of such species include *Allisonella histaminiformans* and *Phascolarctobacterium faecium*, which are intestinal isolates from other mammals (Del Dot *et al.*, 1996; Garner *et al.*, 2002). There are, however, also isolates such as *Aquabacterium commune*, which are typical for non-intestinal ecosystems, in this particular case drinking water (Kalmbach *et al.*, 1999). Thirteen phylotypes

corresponded to organisms recovered from the human oral cavity, and were mainly recovered in studies of microbiota in the upper gastrointestinal tract, rather than faecal samples.

Cultivation-independent studies of gastrointestinal microbiota have revealed the presence of not-yet described phylotypes, of which some are, based on the SSU rRNA gene sequence, only distantly related to any cultured microbe with known SSU rRNA gene sequence. Those phylotypes form novel clusters that are well separated from cultivated species. In the phylogenetic tree of the human gastrointestinal microbiota those clusters are marked as Uncultured Clostridiales I and II (Fig. 1). Both clusters together comprise 63 distinct phylotypes, which will be, once they are cultured, organized in a number of families and genera, as indicated by their SSU rRNA gene similarity. Another cluster of 14 phylotypes within the Firmicutes phylum could not be classified even at the level of order (Table 2). Finally, 19 uncultured phylotypes within the phylum Bacteroidetes form another distinct cluster of exclusively uncultured organisms (Table 2).

Cultivation-based and molecular studies provide a somewhat different description of the diversity of the human gastrointestinal tract microbiota, as indicated by the degree of the overlapping findings (Fig. 2). Remarkably, both approaches show unequivocally that Firmicutes are by far the most diverse group, although the community structure of this group is highly underestimated in the reports of cultivation-dependent studies. Similarly, the Bacteroidetes community in human gastrointestinal samples is more diverse than assessed with cultivation-based approaches. However, about two-thirds of Bacteroidetes isolates, for which the SSU rRNA gene sequence is available, have also been detected by molecular approaches. This indicates that some of the widely distributed and abundant members of this phylum have already been isolated. Representatives of the Proteobacteria, and particularly subdivisions ϵ and γ , have so far been detected relatively poorly by molecular approaches. However, as sequencing of the SSU rRNA gene amplicons from an intestinal sample generally describes only the dominant fractions of the microbiota (Zoetendal *et al.*, 2006), and it is known that Proteobacteria have low abundance in the human gastrointestinal tract (Hopkins *et al.*, 2001), such low recovery of this phylum had to be expected.

Molecular studies showed that members of the phylum Cyanobacteria (one phylotype) can be found in the human gastrointestinal samples (Eckburg *et al.*, 2005), even though isolates from this group have not been reported. Furthermore, six phylotypes of the subphylum of α -Proteobacteria were detected by SSU rRNA gene sequencing, while there is only one known

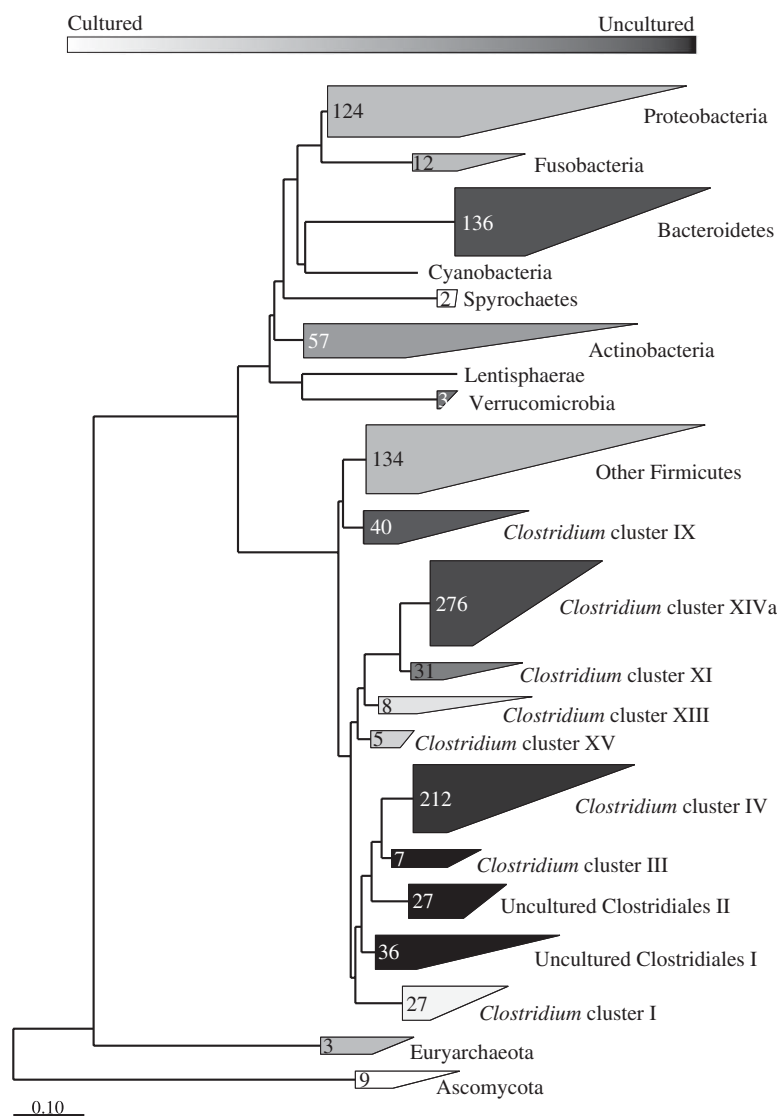


Fig. 1. SSU rRNA-based phylogenetic tree of the distinct phylotypes that have been found in the human gastrointestinal tract. The relative proportion of phylotypes that correspond to cultured representatives is indicated by different darkness of the filling. Black fills indicate phylotypes detected in cultivation-independent studies, while white indicates species detected in cultivation-based studies. The reference bar indicates 10% sequence divergence. Numbers of distinct phylotypes are given for each phylogenetic group.

α -proteobacterial gastrointestinal isolate – *Gemmiger formicilis* (Holdeman *et al.*, 1976; Benno *et al.*, 1986; Moore and Moore, 1995; Macfarlane *et al.*, 2004). Unfortunately, no 16S rRNA gene sequence is available for this species. Phylogenetic analysis revealed that five of the six α -proteobacterial phylotypes did not cluster within the order Rhizobiales to which *Gemmiger* belongs. Strikingly, all seven cyano- and α -proteobacterial phylotypes were detected in studies where mucosa samples from the upper gastrointestinal tract were analysed (Eckburg *et al.*, 2005; Wang *et al.*, 2005). It was previously shown that the microbiota associated with rectal mucosa may differ from that present in faeces (Zoetendal *et al.*, 2002), and recently, two studies indicated that mucosa at upper gastrointestinal parts contains specific phylotypes that are distantly related to any known gastrointestinal isolate (Eckburg *et al.*, 2005; Wang *et al.*, 2005). These data reinforce the notion that faecal samples do not compre-

hensively represent the microbiota that can be found in the lumen or attached to mucosal surfaces along the gastrointestinal tract, and that sampling along the gastrointestinal tract, both for cultivation-based analysis and molecular profiling, would enable better description of the microbial diversity of this ecosystem.

As previously mentioned, the diverse community of bacteria in the human gastrointestinal tract is accompanied by representatives of archaea and eukarya. According to our current knowledge, the archaeal community of the human intestine is very simple and consists of only three isolates, of which *Methanosphaera stadtmanae* is rare (Miller and Wolin, 1985), and *Methanobrevibacter ruminantium* was reported in a single study dating from 1968 (Nottingham and Hungate, 1968). The presence of only one archaeon – the abundant *Methanobrevibacter smithii* – was confirmed by molecular studies (Miller and Wolin, 1986; Eckburg *et al.*, 2005).

Table 2. SSU rRNA gene sequence based phylogenetic distribution of the human gastrointestinal prokaryotic phylotypes given for taxonomic levels of phylum, class, order and family or cluster as proposed by Collins and colleagues (1994).

Phylum		Class		Order		Family/Cluster					
Actinobacteria	57	Actinobacteria	57	Actinomycetales	22	Actinomycetaceae	6				
						Corynebacteriaceae	6				
						Micrococcaceae	4				
						Propionibacteriaceae	6				
						Bifidobacteriales	20				
Bacteroidetes	136	Bacteroidetes	136	Bacteroidales	136	Bifidobacterium	20				
						Coriobacteriaceae	15				
						Rikenellaceae	15				
						Bacteroidaceae	45				
						Prevotellaceae	36				
Cyanobacteria	1	Cyanobacteria	1	Chroococcales	1	Porphyromonadaceae	21				
						Unclassified	19				
Firmicutes	818	Asteroleplasma	1	Anaeroplasmatales	1	Unclassified	1				
						Bacilli	92	Bacillales	18	Bacillaceae	14
										Staphylococcaceae	4
		Lactobacillales	74								
		Clostridia	669	Clostridiales	669	Clostridiales	669	Aerococcaceae	1		
								Carnobacteriaceae	1		
								Lactobacillaceae	36		
								Leuconostocaceae	2		
								Lactococcaceae	2		
								Streptococcaceae	32		
								Cl. cluster I	27		
								Cl. cluster III	7		
								Cl. cluster IV	212		
								Cl. cluster IX	40		
								Cl. cluster XI	31		
Cl. cluster XIII	8										
Cl. cluster XIVa	276										
Cl. cluster XV	5										
Mollicutes	42	Unclassified	42	Unclassified	42	Unclassified	63				
						Cl. cluster XVI	14				
						Cl. cluster XVII	5				
Unclassified	14	Unclassified	14	Unclassified	14	Cl. cluster XVIII	9				
						Unclassified	14				
						Fusobacteria	12				
Fusobacteria	12	Fusobacteria	12	Fusobacteriales	12	Fusobacteriaceae	11				
						Incertae sedis 11	1				
Proteobacteria	124	<i>Alphaproteobacteria</i>	6	Rhizobiales	1	Unclassified	1				
				Sphingomonadales	1	Unclassified	1				
				Unclassified	4	Unclassified	4				
				<i>Betaproteobacteria</i>	19	Burkholderiales	18	Alcaligenaceae	8		
		Oxalobacteriaceae	3								
		Burkholderiaceae	1								
		<i>Gammaproteobacteria</i>	74	Neisseriales	4	Aeromonadales	4	Incertae sedis 5	1		
								Unclassified	5		
								Neisseriaceae	1		
								Aeromonadaceae	2		
								Succinivibrionaceae	2		
								Enterobacteriales	51	Enterobacteraceae	51
								Pasteurellales	2	Pasteurellaceae	2
		Pseudomonadales	11	Moraxellaceae	6	Pseudomonadaceae	5	Vibrionaceae	4		
								Xanthomonadales	2	Xanthomonadaceae	2
<i>Deltaproteobacteria</i>	7							Desulfovibrionales	7	Desulfovibrionaceae	7
		Helicobacteraceae	6								
Verrucomicrobia	3	Verrucomicrobia	3	Verrucomicrobiales	3	Verrucomicrobiaceae	3				
Euryarchaeota	3	Methanobacteria	3	Methanobacteriales	3	Methanobacteriaceae	3				

Inventory period from 1998 until the beginning of 2006.

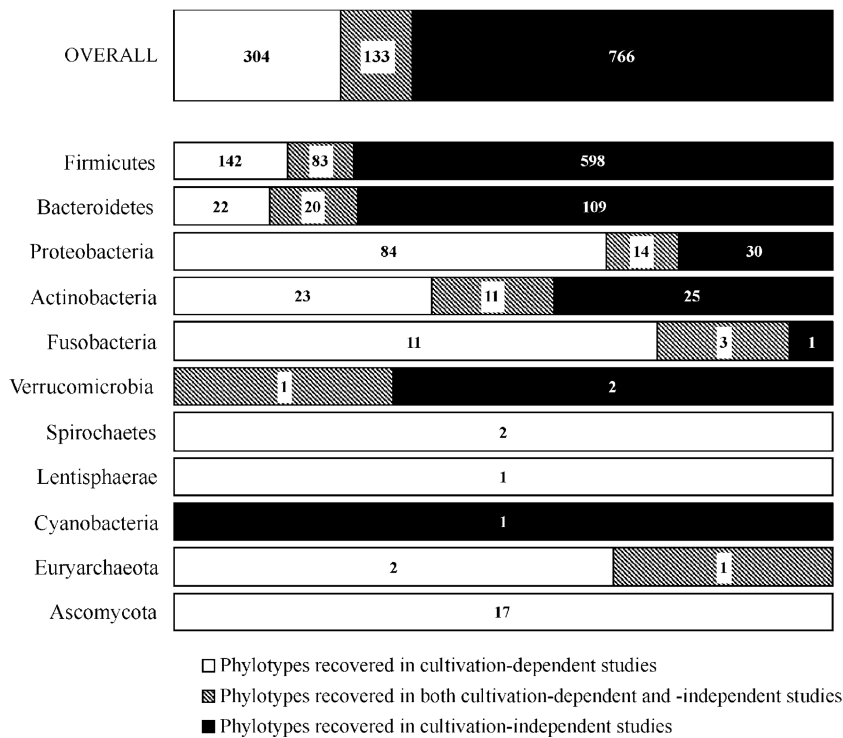


Fig. 2. Distribution of gastrointestinal prokaryotic phylotypes over type of the study, given overall and for eight bacterial, one archaeal and one eukaryal phylum. The diversity of the gastrointestinal isolates that have been fully characterized but lack the SSU rRNA gene sequence were taken into account for construction of this figure.

The diversity of eukarya in the human intestine was assessed exclusively by cultivation-dependent approaches. Seventeen *Candida*, *Aspergillus* and *Penicillium* species were isolated from human intestinal samples (Anderson, 1917; Finegold *et al.*, 1974; 1977; Taylor *et al.*, 1985; Biasoli *et al.*, 2002). However, except for *Candida albicans* and *C. rugosa*, eukarya are neither widely distri-

buted nor abundant in the human intestine, according to the results of the cultivation-based approaches.

Species richness estimation

The recent studies of the human gastrointestinal microbiota diversity by sequencing have provided an exploding

Table 3. Bacterial species typically isolated from other ecosystems but the human gastrointestinal tract that were detected in the human gastrointestinal samples, based on SSU rRNA gene sequence analysis.

Species	Known ecological niche (Reference)	Reference of study in which SSU rRNA sequence was detected
<i>Acinetobacter lwoffii</i>	Oropharynx (Rathinavelu <i>et al.</i> , 2003)	(Wang <i>et al.</i> , 2005)
<i>Actinomyces graevenitzi</i>	Saliva, bronchia (Ramos <i>et al.</i> , 1997)	(Eckburg <i>et al.</i> , 2005)
<i>Actinomyces odontolyticus</i>	Tooth root canal (Peters <i>et al.</i> , 2002)	(Eckburg <i>et al.</i> , 2005)
<i>Allisonella histaminiformans</i>	Cecum of horse (Garner <i>et al.</i> , 2002)	(Hayashi <i>et al.</i> , 2002a; Eckburg <i>et al.</i> , 2005)
<i>Anaerococcus vaginalis</i>	Vaginal discharges (Ezaki <i>et al.</i> , 2001)	(Mangin <i>et al.</i> , 2004; Eckburg <i>et al.</i> , 2005)
<i>Aquabacterium commune</i>	Drinking water (Kalmbach <i>et al.</i> , 1999)	(Wang <i>et al.</i> , 2003)
<i>Bifidobacterium ruminantium</i>	Bovine rumen (Biavati and Mattarelli, 1991)	(Bonnet, 2002)
<i>Burkholderia cepacia</i>	Saliva (Sajjan <i>et al.</i> , 1992)	(Wang <i>et al.</i> , 2003)
<i>Corynebacterium durum</i>	Subgingival plaque (Barrett <i>et al.</i> , 2001)	(Eckburg <i>et al.</i> , 2005)
<i>Corynebacterium sundsvallense</i>	Sinus (Collins <i>et al.</i> , 1999)	(Eckburg <i>et al.</i> , 2005)
<i>Dialister invisus</i>	Oral cavity (Downes <i>et al.</i> , 2003)	(Bonnet, 2002; Mangin <i>et al.</i> , 2004; Eckburg <i>et al.</i> , 2005; Wang <i>et al.</i> , 2005)
<i>Fusobacterium peridonticum</i>	Gingival crevice (Gmur <i>et al.</i> , 2006)	(Wang <i>et al.</i> , 2005)
<i>Granulicatella adiacens</i>	Throat flora (Collins and Lawson, 2000)	(K. Saunier, unpublished)
<i>Leuconostoc argentinum</i>	Raw milk (Dicks <i>et al.</i> , 1993)	(Heilig <i>et al.</i> , 2002)
<i>Mogibacterium vesicum</i>	Peridontal pocket (Nakazawa <i>et al.</i> , 2000)	(Wang <i>et al.</i> , 2005)
<i>Neisseria mucosa</i>	Subgingival plaque (Haffajee <i>et al.</i> , 2005)	(Wang <i>et al.</i> , 2005)
<i>Phascolarctobacterium faecium</i>	Koala faeces (Del Dot <i>et al.</i> , 1996)	(Hayashi <i>et al.</i> , 2002b, 2003)
<i>Prevotella shahii</i>	Saliva (Sakamoto <i>et al.</i> , 2004)	(Wang <i>et al.</i> , 2005)
<i>Rothia mucilaginosa</i>	Oral cavity (Ruoff, 2002)	(Wang <i>et al.</i> , 2005)
<i>Stenotrophomonas maltophilia</i>	Oral cavity (Tada <i>et al.</i> , 2004)	(Wang <i>et al.</i> , 2003)

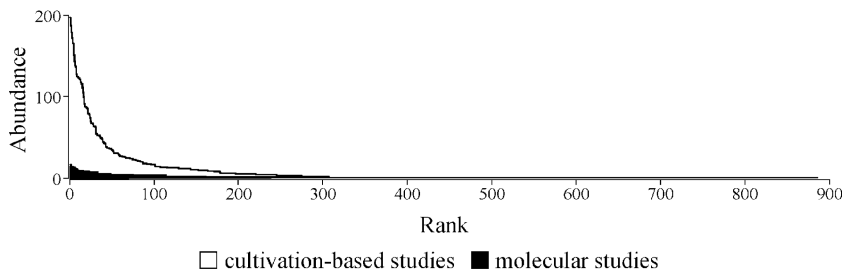


Fig. 3. Rank–abundance curve for the phylotypes of the human gastrointestinal tract found in cultivation-based and molecular studies. The abundance of the phylotype is defined as number of individuals that have been reported to harbour the particular phylotype.

amount of SSU rRNA gene sequence information. However, due to the fact that only a limited number of individuals has been subjected to the analysis of the human gastrointestinal microbiota and because each subject harbours a unique microbial community, a good representation of the ecosystem has not yet been achieved. Consequently, the rank–abundance curve, which can visualize how well an ecosystem has been sampled (Hughes *et al.*, 2001), has a very different shape when compared with the one obtained with data from cultivation-based studies (Fig. 3). The curve is characterized by very strong tailing, as the majority of the phylotypes were found in intestinal samples of only one individual. Still, it is remarkable that molecular studies performed to date revealed almost 900 distinct phylotypes based on the microbiota of only 50 individuals, while cultivation studies, which were employed for characterization of the microbiota of hundreds of subjects have recovered less than half of this species richness. Based on Good's coverage estimates (Moore and Holdeman, 1974), our current knowledge of the normal human gastrointestinal microbiota diversity overall is about 30%, which contrasts with the good coverage of the microbiota diversity that was achieved in individual studies (Suau *et al.*, 1999; Eckburg *et al.*, 2005).

Because a good representation of the intestinal diversity has not yet been achieved, it is not possible to come to reliable number of the total species richness on the basis of the present data set (Hughes *et al.*, 2001). Nevertheless, for the purpose of this review, the Chao1 index was calculated to provide its estimate. Because the archaeal and eukaryal community in the human intestine are rather simple, only the frequency of bacterial phylotypes was taken into account for the calculation. Three values are needed for estimating the species richness based on the Chao1 index: the total number of phylotypes, and the number of the rare phylotypes – these which were detected in only one or two individuals. Based on data from molecular studies that yielded a total of 898 bacterial phylotypes, of which 622 were found in only one individual while 121 in two, the expected diversity of the human gastrointestinal microbiota amounts to about 2500 phylotypes. This estimate exceeds the estimate based on cultivation-based techniques by one order of magnitude

but the true diversity of the human gastrointestinal tract might be even greater, because the Chao1 estimator tends to underestimate the true richness at low sample sizes (Hughes *et al.*, 2001).

Furthermore, the Chao1 estimate, calculated on the basis of frequency of bacterial phylotypes reported in both types of approaches, where out of 1183 phylotypes, 686 were found in one and 128 in two individuals, showed that expected diversity of the gastrointestinal tract exceeds 3000 phylotypes. However, this estimate, similar to the one calculated on the basis of molecular studies alone, probably underestimates the true diversity. Moreover, there are several limitations of the applied approach that originate from the data set that describes the diversity of the human gastrointestinal microbiota, which are discussed below.

Limitations of the approach

The diversity of the gastrointestinal microbiota was analysed on the bases of the SSU rRNA sequence. Unique phylotypes were defined as groups of sequences that have sequence similarity of 98% or less with any other SSU rRNA sequence derived from the human gastrointestinal tract. There is no general agreement about the most appropriate cut-off similarity for a phylotype, and different cut-off values have been applied even in the most recent studies (Eckburg *et al.*, 2005; Gill *et al.*, 2006). This has a major effect on the obtained measure of diversity. As previously reported, within the same clone library, the number of distinct phylotypes ranges from 148 to 643 when the cut-off value is changed from 97% to 99% (Venter *et al.*, 2004). The sequence similarity of different SSU rRNA gene copies in the same genome is rarely below 99% (Acinas *et al.*, 2004), although it can vary substantially (Wang *et al.*, 1997). This variation increases between different strains of the same species but rarely reaches values of 97% (Sacchi *et al.*, 2005). Hence, in this review a cut-off value of 98% was chosen, which has been widely applied by others (Suau *et al.*, 1999; Hayashi *et al.*, 2002a) and represents a compromise between three commonly used values. However, it is well known that many species, especially among Enterobacteriaceae, have more than 98% SSU rRNA gene sequence similarity

with other related species. In fact, from 382 distinct gastrointestinal species for which the SSU rRNA gene is available, only 305 would be recognized as unique phylotypes based on their SSU rRNA gene sequence, when using the 98% similarity criterion. Thus, it is likely that some of the identified distinct phylotypes are, in fact, groups of different phenotypes. Until cultured representatives of these phylotypes are obtained, and while SSU rRNA gene sequence remains the only available information, absolute accuracy will not be achieved.

Another factor affecting the accuracy of our analyses originates from the fact that many of the available SSU rRNA gene sequences are partial, and unless they correspond to the same region within the gene, they do not allow proper direct comparison. Only if two sequences match a common, fully sequenced 'type phylotype', ARB software allow grouping of partial sequences into a single distinct phylotype. Otherwise the sequences are recognized as distinct phylotypes even if they correspond to a different region of the very same SSU rRNA gene. Hence, an overestimation of the true diversity might occur. It would be convenient if there were a general agreement that, if partial sequences are provided, they are always obtained from the same variable region of the SSU rRNA gene. As Ludwig and Klenk (2001) indicated, the sequence of the V1 region appears to be the most variable and consequently the most informative. Therefore, selection of this region seems the most appropriate. However, it has to be acknowledged that sequences generated for analysis by fingerprinting approaches, such as denaturing gradient gel electrophoresis (DGGE) or amplicons obtained by group-specific polymerase chain reaction (PCR) amplification are often confined to different specific regions of the SSU rRNA gene, due to restricted options for primer design.

Each of the techniques used for the assessment of the diversity of the human gastrointestinal microbiota has its limitations and induces biases in the obtained portrayal of the microbial diversity. The limitations of the cultivation-based approaches to adequately resemble the diversity of the gastrointestinal microbiota are nowadays well acknowledged (Tannock, 1999; Vaughan *et al.*, 2000). Moreover, the SSU rRNA gene sequencing-based analyses of the microbial diversity are also giving biased picture, and among many factors that influence the obtained information, effectiveness of cell lyses, and PCR biases are highly relevant. Although it is generally considered that SSU rRNA gene sequence assessed diversity of an ecosystem is representing the most dominant part of the microbial community, this is necessarily accurate as the sequence of the used PCR primers also induces biases. The reason for this effect is that, even when the total bacterial microbiota is targeted, different PCR

primers are used, and primer sets can have preferences for different phylogenetic groups (Suau *et al.*, 1999; Eckburg *et al.*, 2005). This is because none of the universal primer sets has the claimed specificity or universality (Horz *et al.*, 2005). This effect, together with the problems related to the cell lyses, have led to the underestimation of the members of Actinobacterium phylum within the human gastrointestinal microbiota, which appear to be quantitatively relevant members of this ecosystem, based on the results of enumerative techniques such as florescent *in situ* hybridization (Lay *et al.*, 2005). Finally, cloning and sequencing procedures are in most cases performed only once. Remarkably, the analysis of microbiota of a single individual by the use of different numbers of PCR cycles showed some differences in the recovered clone libraries (Suau *et al.*, 1999; Bonnet *et al.*, 2002). This might be an exclusive effect of the number of PCR cycles, as suggested by authors, but it might also indicate limitations of the cloning and sequencing technique itself. If so, the microbiota of each individual might be even more complex than revealed in a single cloning and sequencing experiment.

Concluding remarks

The human gastrointestinal microbiota is an important element of the human body, which is receiving increasing attention. Based only on the data which was generated until the beginning of the 2006, the human gastrointestinal tract harbours over 1200 distinct microorganisms. The complete coverage of the diversity has not yet been achieved, but the total number of already reported distinct phylotypes exceeds the most ambitious estimates previously given for this ecosystem. However, so far the majority of the phylotypes have been detected only in a single individual, indicating that many more gastrointestinal inhabitants are to be discovered, and the total diversity will probably be measured in thousands of species. Therefore, the human gastrointestinal microbiota, often referred to as forgotten organ of the human body (O'Hara and Shanahan, 2006), is still insufficiently described even at the composition level. Because reliable insight into the gastrointestinal diversity is essential for providing a reference framework to study its dynamics in time and space, analyse its functions and characterize host-microbe interactions, more attention should be given to the simple description of this ecosystem. To this end, the availability of curated databases of rRNA sequences, allowing for adequate organization of redundant data without losing useful information from all previously deposited and future sequence entries such as source of isolation, is an absolute requirement. The inventory of currently available data described here is the first important step towards this goal.

Acknowledgements

The authors would like to thank Dr Colin Ingham for reviewing this manuscript and useful suggestions concerning phrasing and sentence structure.

References

- Acinas, S.G., Marcelino, L.A., Klepac-Ceraj, V., and Polz, M.F. (2004) Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J Bacteriol* **186**: 2629–2635.
- Altschul, S.F., Gish, W., Miller, W., Meyers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403.
- Anderson, H.W. (1917) Yeast-like fungi of the human intestinal tract. *J Infect Dis* **21**: 341–386.
- Anonymous (1981) The pioneers of pediatric medicine: Teodor Escherich. *Eur J Pediatr* **137**: 131.
- Bäckhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and Gordon, J.I. (2005) Host-bacterial mutualism in the human intestine. *Science* **307**: 1915–1920.
- Barcenilla, A., Pryde, S.E., Martin, J.C., Duncan, S.H., Stewart, C.S., Henderson, C., and Flint, H.J. (2000) Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl Environ Microbiol* **66**: 1654–1661.
- Barrett, S.L.R., Cookson, B.T., Carlson, L.C., Bernard, K.A., and Coyle, M.B. (2001) Diversity within reference strains of *Corynebacterium matruchotii* includes *Corynebacterium durum* and a novel organism. *J Clin Microbiol* **39**: 943–948.
- Benno, Y., Sawada, K., and Mitsuoka, T. (1986) Comparison of the fecal microflora in rural Japanese and urban Canadians. *Microbiol Immunol* **30**: 521–532.
- Biasoli, S.M., Tosello, E.M., and Magaró, M.H. (2002) Adherence of *Candida* strains isolated from the human gastrointestinal tract. *Mycoses* **45**: 465–469.
- Biavati, B., and Mattarelli, P. (1991) *Bifidobacterium ruminantium* sp. nov. and *Bifidobacterium merycicum* sp. nov. from the rumens of cattle. *Int J Syst Bacteriol* **41**: 163–168.
- Bik, E.M., Eckburg, P.B., Gill, S.R., Nelson, K.E., Purdom, E.A., Francois, F., et al. (2006) Molecular analysis of the bacterial microbiota in the human stomach. *Proc Natl Acad Sci* **103**: 732–737.
- Bonnet, R. (2002) Phylogenetic diversity of the human gut microflora based on small-subunit rRNA. *PhD Thesis*. Reading, UK: University of Reading.
- Bonnet, R., Suau, A., Doré, J., Gibson, G.R., and Collins, M.D. (2002) Differences in rDNA libraries of faecal bacteria derived from 10- and 25-cycle PCRs. *Int J Syst Evol Microbiol* **52**: 757–763.
- Breitbart, M., Hewson, I., Felts, B., Mahaffy, J.M., Nulton, J., Salamon, P., and Rohwer, F. (2003) Metagenomic analyses of an uncultured viral community from human feces. *J Bacteriol* **185**: 6220–6223.
- Cole, J.R., Chai, B., Marsh, T.L., Farris, R.J., Wang, Q., Kulam, S.A., et al. (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res* **31**: 442–443.
- Collins, M.D., Lawson, P.A. (2000) The genus *Abiotrophia* (Kawamura et al.) is not monophyletic: proposal of *Granulicatella* gen. nov., *Granulicatella adiacens* comb. nov., *Granulicatella elegans* comb. nov. and *Granulicatella balaeopterae* comb. nov. *Int J Syst Evol Microbiol* **50**: 365–369.
- Collins, M.D., Lawson, P.A., Willems, A., Cordoba, J.J., Fernandez-Garayzabal, J., Garcia, P. et al. (1994) The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol* **44**: 812–826.
- Collins, M.D., Bernard, K.A., Hutson, R.A., Sjoden, B., Nyberg, A., and Falsen, E. (1999) *Corynebacterium sundsvallense* sp. nov., from human clinical specimens. *Int J Syst Bacteriol* **49**: 361–366.
- Del Dot, T., Osawa, R., and Stackebrandt, E. (1996) *Phascolarctobacterium faecium* gen. nov., spec. nov., a novel taxon of the *Sporomusa* group of bacteria. *Syst Appl Microbiol* **16**: 380–384.
- Derrien, M., Adawi, D., Ahrne, S., Jeppsson, B., Molin, G., Osman, N., et al. (2004) The intestinal mucosa as a habitat of the gut microbiota and a rational target for probiotic functionality and safety. *Microb Ecol Health Dis* **16**: 137–144.
- Dicks, L.M.T., Fantuzzi, L., Gonzalez, F.C., Du Toit, M., and Dellaglio, F. (1993) *Leuconostoc argentinum* sp. nov., isolated from Argentine raw milk. *Int J Syst Bacteriol* **43**: 347–351.
- Downes, J., Munson, M., and Wade, W.G. (2003) *Dialister invisus* sp. nov., isolated from the human oral cavity. *Int J Syst Evol Microbiol* **53**: 1937–1940.
- Drasar, B.S., and Barrow, P.A. (1985) *Intestinal Microbiology*. Washington, DC, USA: American Society for Microbiology.
- Duncan, S.H., Louis, P., and Flint, H.J. (2004) Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Appl Environ Microbiol* **70**: 5810–5817.
- Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., et al. (2005) Diversity of the human intestinal microbial flora. *Science* **308**: 1635–1638.
- Ezaki, T., Kawamura, Y., Li, N., Li, Z.Y., Zhao, L., and Shu, S. (2001) Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov. for members of the genus *Peptostreptococcus*. *Int J Syst Evol Microbiol* **51**: 1521–1528.
- Favier, C.F., Vaughan, E.E., de Vos, W.M., and Akkermans, A.D.L. (2002) Molecular monitoring of succession of bacterial communities in human neonates. *Appl Environ Microbiol* **68**: 219–226.
- Finegold, S.M., Howard, R.A., and Vera, L.S. (1974) Effect of diet on human intestinal fecal flora: comparison of Japanese and American diets. *Am J Clin Nutr* **27**: 1456–1469.
- Finegold, S.M., Vera, L.S., Paul, T.S., Harvey, A.E., Shirley, M.L., and Ronald, L.P. (1977) Fecal microbial flora in Seventh Day Adventist populations and control subjects. *Am J Clin Nutr* **30**: 1781–1792.
- Furrie, E. (2006) A molecular revolution in the study of intestinal microflora. *Gut* **55**: 141–143.
- Garner, M.R., Flint, J.F., and Russell, J.B. (2002) *Allisonella histaminiformans* gen. nov., sp. nov. a novel bacterium that produces histamine, utilizes histidine as its sole energy source, and could play a role in bovine and equine laminitis. *Syst Appl Microbiol* **25**: 298–506.

- Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., *et al.* (2006) Metagenomic analysis of the human distal gut microbiome. *Science* **312**: 1355–1359.
- Gmur, R., Munson, M.A., and Wade, W.G. (2006) Genotypic and phenotypic characterization of fusobacteria from Chinese and European patients with inflammatory periodontal diseases. *Syst Appl Microbiol* **29**: 120.
- Haffajee, A.D., Japlit, M., Bogren, A., Kent, R.L.J., Goodson, J.M., and Socransky, S.S. (2005) Differences in the subgingival microbiota of Swedish and USA subjects who were periodontally healthy or exhibited minimal periodontal disease. *J Clin Periodontol* **32**: 33–39.
- Hayashi, H., Sakamoto, M., and Benno, Y. (2002a) Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods. *Microbiol Immunol* **46**: 535–548.
- Hayashi, H., Sakamoto, M., and Benno, Y. (2002b) Fecal microbial diversity in a strict vegetarian as determined by molecular analysis and cultivation. *Microbiol Immunol* **46**: 819–831.
- Hayashi, H., Sakamoto, M., and Benno, Y. (2003) Molecular analysis of fecal microbiota in elderly individuals using 16S rDNA library and T-RFLP. *Microbiol Immunol* **47**: 557–570.
- Heilig, H.G., Zoetendal, E.G., Vaughan, E.E., Marteau, P., Akkermans, A.D.L., and de Vos, W.M. (2002) Molecular diversity of *Lactobacillus* spp. and other lactic acid bacteria in the human intestine as determined by specific amplification of 16S ribosomal DNA. *Appl Environ Microbiol* **68**: 114–123.
- Hold, G.L., Pryde, S.E., Russell, V.J., Furrie, E., and Flint, H.J. (2002) Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. *FEMS Microbiol Ecol* **39**: 33–39.
- Holdeman, L.V., Good, I.J., and Moore, W.E.C. (1976) Human fecal flora: variation in bacterial composition within individuals and a possible effect of emotional stress. *Appl Environ Microbiol* **31**: 359–375.
- Holmstrom, K., Collins, M.D., Moller, T., Falsen, E., and Lawson, P.A. (2004) *Subdoligranulum variabile* gen. nov., sp. nov. from human feces. *Anaerobe* **10**: 197–203.
- Hopkins, M.J., Sharp, R., and Macfarlane, G.T. (2001) Age and disease related changes in intestinal bacterial populations assessed by cell culture, 16S rRNA abundance, and community cellular fatty acid profiles. *Gut* **48**: 198–205.
- Horz, H.P., Vianna, M.E., Gomes, B.P.F.A., and Conrads, G. (2005) Evaluation of universal probes and primer sets for assessing total bacterial load in clinical samples: general implications and practical use in endodontic antimicrobial therapy. *J Clin Microbiol* **43**: 5332–5337.
- Hughes, J.B., Hellmann, J.J., Ricketts, T.H., and Bohannon, B.J.M. (2001) Counting the uncountable: statistical approaches to estimating microbial diversity. *Appl Environ Microbiol* **67**: 4399–4406.
- Kalmbach, S., Manz, W., Wecke, J., and Szewzyk, U. (1999) *Aquabacterium* gen. nov., with description of *Aquabacterium citratiphilum* sp. nov., *Aquabacterium parvum* sp. nov. and *Aquabacterium commune* sp. nov., three *in situ* dominant bacterial species from the Berlin drinking water system. *Int J Syst Bacteriol* **49**: 769–777.
- Kitahara, M., Sakamoto, M., Ike, M., Sakata, S., and Benno, Y. (2005) *Bacteroides plebeius* sp. nov. and *Bacteroides coprocola* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* **55**: 2143–2147.
- Kleessen, B. (2000) Culture-based knowledge on biodiversity, development and stability of human gastrointestinal microflora. *Microb Ecol Health Dis* **12**: 53–63.
- Lay, C., Rigottier-Gois, L., Holmström, K., Rajilic, M., Vaughan, E.E., de Vos, W.M., *et al.* (2005) Colonic microbiota signatures across five northern European countries. *Appl Environ Microbiol* **71**: 4153–4155.
- Ley, R.E., Turnbaugh, P.J., Klein, S., and Gordon, J.I. (2006) Microbial ecology: human gut microbes associated with obesity. *Nature* **444**: 1022–1023.
- Ludwig, W., and Klenk, H.P. (2001) Overview: a phylogenetic backbone and taxonomic framework for procaryotic systematics. In *Bergey's Manual of Systematic Bacteriology*. Garrity, G.M. (ed.). New York, USA: Springer, pp. 49–66.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, *et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- Macfarlane, S., Furrie, E., Cummings, J.H., and Macfarlane, G.T. (2004) Chemotaxonomic analysis of bacterial populations colonizing the rectal mucosa in patients with ulcerative colitis. *Clin Infect Dis* **38**: 1690–1699.
- Mangin, I., Bonnet, R., Seksik, P., Rigottier-Gois, L., Sutren, M., Bouhnik, Y., *et al.* (2004) Molecular inventory of faecal microflora in patients with Crohn's disease. *FEMS Microbiol Ecol* **50**: 25–36.
- Miller, T.L., and Wolin, M.J. (1985) *Methanosphaera stadtmaniae* gen nov., sp. nov. a species that forms methane by reducing methanol with hydrogen. *Arch Microbiol* **141**: 116–122.
- Miller, T.L., and Wolin, M.J. (1986) Methanogens in human and animal tracts. *Syst Appl Microbiol* **7**: 223–229.
- Moore, W.E.C., and Holdeman, L.V. (1974) Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Appl Microbiol* **27**: 961–979.
- Moore, W.E.C., and Moore, L.H. (1995) Intestinal floras of populations that have a high risk of colon cancer. *Appl Environ Microbiol* **61**: 3202–3207.
- Nakazawa, F., Sato, M., Poco, S.E., Hashimura, T., Ikeda, T., Kalfas, S., *et al.* (2000) Description of *Mogibacterium pumilum* gen. nov., sp. nov. and *Mogibacterium vesicum* gen. nov., sp. nov. and reclassification of *Eubacterium timidum* (Holdeman *et al.* 1980) as *Mogibacterium timidum* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **50**: 679–688.
- Nicholson, J.K., Holmes, E., and Wilson, I.D. (2005) Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Microbiol* **3**: 431–438.
- Nottingham, P.M., and Hungate, R.E. (1968) Isolation of methanogenic bacteria from feces of man. *J Bacteriol* **96**: 2178–2179.
- Noverr, M.C., and Huffnagle, G.B. (2004) Does the microbiota regulate immune responses outside the gut? *Trends Microbiol* **12**: 562–568.
- O'Hara, A.M., and Shanahan, F. (2006) The gut flora as a forgotten organ. *EMBO Rep* **7**: 688–693.
- Peters, L.B., Wesselink, P.R., and van Winkelhoff, A.J. (2002) Combinations of bacterial species in endodontic infections. *Int Endod J* **35**: 698–702.
- Phillips, M.L. (2006) Interdomain interactions: dissecting animal–bacterial symbioses. *Bioscience* **56**: 376–381.

- Rajilić-Stojanović, M. (2007) Diversity of the human gastrointestinal microbiota: novel perspectives from high throughput analyses. *PhD Thesis*. Wageningen, the Netherlands: Wageningen University.
- Ramos, C.P., Falsen, E., Alvarez, N., Akervall, E., Sjoden, B., and Collins, M.D. (1997) *Actinomyces graevenitzi* sp. nov., isolated from human clinical specimens. *Int J Syst Bacteriol* **47**: 885–888.
- Rathinavelu, S., Zavros, Y., and Merchant, J.L. (2003) *Acinetobacter Iwoffii* infection and gastritis. *Microbes Infect* **5**: 651.
- Ruoff, K.L. (2002) Miscellaneous catalase-negative, Gram-positive cocci: emerging opportunists. *J Clin Microbiol* **40**: 1129–1133.
- Sacchi, C.T., Alber, D., Dull, P., Mothershed, E.A., Whitney, A.M., Barnett, G.A., et al. (2005) High level of sequence diversity in the 16S rRNA genes of *Haemophilus influenzae* isolates is useful for molecular subtyping. *J Clin Microbiol* **43**: 3734–3742.
- Sajjan, U.S., Corey, M., Karmali, M.A., and Forstner, J.F. (1992) Binding of *Pseudomonas cepacia* to normal human intestinal mucin and respiratory mucin from patients with cystic fibrosis. *J Clin Invest* **89**: 648–656.
- Sakamoto, M., Suzuki, M., Huang, Y., Umeda, M., Ishikawa, I., and Benno, Y. (2004) *Prevotella shahii* sp. nov. and *Prevotella salivae* sp. nov., isolated from the human oral cavity. *Int J Syst Evol Microbiol* **54**: 877–883.
- Satokari, R.M., Vaughan, E.E., Akkermans, A.D.L., Saarela, M., and de Vos, W.M. (2001) Bifidobacterial diversity in human feces detected by genus-specific PCR and denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **67**: 504–513.
- Suau, A., Bonnet, R., Sutren, M., Godon, J.-J., Gibson, G.R., Collins, M.D., and 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.
- Tada, A., Shiiba, M., Yokoe, H., Hanada, N., and Tanzawa, H. (2004) Relationship between oral motor dysfunction and oral bacteria in bedridden elderly. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* **98**: 184.
- Tannock, G.W. (1999) Analysis of the intestinal microflora: a renaissance. *Antonie Van Leeuwenhoek* **76**: 265–278.
- Taylor, G.R., Kropp, K.D., and Molina, T.C. (1985) Nine-year microflora study of an isolator-maintained immunodeficient child. *Appl Environ Microbiol* **50**: 1349–1356.
- Vaughan, E.E., Schut, F., Heilig, H.G., Zoetendal, E.G., de Vos, W.M., and Akkermans, A.D. (2000) A molecular view of the intestinal ecosystem. *Curr Issues Intest Microbiol* **1**: 1–12.
- Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.
- Wang, M., Ahrne, S., Antonsson, M., and Molin, G. (2004) T-RFLP combined with principal component analysis and 16S rRNA gene sequencing: an effective strategy for comparison of fecal microbiota in infants of different ages. *J Microbiol Methods* **59**: 53.
- Wang, M., Ahrne, S., Jeppsson, B., and Molin, G. (2005) Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. *FEMS Microbiol Ecol* **54**: 219.
- Wang, R.F., Beggs, M.L., Robertson, L.H., and Cerniglia, C.E. (2002) Design and evaluation of oligonucleotide-microarray method for the detection of human intestinal bacteria in fecal samples. *FEMS Microbiol Lett* **213**: 175–182.
- Wang, X., Heazlewood, S.P., Krause, D.O., and Florin, T.H.J. (2003) Molecular characterization of the microbial species that colonize human ileal and colonic mucosa by using 16S rDNA sequence analysis. *J Appl Microbiol* **95**: 508–520.
- Wang, Y., Zhang, Z., and Ramanan, N. (1997) The actinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. *J Bacteriol* **179**: 3270–3276.
- Zhang, T., Breitbart, M., Lee, W.H., Run, J.-Q., Wei, C.L., Soh, S.W.L., et al. (2006) RNA viral community in human feces: prevalence of plant pathogenic viruses. *PLoS Biol* **4**: 108–118.
- Zoetendal, E.G., Akkermans, A.D., and de Vos, W.M. (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.
- Zoetendal, E.G., von Wright, A., Vilpponen-Salmela, T., Ben-Amor, K., Akkermans, A.D.L., and de Vos, W.M. (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.
- Zoetendal, E.G., Vaughan, E.E., and de Vos, W.M. (2006) A microbial world within us. *Mol Microbiol* **59**: 1639–1650.

Supplementary material

The following supplementary material is available for this article online:

Fig. S1. SSU rRNA-based phylogenetic tree of the distinct phylotypes that have been found in the human gastrointestinal tract. Type phylotype representatives of monophyletic groups (> 98% sequence similarity) are presented in the phylogenetic tree while accession numbers for other members of monophyletic groups are listed between parentheses. The reference bar indicates 10% sequence divergence.

This material is available as part of the online article from <http://www.blackwell-synergy.com>