ORIGINAL ARTICLE

The Host Genotype Affects the Bacterial Community in the Human Gastrointestinal Tract

Erwin G. Zoetendal[1,2], Antoon D. L. Akkermans[1], Wilma M. Akkermans-van Vliet[1], J. Arjan G. M. de Visser[1] and Willem M. de Vos[1,2]

From the [1]Laboratory of Microbiology, Wageningen University, Hesselink van Suchtelenweg 4, 6703 CT, Wageningen, [2]Wageningen Center for Food Sciences, PO Box 557, 6700 AL, Wageningen, The Netherlands

Correspondence to: Erwin G. Zoetendal, Tel.: +31 317 483115; Fax: +31 317 483829; E-mail: erwin.zoetendal@algemeen.micr.wau.nl

Microbial Ecology in Health and Disease 2001; 13: 129–134

The gastrointestinal (GI) tract is one of the most complex ecosystems consisting of microbial and host cells. It is suggested that the host genotype, the physiology of the host and environmental factors affect the composition and function of the bacterial community in the intestine. However, the relative impact of these factors is unknown. In this study, we used a culture-independent approach to analyze the bacterial composition in the GI tract. Denaturing gradient gel electrophoresis (DGGE) profiles of fecal bacterial 16S rDNA amplicons from adult humans with varying degrees of genetic relatedness were compared by determining the similarity indices of the profiles compared. The similarity between fecal DGGE profiles of monozygotic twins were significantly higher than those for unrelated individuals ($t_s = 2.73$, $p_{1\text{-tail}} = 0.0063$, df = 21). In addition, a positive relationship ($F_{1,30} = 8.63$, $p = 0.0063$) between the similarity indices and the genetic relatedness of the hosts was observed. In contrast, fecal DGGE profiles of marital partners, which are living in the same environment and which have comparable feeding habits, showed low similarity which was not significantly different from that of unrelated individuals ($t_s = 1.03$, $p_{1\text{-tail}} = 0.1561$, df = 27). Our data indicate that factors related to the host genotype have an important effect on determining the bacterial composition in the GI tract. *Key words*: denaturing gradient gel electrophoresis, host genotytpe, human GI tract, 16S rDNA.

INTRODUCTION

While prokaryotes are the most abundant form of life on our planet (1), it is also well known that microbes are present in vast amounts in the animal gastrointestinal (GI) tract, where they often exceed the number of host cells (2). Therefore, they live in very close contact with each other and with the epithelial cells of the host. The succession of this complex microbial community starts after birth, when empty niches become colonized by the first invading fast-growing microbes. Successive shifts of different microbial populations finally result in a climax community (3, 4). Although most of the interactions in the GI tract are still unknown, several studies indicate that signaling between host and microbes is very important in this ecosystem (5, 6). Recently, the molecular details of the communication between *Bacteroides thetaiotaomicron* and its murine host have been elucidated (7–9). Furthermore, it has been reported for various animal systems that the presence of methanogens in the intestine is a phylogenetic character that obeys 'Dollo's rule', i.e. traits that are lost in the course of evolution do not appear in any of the descendants of the common ancestor that lost these traits (10, 11). On the other hand, a recent study with humans and rats has indicated that shared and unique environmental conditions are important in the ecology of methanogens (12). These rather contradicting observations indicate the complexity of the ecology of methanogens in the GI tract. It is at least remarkable that only a fraction of humans harbors significant numbers of intestinal methanogens (13). Moreover, it has been established that each individual harbors specific strains of *Helicobacter pylori* or *Bifidobacterium* and *Lactobacillus spp*. (14, 15). All these observations argue strongly for the genetic predisposition of the host determining the composition in the GI tract.

A serious problem that limits the global analysis of the whole GI tract community is the inability to isolate and characterize all microbes. Viable plate count techniques only reveal a minor fraction of the GI tract community (16–18). Hence, analysis of the contribution of the host in determining the microbial composition in the GI tract requires the application of culture-independent techniques. These are mainly based on the sequence variability of 16S rRNA genes that have shown to be useful phylogenetic markers (19–21). Recently, it has been shown that the dominant bacterial community in feces remains stable in time (22, 23). In addition, temperature or denaturing gradient gel electrophoresis (TGGE or DGGE) of the PCR-amplified sequences of fecal 16S rRNA and rDNA

 ISSN 0891-060X

indicated that the bacterial composition was host-specific (22).

To study whether factors related to the genetic predisposition of the host rather than environmental factors determine a host-specific composition in the GI tract, we analyzed fecal samples from human individuals with different degrees of genetic relatedness, varying from monozygotic twins to marital partners, and four non-human primate species. Although fecal samples do not reflect the bacterial composition in all parts of the GI tract, the majority of bacteria leave the GI tract via the fecal route and, therefore, an observed effect on the fecal composition reflects an GI tract-related effect. The V6 to V8 regions of the 16S rDNA were amplified using fecal DNA as template and the amplicons were analyzed by DGGE (24, 25). Similarity indices between the DGGE profiles were calculated and statistical analyses were performed.

MATERIALS AND METHODS

Fecal sample collection

Fecal samples from 50 adult human volunteers (21–56 years) with varying genetic relationships (from mono- and dizygotic twins to genetically unrelated individuals) were collected after defecation in sterile plastic bags or collection tubes, and processed or transported to the lab as fast as possible, or stored at -20°C until use (freezing did not affect the procedures). In addition, fresh fecal samples were collected from four other primates (gorilla, chimpanzee, macaque, and orangutan) and transported immediately from the zoo to the lab. Fecal samples from genetically related individuals were analyzed only if they were living at separate locations. Individuals older than 60 years were not included, since it has been suggested that the physiological conditions of aging people affects the microbial community (26). In addition, two fecal samples were taken in a 4-month period from each of four unrelated human adults in order to assess the temporal stability of the fecal community.

DNA isolation, PCR, and DGGE analysis

The DNA isolation from fecal samples was performed as described earlier (22). One microliter of the fecal DNA solution was 10 times diluted and subsequently used as template to amplify the V6–V8 regions of 16S rDNA using primers F-0968-GC and R-1401 (27). DGGE analysis of the amplicons was performed on 8% polyacrylamide gels (PAGE) containing a urea plus formamide gradient from 38 to 48% (100% denaturing solution contains 7 M urea and 40% (v/v) formamide). Electrophoresis was performed in $0.5 \times$ TAE at 85 V at 60°C for 16 h using the DCode or D GENE System apparatus (BioRad, Hercules, CA). After electrophoresis, gels were silver-stained according to the protocol of Sanguinetti and colleagues (28) with some minor modifications.

Calculating similarity indices and statistical analysis

DGGE gels were scanned at 400 dpi and the software of Molecular Analyst 1.12 (Biorad) was used for comparing the DGGE profiles. Similarity indices of the compared profiles were calculated from the densitometric curves of the scanned DGGE profiles by using the Pearsons product-moment correlation coefficient (29). This procedure was performed three times. In this way, comparisons were made between fecal samples originating from human individuals with different degrees of genetic relatedness, with different ages, living in similar or different environments, four non-human primate species. In addition, fecal samples taken in a 4-month period were compared to access the temporal stability of the host-specific bacterial community. To obtain independent comparisons for statistical analysis, a random set of comparisons was selected from each gel in which each DGGE profile occurred only once. Regression analysis and student's t-tests were performed for statistical analysis of the data.

RESULTS

Comparison between fecal samples from monozygotic twins, marital partners, unrelated individuals, and non-human primates

To study whether factors related to the genetic predisposition of the host rather than environmental factors determine a host-specific composition in the GI tract, fecal samples from human individuals with identical genetic relatedness (monozygotic twins), with similar environmental conditions (marital partners), and unrelated individuals were analyzed. Amplicons of the variable regions V6–V8 were analyzed by DGGE, resulting in complex profiles which represent the host-specific dominant bacterial communities in which each band in a profile represents at least one unique bacterial 16S rDNA sequence (Fig. 1). The DGGE profiles of fecal 16S rDNA amplicons from monozygotic twins from three different families showed higher similarity within twin couples than between twin couples (Table I). In addition, the similarity between the marital partners within family C was relatively low compared with the similarity within each twin couple. In general, the similarity indices of all monozygotic twins ($N = 6$) were significantly higher than those of genetically unrelated individuals ($t_s = 2.73$, $p_{1\text{-tail}} = 0.0063$, $\mathrm{df} = 21$), despite some occasional observations of high similarity indices for genetically unrelated individuals (Fig. 2.). On the other hand, a comparison between the fecal communities of the marital partners did not show significant higher similarities than those for unrelated persons ($t_s = 1.03$, $p_{1\text{-tail}} = 0.1561$, $\mathrm{df} = 27$) notwithstanding the fact that the partners lived in the same environment and had in general comparable feeding habits (based on questionnaires). This may indicate that factors related to the host genotype or to the sex difference of the individuals compared have a

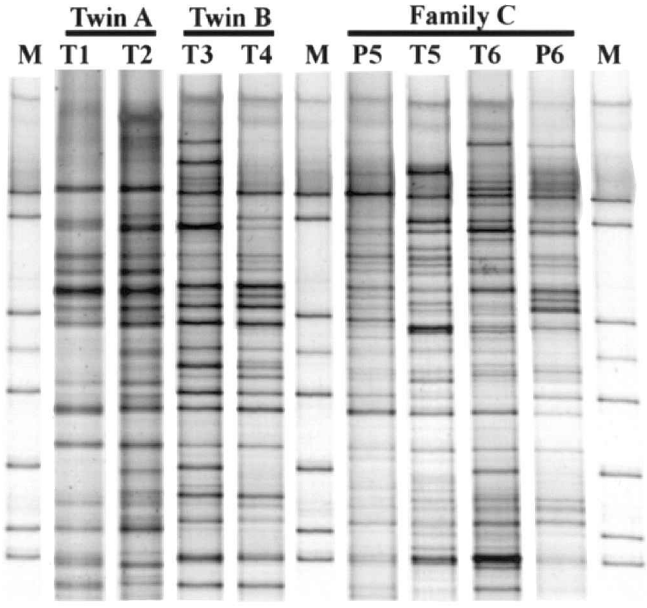

Fig. 1. DGGE profiles of the V6–V8 regions of 16S rDNA from fecal samples of monozygotic twins (T1 and T2, T3 and T4, T5 and T6) and marital partners (T5 and P5, T6 and P6). Twin A, Twin B, and Family C indicate the three different families. M indicates the marker for DGGE analysis.

significant influence on the bacterial community in the human GI tract. However, no significant difference was observed between the similarity indices of unrelated persons of the same sex and those of different sexes ($t_s = 0.41$, $p_{2\text{-tail}} = 0.69$, $df = 15$). This strongly argues for factors related to the host genotype to have an important effect on the GI tract composition. In addition, the similarity between the bacterial communities of genetically unrelated individuals was significantly higher than that of humans compared with other primates that we consider as the baseline level ($t_s = 3.99$, $p_{1\text{-tail}} = 0.0004$, $df = 19$).

DGGE analysis of feces from individuals with varying degrees of genetic relatedness

To study the impact of the genetic predisposition of adult humans in more detail, similarity indices were calculated for comparisons between DGGE profiles of fecal samples from individuals with varying degrees of genetic relatedness (Fig. 3). The similarity between the DGGE profiles of fecal samples appears to show a positive linear relationship with the genetic relatedness (r) between those individuals ($F_{1,30} = 8.63$, $p = 0.0063$), despite the high variation within the similarity indices of individuals with the same genetic relatedness. For example, the SD of the similarity indices for $r = 0$ (genetically unrelated) and $r = 0.5$ (brother and sisters, parents and children) are 14.7 and 16.4, respectively. These high variations could be caused by a variety of factors, such as age, diet, condition of the host, and experimental errors. The variation caused by replicate DNA isolation, PCR amplification, and DGGE analysis was found to be small and did not exceed a SD of 3.1.

The positive relationship between the similarity indices of fecal DGGE profiles and the genetic relatedness between the hosts compared indicates that factors related to the genetic predisposition of the host have a significant influence on the bacterial composition in feces. However, since the mean age difference between the individuals decreases with an increasing genetic relatedness (from $r =$ 0.25 to 1) the effect of the age difference on the similarity index was determined by comparing the DGGE profiles belonging to the group of brothers and sisters with $r = 0.5$ (including dizygotic twins). No positive or negative relation was observed for the similarity indices and the corresponding age differences (up to 14 years difference). This indicates that the age difference between the hosts did not have a significant effect on the observed differences between the hosts.

DISCUSSION

The GI tract is a complex ecosystems consisting of microbial and host cells. The microbial community plays an

Table I

Similarity matrix calculated from the DGGE profiles of Fig. 1. *T indicates a member of a monozygotic twin and P indicates the marital partner of the twin member*

	T1	T2	T3	T4	P5	T5	T6	P6
T1	100							
T2	**81.2**	100						
T3	36.1	44.9	100					
T4	47.4	45.1	**74.7**	100				
P5	13.4	11.8	18.9	23.3	100			
T5	30.0	35.5	58.3	59.4	**17.2**	100		
T6	32.4	34.5	62.1	45.5	16.7	**62.8**	100	
P6	50.2	50.5	60.8	75.7	19.0	52.1	**45.2**	100

[a] T1 and T2; T3 and T4; and T5, T6, P5, and P6 are members of respectively Twin A, Twin B, and Family C. Similarity indices of monozygotic twins are highlighted and underlined, similarity indices of marital partners are highlighted only.

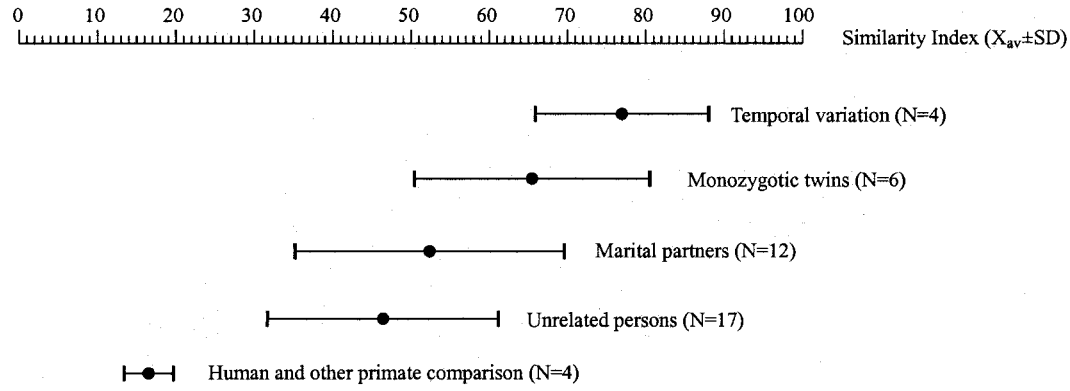

Fig. 2. Plot of the similarity indices from the human and primate, unrelated individuals, monozygotic twins, marital partners, and temporal variation comparisons. The mean (bullet) ± SD (black bar) are plotted.

important metabolic role in the GI tract by converting dietary components that escaped digestion by the host, and polymers excreted by the host into readily accessible nutrients and other compounds, such as vitamins (3). In addition, the microbial community participates in protecting the host against pathogens (30). Interactions between the host and the microbial community are, therefore, of considerable importance, very complex and just starting to be understood (5, 7–9). In the current study, we describe a culture-independent analysis of the dominant bacterial composition in the GI tract based on the 16S rRNA sequence variability of bacteria. DGGE and related electrophoretic analyses of 16S rDNA amplicons have shown to be powerful in studying the ecology of bacteria in different ecosystems including the GI tract (reviewed by (31, 32)). In an earlier study, we have shown that the diversity and stability of bacterial communities in the human GI tract could easily be analyzed using such an approach (22). In addition, the determination of similarity indices for DGGE profiles has shown to be a suitable tool to make the comparisons objectively (33, 34). In the current study, we determined the similarity indices of fecal DGGE profiles from hosts with varying degrees of genetic relatedness. It was observed that the host genotype has a significant effect on determining the dominant bacterial composition in the GI tract. The effect of the environment seems to be of less importance as indicated by the similarity indices of the bacterial communities between monozygotic twins and marital partners.

A strong positive correlation between the similarity indices and the genetic relatedness of the hosts was found, suggesting that either the host genotype or the colonization history, presumably via the fecal oral route from mother to child, has a significant effect on the bacterial composition in the GI tract. Another possible factor influencing the fecal comparison could be the age difference between the hosts compared. As reported earlier, the fecal communities might be affected by the physiological conditions of aging people (26). Therefore, individuals older than the arbitrary 60 years were excluded from this study. Since no significant relationship was found for the age differences between brothers and sisters with $r = 0.5$ and the corresponding similarity indices, it is not likely that our findings can be explained by the effect of age differences. Furthermore, the combination of earlier findings that the bacterial community in adults is stable in time (22, 23) and the fact that all genetic relatives in this study are already living separately for a long period of time (more than 5 years, some more than 15 years) also argues for a strong effect of the host genotype.

Our findings may explain why the fecal composition is host-specific and stable in time in adults, and the presence or absence of certain strains in the GI tract as described earlier (10, 11, 13–15, 22). Currently, we can only specu-

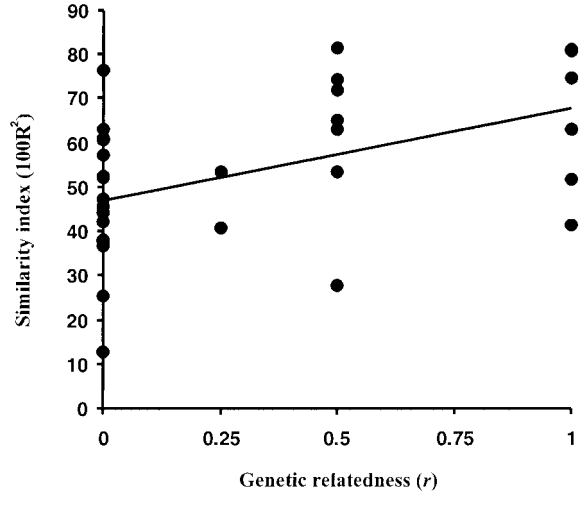

Fig. 3. Diagram showing the positive regression between the similarity indices of fecal DGGE profiles and the genetic relatedness between the individuals compared. The comparisons for each value of genetic relatedness and the linear regression that best fits the data are plotted. Comparisons were made for genetically unrelated individuals ($r = 0$); Aunts/uncles and nephews/nieces ($r = 0.25$); parents and children, brothers and sisters, dizygotic twins ($r = 0.5$), and monozygotic twins ($r = 1$).

late if immunological properties of the host, specific receptors for GI tract bacteria, or other communication systems between the host and the microbial community are responsible for the observed findings. Recently, the molecular details of the communication between *Bacteroides thetaiotaomicron* and its gnotobiotic murine host have been elucidated (7–9). Unfortunately, such studies can only be performed under well-controlled laboratory conditions and cannot be used to study the communication between humans and the GI tract community. Nevertheless, it is evident that the composition of the GI tract community is not only affected by the colonization history, the physiological (aging) effects in the GI tract and environmental factors, but also by the host genotype. Therefore, the effects of the consumption of for example probiotics or prebiotics may be difficult to determine in different individuals as a consequence of the variation in the bacterial community in the GI tract. To correct for these host-specific effects, the volunteers in a feeding trial should include individuals, which are genetically closely related.

ACKNOWLEDGEMENTS

We thank P. Knossenburg for assistance in standardizing the application of the Molecular Analyst software to analyze DGGE profiles, G.H.J. Heilig for technical assistance, Dr G. Welling and Dr H. Harmsen for discussions, and Dr G.W. Tannock and Dr J.H.P. Hackstein for critically reading the manuscript. In addition, we thank all volunteers and M. Spaans (Burger's Zoo, Arnhem, The Netherlands) for providing fecal samples.

REFERENCES

1. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: The unseen majority. Proc Natl Acad Sci USA 1998; 95: 6578–83.
2. Savage DC. Microbial ecology of the gastrointestinal tract. Ann Rev Microbiol 1977; 31: 107–33.
3. Tannock GW. Normal Microflora. An Introduction to Microbes Inhabiting the Human Body. London: Chapman and Hall, 1995.
4. Mackie RI, Sghir A, Gaskins HR. Developmental microbial ecology of the neonatal gastrointestinal tract. Am J Clin Nutr 1999; 69 (suppl): 1035S–45S.
5. Falk PG, Hooper LV, Midtvedt T, Gordon JI. Creating and maintaining the gastrointestinal ecosystem: What we know and need to know from gnotobiology. Microbiol Mol Biol Rev 1998; 62: 1157–70.
6. Strauss E. Microbes feature as pathogens and pals at gathering. Science 1999; 284: 1916–7.
7. Bry L, Falk PG, Midtvedt T, Gordon JI. A model of host-microbial interactions in an open mammalian ecosystem. Science 1996; 273: 1381–3.
8. Hooper LV, Xu J, Falk PG, Midtvedt T, Gordon JI. A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. Proc Natl Acad Sci USA 1999; 96: 9833–8.
9. Hooper LV, Wong MH, Thelin A, Hansson L, Falk PG, Gordon JI. Molecular analysis of host-microbial relationships in the intestine. Science 2000; 291: 881–4.
10. Hackstein JHP, Stumm CK. Methane production in terrestrial arthropods. Proc Natl Acad Sci USA 1994; 91: 5441–5.
11. Hackstein JHP, van Alen TA. Fecal methanogens and vertebrate evolution. Evolution 1996; 50: 559–72.
12. Florin THJ, Zhu G, Kirk KM, Martin NG. Shared and unique environmental factors determine the ecology of methanogens in humans and rats. Am J Gastroenterol 2000; 95: 2872–9.
13. Miller TL, Wolin MJ. Methanogens in human and animal intestinal tracts. Syst Appl Microbiol 1986; 7: 223–9.
14. McCartney AL, Wenzhi W, Tannock GW. Molecular analysis of the composition of the Bifidobacterial and *Lactobacillus* microflora of humans. Appl Environ Microbiol 1996; 62: 4608–13.
15. Covacci A, Telford JL, Del Giudice G, Parsonnet J, Rappouli R. *Helicobacter pylori* virulence and genetic geography. Science 1994; 284: 1328–33.
16. McFarlene GT, Gibson GR. Metabolic activities of the normal colonic microflora. In: Gibson SAW, ed. Human Health: Contribution of Micro Organisms. Frankfurt: Springer, 1994: 17–38.
17. Langendijk PS, Schut F, Jansen GJ, Raangs GC, Camphuis GR, Wilkinson MF, Welling GW. Quantative fluorescence in situ hybridization of *Bifidobacterium* spp. with genus-specific 16S rRNA-targeted probes and its application in fecal samples. Appl Environ Microbiol 1996; 61: 3069–75.
18. Wilson KH, Blitchington RH. Human colonic biota studied by ribosomal DNA sequence analysis. Appl Environ Microbiol 1996; 62: 2273–8.
19. Amann RI, Ludwig W, Schleifer K-H. Phylogenetic identification and in situ detection of individual cells without cultivation. Microbiol Rev 1995; 59: 143–69.
20. Woese CR. Bacterial Evolution. Microb Rev 1987; 51: 221–71.
21. Woese CR. A definition of the domains Archaea, Bacteria, and Eucarya in terms of small subunit ribosomal characteristics. Syst Appl Microbiol 1990; 14: 305–10.
22. Zoetendal EG, Akkermans ADL, de Vos WM. Temperature gradient gel electrophoresis analysis from human fecal samples reveals stable and host-specific communities of active bacteria. Appl Environ Microbiol 1998; 64: 3854–9.
23. Franks AH, Harmsen HJ, Raangs GC, Jansen GJ, Schut F, Welling GJ. Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Appl Environ Microbiol 1998; 64: 3336–45.
24. Fischer SG, Lerman LS. Lenght-independent separation of DNA restriction fragments in two-dimensional gel electrophoresis. Cell 1979; 16: 191–200.
25. Muyzer G, de Waal EC, Uitterlinden GA. Profiling of complex populations by denaturating gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 1993; 59: 695–700.
26. Mitsuoka T. Intestinal flora and aging. Nutr Rev 1992; 50: 438–46.
27. Nübel U, Engelen B, Felske A, et al. Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. J Bacteriol 1996; 178: 5636–43.
28. Sanguinetti CJ, Dias Neto E, Simpson AJG. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. Biotechniques 1994; 17: 915–9.
29. Häne BG, Jäger K, Drexler H. The pearson product-moment correlation coefficient is better suited for identification of DNA fingerprint profiles than band matching algorithms. Electrophoresis 1993; 14: 967–72.

30. van der Waaij D, Berghuis-de Vries JM, Lekkerkerk-van der Wees JEC. Colonization resistance of the digestive tract in conventional and antibiotic-treated mice. J Hyg 1971; 67: 405–11.
31. Muyzer G, Smalla K. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. Antonie van Leeuwenhoek 1998; 73: 127–41.
32. Vaughan EE, Schut F, Heilig GHJ, Zoetendal EG, de Vos WM, Akkermans ADL. A molecular view of the intestinal ecosystem. Curr Issues Intest Microbiol 2000; 1: 1–12.
33. Murray AE, Preston CM, Massana R, et al. Phylogenetic compositions of bacterioplankton from two California estuaries compared by denaturing gradient gel electrophoresis of 16S rDNA fragments. Appl Environ Microbiol 1998; 62: 2676–80.
34. Simpson JM, McCracken VJ, White BA, Gaskins HR, Mackie RI. Application of denaturing gradient gel electrophoresis for the analysis of the porcine gastrointestinal microbiota. J Microbiol Methods 1999; 36: 167–79.