APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Feb. 2007, p. 1388–1392 Vol. 73, No. 4
0099-2240/07/$08.00+0 doi:10.1128/AEM.01921-06

Metaproteomics Approach To Study the Functionality of the Microbiota in the Human Infant Gastrointestinal Tract▽

Eline S. Klaassens,* Willem M. de Vos, and Elaine E. Vaughan

Laboratory of Microbiology, Wageningen University, 6703 CT Wageningen, The Netherlands

Received 11 August 2006/Accepted 3 December 2006

A metaproteomics approach comprising two-dimensional gel electrophoresis and matrix-assisted laser desorption ionization–time of flight (mass spectrometry) was applied to the largely uncultured infant fecal microbiota for the first time. The fecal microbial metaproteome profiles changed over time, and one protein spot contained a peptide sequence that showed high similarity to those of bifidobacterial transaldolases.

The human gastrointestinal tract is rapidly colonized during the first days of life by microbes (4, 10) which ultimately contribute significantly to host nutrition and immunity among other beneficial effects (1, 8). Analysis of 16S ribosomal DNA clone libraries revealed numerous undescribed species in intestinal samples, emphasizing the importance of techniques that bypass cultivation of microbiota (3, 4, 15, 17, 20) such as proteomics. So far there have been two reports of the use of metaproteomics to characterize complex bacterial ecosystems, illustrating the feasibility of this approach in analyzing complex ecosystems (12, 19). In the present study we investigated the potential use of metaproteomics for characterization of the human fecal microbiota.

Dynamics of the intestinal microbiota. The relatively simple infant fecal ecosystem was chosen for this study as it is dominated by bifidobacteria, which were used in this study to optimize the protein isolation and two-dimensional (2D) gel procedures (data not shown). Informed consent was obtained from parents for use of the fecal samples of their infants. Infant fecal samples containing high bifidobacterial content were selected by bifidobacterium-specific fluorescent in situ hybridization (6, 22). Infants A (8 days old) and B (117 days old) harbored 45% and 63% bifidobacteria in the total microbiota, respectively. PCR-denaturing gradient gel electrophoresis (DGGE) (14, 21, 23) of the 16S rRNA gene was performed to monitor the total bacterial community (Fig. 1). The profiles were relatively simple (Fig. 1) (4), as expected, and the predominance of bifidobacteria was supported by the presence of abundant 16S rRNA gene amplicons that comigrated with the control *Bifidobacterium longum* strain and was confirmed by sequencing (14).

Metaproteome production of infant fecal microbiota. Fecal samples were collected from the infants prior to weaning: for infant A, at days 8, 24, and 41, and for infant B, at days 103, 117, and 144. Microbial cells were released from the feces and washed as previously described (5, 22). The bead beating method was confirmed to be applicable for protein extraction of infant fecal microbiota containing bifidobacteria (data not shown), as expected from previous systematic studies (5, 14, 20, 23, 24). Total soluble protein was obtained by three treatments of 45 s of bead beating (FastPrep; Qbiogene), interspersed by 1 min on ice, in 500-μl isoelectric focusing (IEF) buffer (Fluka, Switzerland) (10 M urea) containing 2% CHAPS (3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate) (Roche, Switzerland), 0.65 mM dithiothreitol (DTT) (Sigma, Switzerland), 0.2% Biolytes 3/10 (Bio-Rad), Pefabloc Sc (Fluka), and glass beads (Sigma) (≤0.1 mm). Protein concentrations were determined using a 2-D Quant kit (Amersham Biosciences). Each protein sample (100 μg) was loaded onto 11-cm immobilized pH gradient (IPG) ReadyStrips (Bio-Rad) (pH 4 to 7) and rehydrated for 12 h. IEF was carried out for 98,000 Vh (Protean IEF; Bio-Rad). Reduction and alkylation of proteins were performed prior to electrophoresis in the second dimension by incubating for 10 min in 6 M Urea–0.1 M Tris-HCl (pH 8.8)–2% (wt/vol) sodium dodecyl sulfate (SDS) with 130 mM DTT followed by a 10 min incubation in the presence of 216 mM iodoacetamide (instead of DTT). The IPG strip was positioned on an SDS-polyacrylamide electrophoresis gel (Criterion gel, Bio-Rad) (12.5% polyacrylamide) with 1% low melting agarose in 40 mM Tris-HCl (pH 6.8). Electrophoresis was run at 100 V and 30 W. Silver staining was performed as described previously (16). Protein maps were scanned with a GS-800 densitometer (Bio-Rad) and analyzed with PDQuest software (Bio-Rad). Triplicate 2D gels for each sample were grouped. For each infant, three groups, each representing one time point in life, were compared using the quantitative function within PDQuest software. The mean coefficient of variation (CV) (standard deviation/mean × 100) is a quantitative index for variation of quantities among matched spots and was computed for gel-to-gel variations within each replicate group. More than 200 protein spots were visualized on each gel. A comparison of gels of infant A revealed changes in number and intensities of protein spots during the 33 days, although the patterns remained similar (Fig. 2A to C). In order to clearly demonstrate the changes a section of the gel corresponding to each time point was enlarged (Fig. 3). Comparisons of differential protein production values with percent CV values for various time points for the marked spots are illustrated with bar graphs (Fig. 3C). The metaproteome profiles of infant A (Fig. 2A to C) were different from those of infant B (see, e.g., Fig.

* Corresponding author. Mailing address: Laboratory of Microbiology, Wageningen University, Hesselink van Suchtelenweg 4, 6703 CT Wageningen, The Netherlands. Phone: 31 317 483486. Fax: 31 317 483829. E-mail: Eline.klaassens@wur.nl.

▽ Published ahead of print on 8 December 2006.

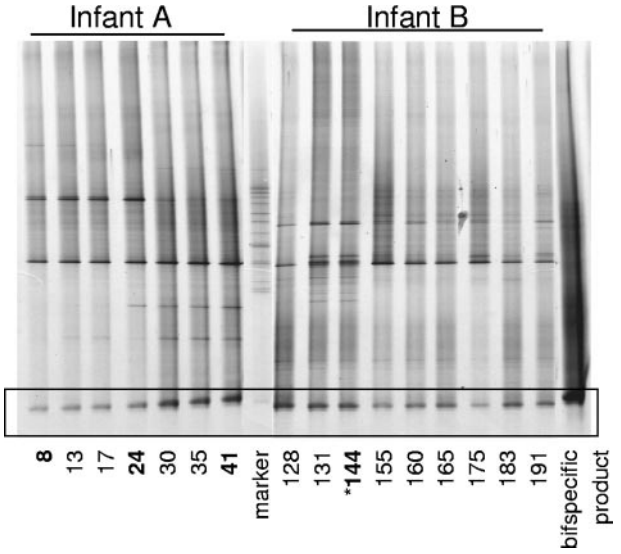

FIG. 1. PCR-DGGE profiles of regions V8 to V6 of 16S rRNA genes of fecal samples from two infants at different ages (infant A, 8 to 41 days; infant B, 128 to 191 days), reflecting the predominant bacterial community. The bands in the box represent amplicons of bifidobacterial species, including *Bifidobacterium longum* DSM 20219. *, start of weaning period for infant A. Sampling days for protein extraction are shown in bold. marker, mix of DNA size standards; bifspecific product, amplicon of 16S rRNA gene of *B. longum* DSM 20219.

2D) in accordance with the uniqueness of each individual's microbiota (20), as was also observed in the PCR-DGGE profiles obtained as described above.

Identification of proteins. A total of 55 protein spots were excised from the infant A 2D gel at day 41 (Fig. 3A). Peptides were extracted after tryptic digestion and analyzed by matrix-assisted laser desorption ionization–time of flight (mass spectrometry) [MALDI-TOF (MS)] (2) (Applied Biosystems 4700 proteomics analyzer; Technology Facility, Department of Biology, The University of York, York, United Kingdom), which provided a catalogue of 21 good mass peptide fingerprints. Peptide sequences were searched using BLASTP, with the default settings of the "Search for short, nearly exact matches" function, as well as MS-Pattern in Protein Prospector 4.0.5 (The University of California). As expected, the peptides' mass spectra showed low similarity to those of database entries and no similarities to those of human proteins. The mean differential protein amounts of the 21 spots with clean signals for days 8, 24, and 41 are indicated in Fig. 3A and presented in Table 1. MALDI-TOF (MS) of spots 1, 2, 7, and 11 was performed for de novo peptide sequencing (Fig. 3B), which resulted in determination of 11 N-terminal sequences that shared similarity to those of bacterial proteins, a viral protein, and four eukaryal proteins, but the collective sequences were not significantly similar to any previously reported.

One complete peptide fragment (ELAEATDFVDGR) of

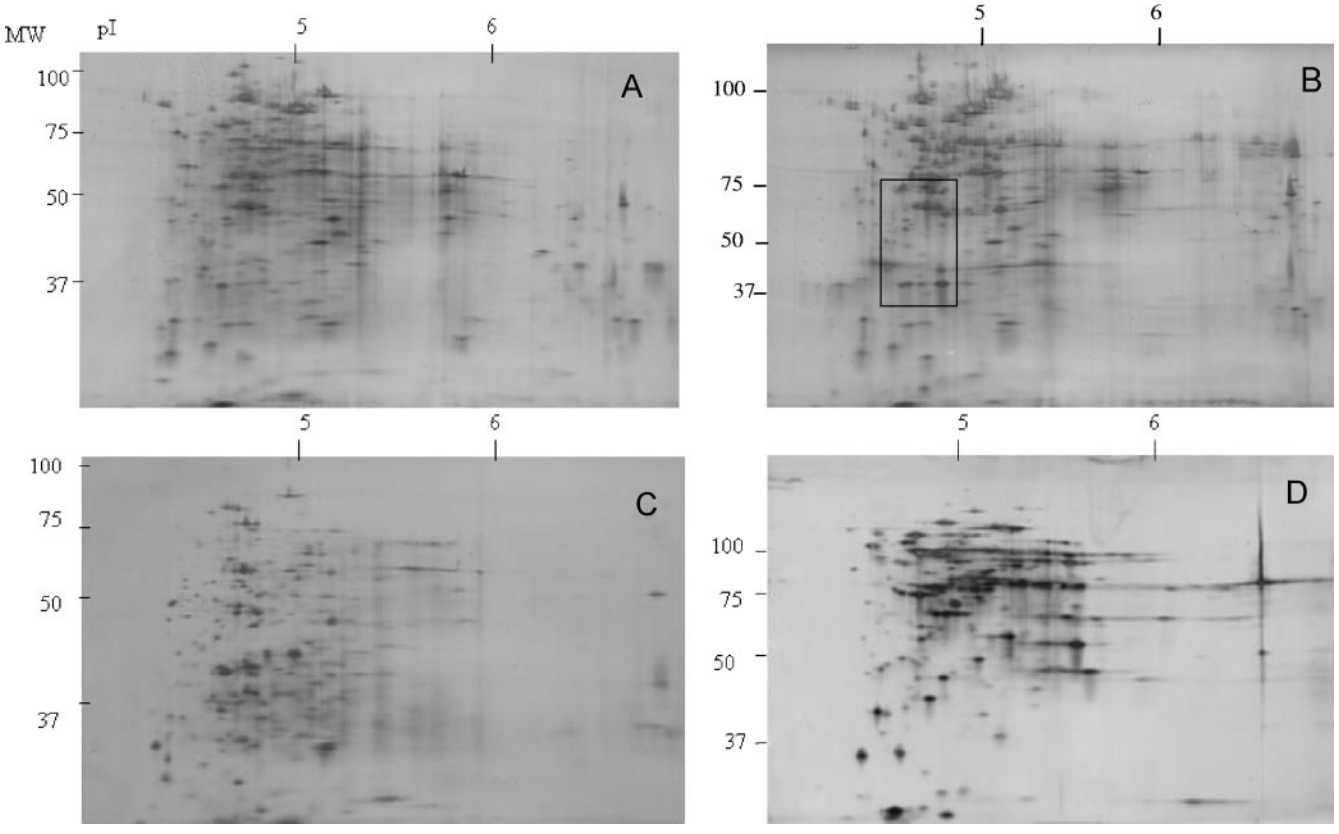

FIG. 2. Silver-stained 2D gels of fecal samples from infant A at the ages of 8 days (A), 24 days (B), and 41 days (C) and from infant B at 117 days (D); data were obtained using IPG (pH 4 to 7). MW, molecular weight.

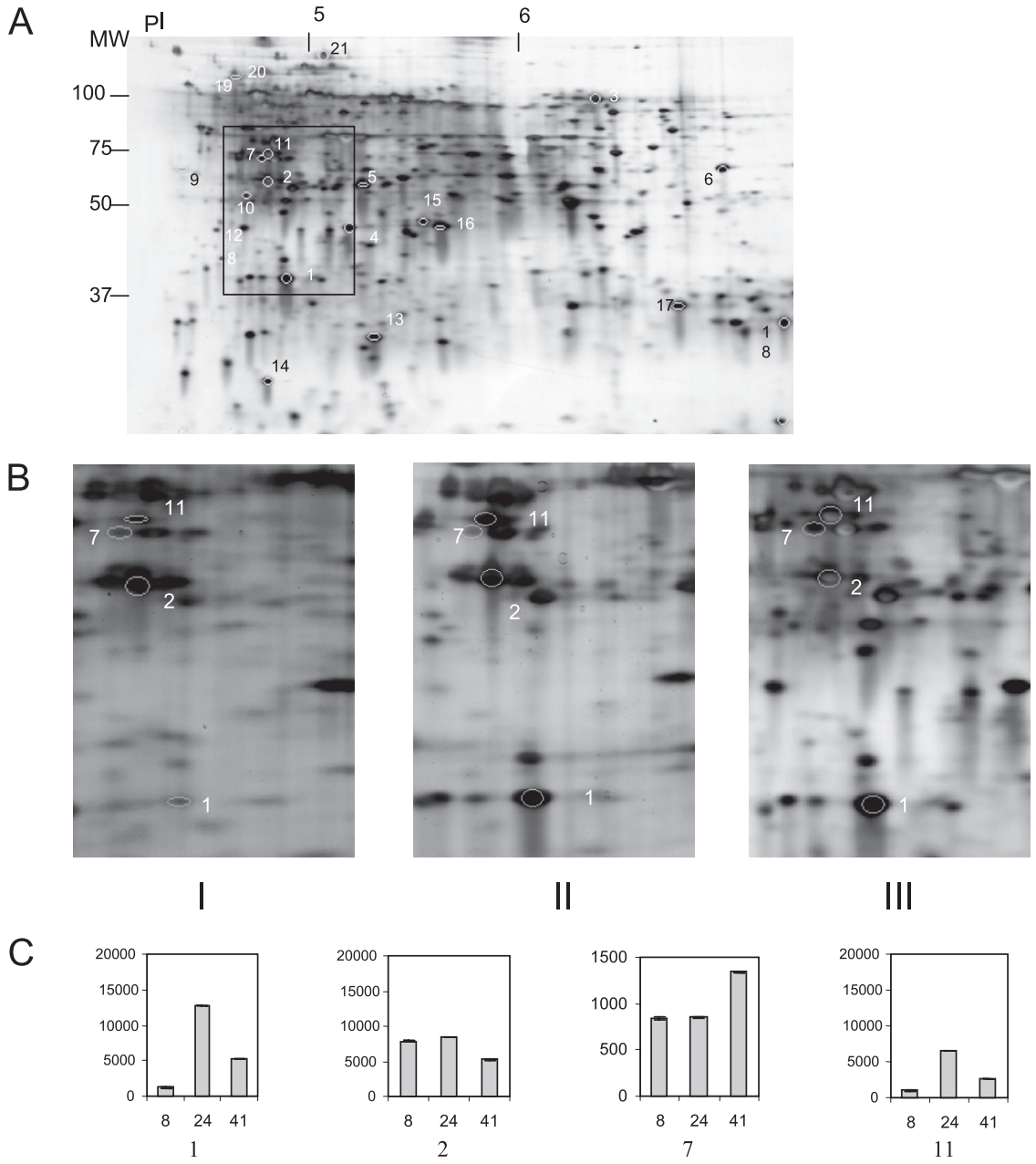

FIG. 3. (A) Silver-stained 2D gel of a fecal sample from infant A at the age of 41 days obtained using IPG (pH 4 to 7). The circled protein spots were excised, and mass fingerprints were obtained by MALDI-TOF (MS). Table 1 shows the quantity of the selected protein spots. MW, molecular weight. (B) Areas of interest enclosed in the boxes are enlarged from gels of samples from infant A at the ages of 8 (I), 24 (II), and 41 (III) days. The circled protein spots (1, 2, 7, and 11) gave good peptide sequencing results with MALDI-TOF (MS). (C) The diagrams demonstrate changes in intensity of the protein spots (1, 2, 7, and 11) (see panel BI) with percent CV values at ages 8, 24, and 41 days.

protein spot 4 gave a result showing high-level database matches of 91% identity (i.e., one mismatch) with the *Bifidobacterium infantis* and *B. longum* NCC2705 transaldolases and 83% identity with the DJO10A transaldolase. Production for this protein spot was relatively high at day 8 but increased approximately 10-fold and 6-fold at days 24 and 41, respectively, possibly due to the increase in numbers and activity of bifidobacteria in the infants' microbiota, based on PCR-DGGE (Table 1 and Fig. 1). The data obtained for the digestion sites of trypsin and the theoretical pI (4.87) and molecular mass (39.6 kDa) of the *B. infantis* transaldolase protein agreed with the position of the protein spot in the 2D gel (Fig. 3A). This transaldolase gene is a common target for PCR used to detect and enumerate bifidobacteria (13). Transaldolase was identified in a proteomic study by Vitali et al. (18) of *B. infantis* B107, where it represented, together with nine other proteins, the most abundant portion of the proteome. The dominance of the (tentatively identified) bifidobacterial transaldolase protein in the feces of a newborn infant may explain its detection by a metaproteomics approach.

TABLE 1. Mean quantity of selected protein spots over time with % CV and experimental pI and MW values

Protein quantity change category and protein spot no.	Mean protein quantity in ppm per spot area (% CV) for infant A at age:			Exptl value	
	8 days	24 days	41 days	MW	pI
Increasing					
1[a]	1,263 (65)	12,780 (18)	6,988 (41)	39	4.8
2[a]	5,534 (141)	8,415 (0)	5,240 (52)	50	4.8
3	ND[b]	8,221 (38)	3,521 (1)	88	6.0
4	2,344 (12)	5,135 (2)	4,060 (35)	45	5.2
5	11,283 (20)	8,266 (19)	3,480 (0)	50	5.2
6	S	10,547 (18)	7,755 (8)	55	6.5
Limited change					
7[a]	648 (55)	705 (37)	1,340 (3)	65	4.7
8	2,308 (21)	1,363 (25)	2,093 (37)	41	4.6
9	1,631 (15)	837 (22)	804 (17)	60	4.4
10	1,415 (28)	4,207 (59)	3,817 (1)	46	4.6
11[a]	1,044 (10)	786 (2)	1,007 (5)	70	4.8
12	387 (15)	290 (23)	351 (23)	42	4.6
Decreasing					
13	2,722 (1)	3,472 (0)	7,533 (14)	37	5.2
14	2,399 (21)	3,902 (10)	7,339 (41)	28	4.8
15	ND	1,571 (3)	4,974 (22)	47	5.3
16	2,148 (27)	2,899 (24)	6,472 (30)	46	5.4
17	ND	3,545 (23)	10,520 (3)	37	6.3
18	2,844 (10)	2,652 (23)	9,896 (13)	35	6.8
19	ND	ND	925 (23)	90	4.5
20	1,711 (6)	1,474 (1)	2,952 (15)	95	4.6
21	S[c]	4,661 (18)	4,685 (3)	100	5.0

[a] Protein spots selected for peptide sequencing.
[b] ND, not detectable.
[c] S, saturated.

Perspectives. For the first time, reproducible 2D gels, extraction of proteins, and tentative identification using MALDI-TOF (MS) demonstrated the applicability of the proteomics approach for the complex intestinal ecosystem. Currently insufficient microbiome sequence information confounds identification of the proteins, but ongoing metagenomic library analysis will enable meaningful identification in time (7, 9, 11). Furthermore, peptide mass fingerprints from sequence data will be sufficient to produce more statistically valid database matches, as was recently demonstrated by Ram et al. (12), who matched 6,000 peptide fragments to DNA sequences of an accompanying metagenomic library from a low-complexity natural microbial biofilm. Metaproteomics approaches may become a useful tool to monitor the functional products of the microbiota in feces over time as affected by dietary intervention, length life, health, and disease.

Nucleotide sequence accession numbers and peptide sequences. The bifidobacterial nucleotide sequences of partial 16S rRNA genes have been deposited in the GenBank database under accession no. DQ323457, DQ323458, and DQ323459; the 11 peptide sequences are DLAVALSENKR, ATNSEL MHVGVSR, ELAEATDFVDGR, and PSSKVGSGSSGA GALK (for spot 1), DVAPDLALMHTKLSR (for spot 2), ADNFEGDDR and TAFTGYETLR (for spot 7), and KTGP KLFAADEALK, HYGLASDALANGGCVDSVSDSPA, AP MVALSELER, and TANSELLEAELAR (for spot 11).

This work was supported by the Dutch Ministry of Economic Affairs through the Innovation Oriental Research Program on Genomics (IOP Genomics: IGE01016).

REFERENCES

1. **Bäckhed, F., R. E. Ley, J. L. Sonnenburg, D. A. Peterson, and J. I. Gordon.** 2005. Host-bacterial mutualism in the human intestine. Science **307:**1915–1920.
2. **Curwen, R. S., P. D. Ashton, D. A. Johnston, and R. A. Wilson.** 2004. The *Schistosoma mansoni* soluble proteome: a comparison across four life-cycle stages. Mol. Biochem. Parasitol. **138:**57–66.
3. **Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson, and D. A. Relman.** 2005. Diversity of the human intestinal microbial flora. Science **308:**1635–1638.
4. **Favier, C. F., E. E. Vaughan, W. M. de Vos, and A. D. L. Akkermans.** 2002. Molecular monitoring of succession of bacterial communities in human neonates. Appl. Environ. Microbiol. **68:**219–226.
5. **Fitzsimons, N. A., A. D. L. Akkermans, W. M. de Vos, and E. E. Vaughan.** 2003. Bacterial gene expression detected in human faeces by reverse transcription-PCR. J. Microbiol. Methods **55:**133–140.
6. **Franks, A. H., H. J. M. Harmsen, G. C. Raangs, G. J. Jansen, F. Schut, and G. W. Welling.** 1998. Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Appl. Environ. Microbiol. **64:**3336–3345.
7. **Gill, S. R., M. Pop, R. T. DeBoy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett, and K. E. Nelson.** 2006. Metagenomic analysis of the human distal gut microbiome. Science **312:**1355–1359.
8. **Hooper, L. V., T. Midtvedt, and J. I. Gordon.** 2002. How host-microbial interactions shape the nutrient environment of the mammalian intestine. Annu. Rev. Nutr. **22:**283–307.
9. **Ley, R. E., D. A. Peterson, and J. I. Gordon.** 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell **124:**837.
10. **Mackie, R. I., A. Sghir, and H. R. Gaskins.** 1999. Developmental microbial ecology of the neonatal gastrointestinal tract. Am. J. Clin. Nutr. **69:**S1035–S1045.

11. **Manichanh, C., L. Rigottier-Gois, E. Bonnaud, K. Gloux, E. Pelletier, L. Frangeul, R. Nalin, C. Jarrin, P. Chardon, P. Marteau, J. Roca, and J. Dore.** 2006. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut **55:**205–211.
12. **Ram, R. J., N. C. VerBerkmoes, M. P. Thelen, G. W. Tyson, B. J. Baker, R. C. Blake, I. I., M. Shah, R. L. Hettich, and J. F. Banfield.** 2005. Community proteomics of a natural microbial biofilm. Science **308:**1915–1920.
13. **Requena, T., J. Burton, T. Matsuki, K. Munro, M. Simon, R. Tanaka, K. Watanabe, and G. W. Tannock.** 2002. Identification, detection, and enumeration of human *Bifidobacterium* species by PCR targeting the transaldolase gene. Appl. Environ. Microbiol. **68:**2420–2427.
14. **Satokari, R. M., E. E. Vaughan, A. D. L. Akkermans, M. Saarela, and W. M. de Vos.** 2001. Bifidobacterial diversity in human feces detected by genus-specific PCR and denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. **67:**504–513.
15. **Satokari, R. M., E. E. Vaughan, H. Smidt, M. Saarela, J. Mättö, and M. de Vos.** 2003. Molecular approaches for the detection and identification of bifidobacteria and lactobacilli in the human gastrointestinal tract. Syst. Appl. Microbiol. **26:**572–584.
16. **Shevchenko, A., M. Wilm, O. Vorm, and M. Mann.** 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal. Chem. **68:**850–858.
17. **Suau, A., R. Bonnet, M. Sutren, J. Godon, G. R. Gibson, M. D. Collins, and J. Dore.** 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.
18. **Vitali, B., V. Wasinger, P. Brigidi, and M. Guilhaus.** 2005. A proteomic view of *Bifidobacterium infantis* generated by multi-dimensional chromatography coupled with tandem mass spectrometry. Proteomics **5:**1859–1867.
19. **Wilmes, P., and P. L. Bond.** 2004. The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. Environ. Microbiol. **6:**911–920.
20. **Zoetendal, E. G., A. D. L. Akkermans, and W. M. 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.
21. **Zoetendal, E. G., K. Ben-Amor, A. D. L. Akkermans, T. Abee, and W. M. de Vos.** 2001. DNA isolation protocols affect the detection limit of PCR approaches of bacteria in samples from the human gastrointestinal tract. Syst. Appl. Microbiol. **24:**405–410.
22. **Zoetendal, E. G., K. Ben-Amor, H. J. M. Harmsen, F. Schut, A. D. L. Akkermans, and W. M. de Vos.** 2002. Quantification of uncultured *Ruminococcus obeum*-like bacteria in human fecal samples by fluorescent in situ hybridization and flow cytometry using 16S rRNA-targeted probes. Appl. Environ. Microbiol. **68:**4225–4232.
23. **Zoetendal, E. G., C. C. G. M. Booijink, E. S. Klaassens, H. G. H. J. Heilig, M. Kleerebezem, H. Smidt, and W. M. de Vos.** 2006. Isolation of DNA from bacterial samples of the human gastrointestinal tract. Nat. Protocols **1:**870–873.
24. **Zoetendal, E. G., C. C. G. M. Booijink, E. S. Klaassens, H. G. H. J. Heilig, M. Kleerebezem, H. Smidt, and W. M. de Vos.** 2006. Isolation of RNA from bacterial samples of the human gastrointestinal tract. Nat. Protocols **1:**954–959.