Appendix 1: Complete study protocols

Sample collection

Fecal samples were collected from infants at 3–4 months of age as part of a scheduled home visit. Mothers placed a liner in their child's diaper the day before the home visit and changed the liner with the diaper until stool was obtained. The diaper with liner was placed in a collection bag, sealed and kept in a refrigerator until collected by trained research assistants using a screw-cap container with spoon (Globe Scientific, Paramus, New Jersey). The container was placed in a cooler for transport and then placed in a –80°C freezer for long-term storage. The times of passing the stool, and collecting and freezing the sample were noted.

DNA extraction and amplification

Whole-genome DNA was extracted from 40 mg of stool with use of the FastPrep DNA for Soil Kit (MP Biomedicals Inc., Solon, Ohio).[1] The bacterial 16S rRNA gene, hypervariable regions V5–V7, was amplified by means of polymerase chain reaction (PCR) using primers optimized for gut-occurring taxa, including the genus *Bifidobacterium*: 5′-GGGKAKCRAACVGGATTAGATACCCBGGTAGTCCWNRCH STAAACGDTG-3′ and 5′-GGSCRTRMKGAYTTGACGTCRYCCCCDCCTTCCTCC-3′. The primers were barcoded so that each sample could be uniquely identified after sequencing. Each PCR mixture (50 µL) contained 5 µL of 10X Hotstart Buffer, 400 mM of deoxyribonucleotide triphosphates (dNTPs), 1.5 mM of magnesium chloride, 2.5 U of Hotstart Taq polymerase (Fermentas, Glen Burnie, Maryland), 1 mg of Ultrapure Bovine Serum Albumin (Ambion, Austin, Texas), molecular biology reagent grade water (Sigma-Aldrich, St. Louis, Missouri), 0.16 µM of primer and 2 µL of bacterial template DNA (10 ng/µL). The PCR program consisted of an initial DNA denaturation step at 94°C (4 minutes), followed by 18 cycles of DNA denaturation at 94°C (45 seconds), an annealing step at 56°C (30 seconds) and an elongation step at 72°C (2 minutes and 30 seconds), and was performed on the PTC-200 Peltier Thermal Cycler (MJ Research, St. Bruno, Québec). The PCR product was cleaned with GENECLEAN Turbo Kit (MP Biomedicals Inc, Solon, Ohio) and gel purified using the E-gel SizeSelect 2% agarose gel cutting system (Invitrogen, Carlsbad, California). Fifty nanograms of cleaned/extracted product from each sample was combined and concentrated for sequencing using an Amicon Ultra-4 30K centrifugal filter (Millipore, Billerica, Massachusetts).

16S rRNA sequencing and taxonomic classification

Pooled PCR amplicons were sequenced at portions of the V5, V6 and V7 hypervariable regions of the 16S rRNA gene using serial Illumina sequencing (SI-Seq) at the University of Toronto Centre for the Analysis of Genome Evolution and Function (CAGEF), as previously described.[2]

These reads were concatenated to a final length of 144 bp and then processed through the SI-Seq analysis pipeline for de-barcoding and quality filtering, which removed reads having more than 10 sites with a Phred quality score of less than 20. The resulting high-quality reads were denoised, cleared of chimeras and clustered into operational taxonomic units (OTUs) using the otupipe scripts.[3] An empirically derived nucleotide identity threshold of 87% was used for OTU clustering.[2] One representative sequence from each OTU was classified according to the SILVA taxonomy by 95% identity (i.e., genus level) clustering, with the SILVA database sequences formatted to SI-Seq read structure. Within each sample, OTUs with abundances of less than 0.18% were removed from the analysis based on an empirically derived misclassification/sequencing error rate.[2] After cleaning and processing, a total of 2.86 million reads were retained (median 9.6×10^4 per sample, range 1.5×10^4 to 4.5×10^5).

Appendix to: Azad MB, Konya T, Maughan H, et al.; the CHILD Study Investigators. Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. *CMAJ* 2013. DOI:10.1503/cmaj.121189.

Quantitative PCR

We followed the method of Penders and associates[4] for quantitative PCR analysis of *Clostridium difficile*. Oligonucleotides were manufactured by IDT (Integrated DNA Technologies Inc., Coralville, Iowa). All reactions were performed on the MiniOpticon Real-Time PCR System (Bio-Rad, Hercules, California).

References

1. Ariefdjohan MW, Savaiano DA, Nakatsu CH. Comparison of DNA extraction kits for PCR-DGGE analysis of human intestinal microbial communities from fecal specimens. *Nutr J* 2010;9:23.
2. Maughan H, Wang PW, Diaz CJ, et al. Analysis of the Cystic Fibrosis Lung Microbiota via Serial Illumina Sequencing of Bacterial 16S rRNA Hypervariable Regions. *PLoS One* 2012;7:e45791.
3. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;26:2460-1.
4. Penders J, Thijs C, Vink C, et al. Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* 2006;118:511-21.

Appendix to: Azad MB, Konya T, Maughan H, et al.; the CHILD Study Investigators. Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months. *CMAJ* 2013. DOI:10.1503/cmaj.121189.