SUPPLEMENTAL METHODS

Gut microbiota analysis by qPCR

The primers used to detect *Bifidobacterium* spp., *Lactobacillus* spp. and *Lactobacillus acidophilus* were based on 16S rRNA gene sequences: F-*Bifidobacterium* spp. TCGCGTCYGGTGTGAAAG, R-*Bifidobacterium* spp. CCACATCCAGCRTCCAC, F-*Lactobacillus* spp. AGCAGTAGGGAATCTTCCA, R-*Lactobacillus* spp. CACCGCTACACATGGAG, F-*L. acidophilus* CCTTTCTAAGGAAGCGAAGGAT and R-*L. acidophilus* AATTCTCTTCTCGGTCGCTCTA.

PCR amplification was carried out as follows: 10 min at 95°C followed by 45 cycles of 3 s at 95°C, 26 s at 58°C or 60°C (*Lactobacillus* spp. and *L. acidophilus* or *Bifidobacterium* spp., respectively) and 10 s at 72°C. Detection was achieved with an STEP one PLUS instrument and software (Applied Biosystems, Foster City, CA) using MESA FAST qPCR MasterMix Plus for SYBR Assay (Eurogentec, Verviers, Belgium). Each assay was performed in duplicate in the same run. The cycle threshold of each sample was then compared with a standard curve (performed in duplicate) made by diluting genomic DNA (fivefold serial dilution) (BCCM/LMG, Ghent, Belgium). Prior to isolating the DNA, the cell counts were determined in culture and expressed as “colony-forming unit” (CFU). Data are expressed as log CFU/g of feces.

Metabolic profiling by ^{1}H NMR spectroscopy

Urine spectra were acquired using a standard 1-dimensional pulse sequence [recycle delay (RD)-90°-t1-90°-tm-90°-acquire free induction decay (FID)] with water suppression applied during RD of 2 s, a mixing time (tm) of 100 ms and a 90° pulse set at

7.70 µs. For each spectrum, a total of 128 scans were accumulated into 64 K data points with a spectral width of 9803 Hz. The FIDs were multiplied by an exponential function corresponding to 0.3 Hz line broadening. All spectra were manually phased, baseline corrected and calibrated to the chemical shift of DSS (δ 0.00).

Plasma spectra were acquired using a Carr Purcell Meiboom Gill sequence,(1) with water suppression, using a 90° pulse set at 7.70 µs, a delay (D20) of 152 µs and 108 loops (L4). A total of 128 scans were accumulated into 32 K data points with a spectral width of 9803 Hz. The FIDs were multiplied by an exponential function corresponding to 0.3 Hz line broadening. All spectra were manually phased, baseline corrected and calibrated to the chemical shift of glucose (δ 5.22).

A range of 2D NMR spectra were performed on the same equipment for selective samples, including correlation spectroscopy (COSY),(2), total correlation spectroscopy (TOCSY),(3) and heteronuclear single quantum coherence (HSQC) NMR spectroscopy,(4). Metabolites were assigned using our in house standard database, data from literature,(5;6) and confirmed by 2D NMR experiments.

Statistical analysis of metabolic profiling

All spectral regions between 0.5 – 10 p.p.m. were imported in Matlab® version R2010a (Mathworks® UK) and statistical algorithms were provided by Korrigan Sciences (Korrigan Sciences Ltd, UK). To minimise variability due to water presaturation, the water resonance region (δ 4.70 – 5.05) was removed. Urine data were then aligned and normalized to the probabilistic quotient as previously described,(7;8). Plasma data were neither aligned, nor normalized since plasma concentration is considered constant among

patients. All statistical models were performed using unit variance scaling. For each dataset (urine and plasma), principal component analyses (PCA) were performed on all spectra in order to detect any outlier and to identify potential patterns associated with volunteers, or prebiotic effect. PCA is a latent variable projection method that produces linear combinations of the original variables to generate the axes, also known as principal components (PCs),(9). A series of partial least squares (PLS) regressions were then performed on each metadata and microbiological data (Y predictors) using metabolic profiles as independent variables (X matrix). These models were all performed using one predictive component and validated by random permutations (500 times) to simulate the null hypothesis. A p-value was calculated by rank determination of the model actual Q^2Y value (representing the goodness of prediction) among the Q^2Y values calculated for the 500 permutated models. The model R^2Y value represents the goodness of fit and the R^2X value, the percentage of variance of X explaining Y.

DATA SUPPLEMENT

Metabolic profiling by ^{1}H NMR spectroscopy

PLS parameters of the models listed in the main text:

- Correlation between PLS scores of plasma NMR spectra and *Propionibacterium*: $R^2Y=0.30$, $Q^2Y=0.16$, $R^2X=0.18$; $p=0.002$ after 500 random permutations.
- Correlation between PLS scores of plasma NMR spectra and *Bacteroides vulgatus*: $R^2Y=0.25$, $Q^2Y=0.08$, $R^2X=0.21$; $p=0.01$ after 500 random permutations.
- Correlation between PLS scores of urine NMR spectra and waist/hip ratio: $R^2Y=0.26$, $Q^2Y=0.09$, $R^2X=0.12$; $p=0.01$ after 500 random permutations.

- Correlation between PLS scores of urine NMR spectra and post-OGTT insulin: $R^2Y=0.31$, $Q^2Y=0.07$, $R^2X=0.10$; $p=0.03$ after 500 random permutations.
- Correlation between PLS scores of urine NMR spectra and *Collinsella*: $R^2Y=0.64$, $Q^2Y=0.29$, $R^2X=0.10$; $p=0.01$ after 500 random permutations.

SUPPLEMENTAL TABLE

	Relative contribution (%)	Preb-1	SD	Preb-2	SD	p values	corr. p values	Plac-1	SD	Plac-2	SD
Actinobacteria		**0,23**	**0,20**	**1,92**	**2,18**	**0,0012**		**0,44**	**0,59**	**0,58**	**0,96**
	Bifidobacterium	0,17	0,19	1,85	2,15	0,0009	0,05	0,39	0,56	0,52	0,90
	Collinsella	0,03	0,03	0,06	0,04	0,0052	0,13	0,03	0,05	0,04	0,07
	Propionibacterium	0,03	0,03	0,01	0,01	0,0203	0,29	0,02	0,02	0,02	0,02
Bacteroidetes		**67,53**	**12,03**	**57,49**	**9,19**	**0,0107**		**61,22**	**11,00**	**58,29**	**15,15**
	Bacteroides intestinalis et rel.	2,09	0,84	1,54	0,68	0,0494	0,39	2,00	0,80	1,95	0,83
	Bacteroides vulgatus et rel.	4,20	2,52	2,57	1,42	0,0353	0,34	3,69	1,61	3,80	1,83
	Prevotella tannerae et rel.	3,85	1,88	2,60	1,34	0,0052	0,13	3,07	1,02	2,89	0,90
Bacilli		**0,26**	**0,38**	**0,95**	**1,46**	**0,007**		**0,44**	**0,96**	**0,55**	**1,03**
	Lactobacillus gasseri et rel.	0,01	0,01	0,27	0,78	0,0085	0,16	0,02	0,01	0,02	0,03
	Streptococcus bovis et rel.	0,08	0,05	0,18	0,18	0,0245	0,31	0,11	0,19	0,15	0,14
C. cluster IV		**13,82**	**7,79**	**18,65**	**4,85**	**0,0085**		**15,39**	**8,74**	**16,60**	**8,28**
	Faecalibacterium prausnitzii et rel.	4,51	4,17	8,19	5,67	0,0023	0,10	5,65	4,45	5,25	2,89
C. cluster IX	*Megasphaera elsdenii* et rel.	0,07	0,20	0,29	0,81	0,0001	0,02	0,09	0,31	0,14	0,50
C. cluster XIVa	*Anaerostipes caccae* et rel.	0,37	0,38	0,71	0,54	0,0067	0,14	0,45	0,43	0,57	0,44
C. cluster XVI		**0,05**	**0,08**	**0,23**	**0,34**	**0,0203**		**0,03**	**0,04**	**0,04**	**0,07**
	Eubacterium biforme et rel.	0,05	0,07	0,21	0,33	0,0295	0,31	0,03	0,04	0,03	0,07
Proteobacteria	*Oxalobacter formigenes* et rel.	0,08	0,10	0,09	0,09	0,0295	0,31	0,14	0,15	0,15	0,16

Table S1

Phylum	**Group**	**Genus-like group**
Firmicutes	Bacilli	*Aerococcus*
Firmicutes	Bacilli	*Aneurinibacillus*
Firmicutes	Bacilli	*Bacillus*
Firmicutes	Bacilli	*Enterococcus*
Firmicutes	Bacilli	*Gemella*
Firmicutes	Bacilli	*Granulicatella*
Firmicutes	Bacilli	*Lactobacillus gasseri* et rel.
Firmicutes	Bacilli	*Lactobacillus plantarum* et rel.
Firmicutes	Bacilli	*Lactobacillus salivarius* et rel.
Firmicutes	Bacilli	*Lactococcus*
Firmicutes	Bacilli	*Staphylococcus*
Firmicutes	Bacilli	*Streptococcus bovis* et rel.
Firmicutes	Bacilli	*Streptococcus intermedius* et rel.
Firmicutes	Bacilli	*Streptococcus mitis* et rel.
Firmicutes	Bacilli	*Weissella* et rel.
Firmicutes	*Clostridium* cluster IV	*Anaerotruncus colihominis* et rel.
Firmicutes	*Clostridium* cluster IV	*Clostridium cellulosi* et rel.
Firmicutes	*Clostridium* cluster IV	*Clostridium leptum* et rel.
Firmicutes	*Clostridium* cluster IV	*Clostridium orbiscindens* et rel.
Firmicutes	*Clostridium* cluster IV	*Eubacterium siraeum* et rel.
Firmicutes	*Clostridium* cluster IV	*Faecalibacterium prausnitzii* et rel.
Firmicutes	*Clostridium* cluster IV	*Oscillospira guillermondii* et rel.
Firmicutes	*Clostridium* cluster IV	*Papillibacter cinnamivorans* et rel.
Firmicutes	*Clostridium* cluster IV	*Ruminococcus bromii* et rel.
Firmicutes	*Clostridium* cluster IV	*Ruminococcus callidus* et rel.
Firmicutes	*Clostridium* cluster IV	*Sporobacter termitidis* et rel.
Firmicutes	*Clostridium* cluster IV	*Subdoligranulum variable* at rel.
Firmicutes	*Clostridium* cluster XVI	*Bulleidia moorei* et rel.
Firmicutes	*Clostridium* cluster XVI	*Eubacterium biforme* et rel.
Firmicutes	*Clostridium* cluster XVI	*Eubacterium cylindroides* et rel.

Table S2

	Placebo			Prebiotic			
	T0	T3months	Δ	T0	T3months	Δ	p value
HbA1c (%)	5.8 ± 0.5	5.8 ± 0.5	0.0 ± 0.1	5.6 ± 0.5	5.7 ± 0.5	0.0 ± 0.1	0.32
Fasting glycaemia (mg/dl)	99 ± 15	99 ± 18	1 ± 3	97 ± 18	95 ± 16	-2 ± 6	0.36
Post-OGTT glycemia (mg/dl)	109 ± 30	124 ± 44	15 ± 13	130 ± 40	122 ± 47	-9 ± 17	0.008
Fasting insulinaemia (µU/ml)	13 ± 5	13 ± 6	-1 ± 3	18 ± 14	16 ± 9	-2 ± 5	0.90
Post-OGTT insulin (µU/ml)	56 ± 36	53 ± 30	-3 ± 16	69 ± 38	59 ± 38	-9 ± 16	0.29
HOMA index	3.33 ± 1.77	3.40 ± 1.93	0.07 ± 0.69	4.64 ± 5.0	3.84 ± 2.76	-0.80 ± 2.04	0.56
Adiponectinaemia (µg/ml)	21.7 ± 10.8	22.0 ± 10.6	0.3 ± 1.9	20.2 ± 10.5	20.2 ± 9.7	-0,1 ± 1.0	0.30
Total Cholesterol (mg/dl)	206 ± 42	208 ± 48	2 ± 16	201 ± 48	196 ± 50	-7 ± 9	0.69
LDL-cholesterol (mg/dl)	132 ± 37	134 ± 40	1 ± 14	126 ± 40	122 ± 45	-4 ± 7	0.73
HDL-cholesterol (mg/dl)	53 ± 15	53 ± 12	-1 ± 3	51 ± 12	51 ± 14	-1 ± 2	0.63
HDL/LDL ratio	0.44 ± 0.21	0.42 ± 0.12	-0.02 ± 0.07	0.44 ± 0.16	0.48 ± 0.23	0.03 ± 0.05	0.45
Triglycerides (mg/dl)	102 ± 41	108 ± 33	6 ± 18	123 ± 74	113 ± 66.1	-8 ± 16	0.19

Table S3

Reference List

(1) Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. *The review of scientific instruments* 1958;**29**:688-91.

(2) Aue WP, Bartholdi E, Ernst RR. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *Journal of Chemical Physics* 1976;**64**(5):2229-46.

(3) Glaser SJ, Schwalbe H, Marino JP *et al*. Directed TOCSY, a method for selection of directed correlations by optimal combinations of isotropic and longitudinal mixing. *J Magn Reson B* 1996;**112**(2):160-80.

(4) Bodenhausen G, Ruben DJ. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters* 1980;**69**(1):185-9.

(5) Nicholson JK, Foxall PJ, Spraul M *et al*. 750 MHz 1H and 1H-13C NMR spectroscopy of human blood plasma. *Anal Chem* 1995;**67**(5):793-811.

(6) Fan T, Lane A. Structure-based profiling of metabolites and isotopomers by NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* 2008;**52**(2-3):69-117.

(7) Veselkov KA, Lindon JC, Ebbels TM *et al*. Recursive segment-wise peak alignment of biological (1)h NMR spectra for improved metabolic biomarker recovery. *Anal Chem* 2009;**81**(1):56-66.

(8) Dieterle F, Ross A, Schlotterbeck G *et al*. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem* 2006;**78**(13):4281-90.

(9) Rajalahti T, Kvalheim OM. Multivariate data analysis in pharmaceutics: a tutorial review. *Int J Pharm* 2011;**417**(1-2):280-90.

SUPPLEMENTAL FIGURE/TABLE LEGEND

Table S1: HITChip analysis: relative contribution of phyla, groups, genus and species of bacteria significantly modified by the prebiotic treatment as shown by the Wilcoxon p value calculated between values at the beginning (Preb-1) and the end (Preb-2) of treatment. The corrected p values were obtained after correcting for the false discovery

rate. The results are given as the mean ± SD. Preb-1: prebiotic group, T0; Preb-2: prebiotic group, T3months; Plac-1: placebo group, T0; Plac-2: placebo group, T3months.

Table S2: HITChip analysis: description of the genus-like groups belonging to bacilli, *Clostridium* clusters IV and XVI within the Firmicutes.

Table S3: Glucose and lipid homeostasis in both groups (placebo and treated) before (T0) and after (T3months) treatment. Raw data are given as the mean ± SD. Differential values (Δ) are given as the mean ± 95% confidence intervals. P values according to the Mann-Whitney test performed on differential values to assess treatment effect. OGTT: oral glucose tolerance test; HOMA: homeostasis model assessment; LDL: low-density lipoprotein; HDL: high-density lipoprotein.

Figure S1: Hierarchical clustering of the HITChip profiles of the 58 faecal samples analysed (29 patients at 2 time points). The samples from the same subject at time points 1 and 2 (T0 and T3months) clustered together. The subjects belonging to the prebiotic group are shown with green arrows.

Figure S2: HITChip analysis: RDA plot of samples belonging to the placebo (red and light red) and prebiotic (green and light green) groups at T0 and T3months. The first and second ordination axes are plotted explaining up to 10% of the variability in the dataset.

Figure S3: Partial least square (PLS) regression analysis between urinary metabolic profiles and the waist/hip ratio or post-OGTT insulin. Score plots showing the correlation between metadata (y axis) and PLS scores (x axis).