Supplementary Material (Supplementary Material and Methods and Supplementary Figures) for:

Duodenal *Anaerobutyricum soehngenii* infusion stimulates GLP-1 production, ameliorates glycemic control and beneficially shapes the duodenal transcriptome in metabolic syndrome subjects; a randomized double-blind placebo-controlled cross-over study.

Supplementary Material and Methods

Culturing of *A. soehngenii*

The cells were obtained as described previously[1] by culturing *A. soehngenii L2-7* at 500-liter scale in a basic phosphate-bicarbonate salt medium containing 2% yeast extract, 0.4% soy peptone, and 2% sucrose, at pH 6.8 and 37C. Following autoclaving, filter-sterilized components were added, including cysteine (final concentration 0.05%) and a 1 ml per liter of a vitamin solution (containing per liter 10 mg biotin, 10 mg cobalamin, 30 mg para-aminobenzoic acid, 50 mg folic acid, and 150 mg pyridoxamine). *A. soehngenii* L2-7 cells were harvested by microfiltration, washed with PBS, and finally stored in PBS containing 10% glycerol at a concentration of 10^{10} cells/ml in 10 ml tubes at -80°C. *A. soehngenii* L2-7 was handled under strict anaerobic conditions which were maintained during all stages of the production of the concentrated cells: during growth, microfiltration, glycerol mixing, and filling of the tubes with a nitrogen atmosphere. The viability of *A. soehngenii* L2-7 in randomly selected tubes (stored at -80°C at the AMC Department of Clinical Pharmacy) was tested every 6 months during the study using most probable number (MPN) analysis in YCFA medium. MPN analyses were performed in duplicate in anoxic YCFA medium containing sucrose incubated at 37°C for 5 days[1]. Growth was scored by visual and microscopic inspection. Viability stayed constant at 10^{10} cells/ml during the time of the study.

Duodenal RNA sequencing and differential gene expression analysis

RNA for RNA sequencing analysis was isolated from duodenum biopsies, which were directly snap-frozen in liquid nitrogen after biopsy and stored at -80˚C until analysis, from all 12 included participants, using an RNA isolation protocol optimized for small tissue biopsies. In short, biopsies were mixed with 300 µl TriPure (Roche, Basel, Switzerland) and homogenized on ice using a sterile, RNAse free pestle. After short centrifugation, 60 µl of chloroform was added. Samples were then added to a Heavy Phase Lock gel tube (Quanta Bio, Beverly, USA) and centrifuged (15 min, 12.000 x g, 4˚C). The aqueous phase was transferred and mixed with 1 volume of 70% ethanol. The mixture was added to a RNeasy MinElute spin column (QIAgen, Tegelen, the Netherlands). RNA was washed according to manufacturer's protocol and eluted in 14 µl RNAse free water. RNA concentration was measured using the NanoDrop 1000 (Thermo Scientific, Landsmeer, the Netherlands). RIN scores were assessed on a Bioanalyzer 2100 using Eukaryote Total RNA Nano chips (Agilent Technologies, Santa

Koopen A, *et al. Gut* 2022; 71:1577–1587. doi: 10.1136/gutjnl-2020-323297

Clara, USA). RNA was depleted from rRNA and sequenced on a HiSeq4000 (paired-end, 150 bp) by Genomescan BV, Amsterdam, The Netherlands.

RNA raw sequence quality was checked using FastQC (v0.11.9)[2] and quality trimmed and filtered using Trimmomatic (v0.38)[3]. The single, 50 bp reads were processed by removing the first 5 bases, applying a sliding-window quality trim 4 bp wide with a threshold of Q15, then removing all reads shorter than 36 bp after trimming. Quality checked and trimmed reads were subsequently pseudo aligned to the human transcriptome (GRCh38 release 97) using Kallisto (v0.46.0)[4]. The transcript-level counts from Kallisto output were used to perform differential gene expression analysis using 3 different R packages: sleuth (v0.30.0)[5], DESeq2 (v1.28.1)[6], and edgeR (v3.30.3)[7–10]. Count data was imported to DESeq2 and edgeR using the tximport package[11]. Transcript to gene mappings were obtained using the biomaRt package (v2.44.0)[12]. In all 3 workflows, likelihood ratio tests were applied using '~ Subject + Visit' as the full model design and '~ Subject' as the reduced model in order to detect genes that were differentially expressed after *A. soehngenii L2-7* infusion compared to after placebo infusion. All p-values were adjusted for multiple comparisons using the Benjamini-Hochberg method[13]. Significance thresholds were 0.05 for sleuth and edgeR and 0.10 for DESeq2. Only genes found to be differentially expressed by all 3 workflows were examined in downstream analyses.

Real Time quantitative polymerase chain reaction (RT-qPCR)

Total RNA was extracted from frozen duodenal biopsies from 12 patients as described above. Briefly, 1μg of RNA was converted to cDNA with iScript cDNA synthesis kit (BioRad, Veenendaal, The Netherlands). qPCR was performed on a ViiA7 PCR machine (Applied Biosystems, Bleiswijk, The Netherlands). using Sybr Green Fast (Bioline Meridian Bioscience, Cincinnati, Ohio, USA). Gene expression was normalized towards the housekeeping gene Actin, and relative gene expression was calculated with the "delta delta Ct" method and shown as 2ˆ-delta delta Ct. Primer sequences are outlined in Supplementary Table 2; all primers were manufactured by Sigma-Aldrich (Zwijndrecht, The Netherlands).

Western blotting

Duodenal biopsies (from 8/9 patients) were lysated in RIPA buffer (Thermo Fisher Scientific, Breda, The Netherlands) containing protease and phosphatase inhibitors (cOmplete™ Protease inhibitor and PhosSTOP Phosphatase Inhibitor Cocktails, Sigma, Zwijndrecht, The Netherlands) using a ceramic beads homogenizer. For westernblotting of cell lysates, Caco-2 cells were incubated for 30 minutes 4˚C in RIPA buffer supplemented with protease and phosphatase inhibitors. BCA protein assay kit (Thermo Fisher) was used to determine protein concentrations.β-mercaptoethanol was added as reducing agent to all sample lysates, which were run on 4-12% polyacrylamide gels (BioRad, GE, Boston, USA) in MES running buffer. Proteins were transferred to PVDF membranes (BioRad) and were blocked using 5% milk in TBS-T (Tris Buffered Saline – Tween-20). Membranes were incubated

overnight a 4˚C with primary polyclonal rabbit antibodies anti-Reg1B (for Figure 6A/6G: 1:500, E-AB-52897, Elabscience, Houston, USA) and Actin (AB306371, Abcam, Cambridge, UK) or anti –Reg1A (for Figure S5G: 1:500, orb100720, Biorbyt, Cambridge, UK). Horseradish peroxidase (HRP)-conjugated secondary antibodies (R&D systems, Minneapolis, USA) were incubated for 1hour at room temperature. HRP activity was visualized with peroxidase substrate for enhanced chemiluminescence and imaged with ChemiDoc MP Imaging System (BioRad) using Image Lab software (BioRad). Densitometric quantification analysis was performed using the Image J software. All protein levels were normalized to the loading control (β-actin). Reg1B protein expression shown as ratio of densitometric quantification of Reg1B versus β-actin and as intraindividual fold-changes of treatment versus placebo (expression rates of placebo group normalized to mean to represent distribution in expression levels among subjects). In order to show lack of cross-reactivity of the antibodies anti-Reg1B and anti-Reg1A, recombinant Reg1A (Prospec) and Reg1B (Sino Biological) proteins were loaded on polyacrylamide gels and immunoblotted using antibodies anti-Reg1A (orb100720, Biorbyt) or Reg1B (E-AB-52897, Elabscience) (Figure S5I,K).

Immunohistochemical staining

Formalin-fixed paraffin-embedded (FFPE) duodenal 4 μm sections were utilized for immunohistochemical staining. Slides were deparaffinized in 100% Xylene and rehydrated in ethanol (100%, 96% and 70%) and H2O, following by block of endogenous peroxidase in 3% H2O2 methanol for 20 minutes and heat-induced epitope retrieval (HIER) in citrate buffer pH 6.0 at 98°C for 10 minutes in the Thermo Scientific PT Module. After incubation for 10 minutes with Ultravision protein block (Thermo Fisher Scientific, Breda, The Netherlands), FFPE sections were incubated with anti-Reg1B primary antibody (for Figure 6D,E: 1:100 dilution in TBS, MA5-29517, Invitrogen, Waltham, Massachusetts, USA) for 1 hour at room temperature (RT), following by incubation with the secondary antibodies BrightVision Poly-HRP-conjugates goat anti-rabbit IgG (undiluted) for 30 minutes. For assessing the expression and localization of Reg1B in duodenal tissue with different antibodies and to compare its expression to the one of Reg1A (Figure S5H), following 10-minutes HIER in citrate buffer, duodenal sections were incubated with primary antibodies against Reg1B (1:2000, E-AB-52897, Elabscience, Houston, USA) or against Reg1A (1:2000, orb100720, Biorbyt, Cambridge, UK). All single-stainings were visualized with 3,3'Diaminobenzidine (DAB) kit (Sigma Aldrich, Zwijndrecht, The Nederlands).

For the triple staining of REG1B, lysozyme and mucin were stained in sequential order using sequentially cut duodenal FFPE sections. Reg1B expression was visualized by: 10-minute HIER in citrate buffer pH 6.0 at 98°C, 1-hour incubation at RT with anti-Reg1B rabbit IgG (Invitrogen, MA5-29517, 1:100 dilution in TBS), 30-minute incubation with BrightVision Poly-alkaline phosphatase (AP)-conjugated goat anti-rabbit IgG (undiluted) and staining development with Perma Red/AP kit (Diagnostic BioSystem, Pleasanton, California, USA). Lysozyme (as marker of Paneth cells) was stained

Koopen A, *et al. Gut* 2022; 71:1577–1587. doi: 10.1136/gutjnl-2020-323297

following: 5-minute HIER in Tris-EDTA pH 9.0 buffer at 98°C, 30-minute incubation at RT with polyclonal rabbit anti-lysozyme (1:2000 dilution in TBS, Dako EC 3.2.1.17, Agilent Technologies, Santa Clara, California, USA), 30-minute incubation with BrightVision Poly-HRP-conjugated goat anti-rabbit IgG (1:2 dilution in TBS) and staining development with Perma Yellow/HRP kit (Diagnostic BioSystem). For the final detection of acidic mucins (to mark Goblet cells), FFPE double-stained slides were incubated with Alcian Blue solution (1% in 3% acetic acid, pH 2.5, Sigma Aldrich) for 5 minutes and eventually covered with coverslips using VectaMount mounting medium (Thermo Fisher Scientific, H-5000).

Cell culture and assay procedure

Caco-2 cells were cultured in Gibco Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 IU/ml penicillin, 100 IU/ml streptomycin, and 2mM glutamine (Thermo Fisher Scientific) in T75 flasks. The day prior to the stimulation assays, cells were seeded at $1x10^5$/well in a 12-well plates; after resting overnight, cells were exposed to 1mM butyrate or 1 µg/ml muramyl dipeptide (MDP) (both stock solutions diluted in water) for 6 hours. Afterwards cells were washed once in PBS and lysated in RIPA buffer. Reg1B concentrations were determined by ELISA (Cloud-Clone Corp., Uscn Life Science Kit Inc., Wuhan, China) accordingly to the manufacturer's instructions in Caco-2 cells exposed to increasing concentrations of *A. soehngenii* L2-7 cells (10^5/ml and 10^6/ml). Prior to use in cell culture, bacteria were heat-inactivated at 65°C for 20 minutes. BCA protein assay kit (Thermo Fisher) was used to assess protein concentrations.

Measurements of fecal SCFA and plasma SCFA, incretins, bile acids and Reg1B

Fecal SCFAs (butyrate, acetate, propionate) were measured in morning stool samples (N=11), directly frozen at -20°C after collection, using gas chromatography coupled to tandem mass spectrometry detection (GC-MS/MS) as described previously[14]. Briefly, approximately 20-100 mg of fecal samples were mixed with internal standards, added to glass vials and freeze dried. All samples were then acidified with HCl, and SCFAs were extracted with two rounds of diethyl ether extraction. The organic supernatant was collected, the derivatization agent N-tert-butyldimethylsilyl-N-methyltrifluoroacetamide (Sigma-Aldrich, Stockholm, Sweden) was added and samples were incubated at room temperature overnight. SCFAs were quantified with a gas chromatograph (Agilent Technologies 7890A, Santa Clara, California, USA) coupled to a mass spectrometer (Agilent Technologies 5975C). Short chain fatty acid standards were attained from Sigma-Aldrich (Stockholm, Sweden).

Plasma SCFA (butyrate, acetate, propionate) were measured at Cleveland Clinic (OH, USA) in heparin plasma samples (N=12), directly frozen at -80-°C after collection, using gas chromatography coupled to TANDEM mass spectrometry (GC-MS/MS) as previously described[15]. Briefly, 30 µl aliquots of plasma were mixed with 50 µl 2-Butanol/Pyridine (3:2) and 5 µl containing the heavy labeled internal

standards. Afterwards, the carboxylic acids were derivatized by mixing 50 µl supernatant with 10 µl isobutyl chloroformate, followed by vortexing and sonicating the mixture. After derivatization, 50 µl hexane were added and mixed; following centrifugation, the top hexane layer was removed for GC-MS/MS analysis and 1 µl was injected into GC column. The quantitation of butyric acid, acetic acid, and propionic acid was performed using isotope dilution GC-MS/MS and the absolute concentration of each SCFA was determined using calibrations curves measured for each analyte. Samples were analyzed on a Thermo TSQ-Evo triple quadrupole mass spectrometer interfaced with the Trace 1310 gas chromatograph (Thermo Fisher Scientific). Chromatographic separation was achieved by using an HP-5MS fused-silica capillary column (30 m × 0.250 mm × 0.25 µm; Agilent Technologies, Santa Clara, CA, USA) coated with 5% phenymethyl siloxane as previously described[15]. The mass spectrometer was used in MRM mode with the following parent to daughter ion transitions: m/z 61.0 → 43.0 for acetic acid, m/z 63.0 → 45.0 for [13C2]-acetic acid, m/z 61.0 → 43.0 m/z 71.0 → 41.0 for butyric acid, m/z 78.1 → 46.1 for D7-butyric acid, m/z 75.1 → 57.0 for propionic acid, m/z 77.1 → 59.0 for D2-propionic acid.

Plasma incretin levels of all 12 individuals were determined in postprandial (2-hour mixed meal test) samples as previously described[16]. Plasma concentrations of GIP (total) and GLP-1 (total) were measured by Holst group with ELISA (cat no. 10-1258-01 and 10-1278-01, Mercodia, Sweden). All quality controls provided by the manufacturer were within allowed limits. All samples from the same individual were measured in the same assay run.

Concentrations of the secondary bile acids tauro-omega-muricholic acid (TOMCA), tauroursodeoxycholic acid (TUDCA), taurodeoxycholic acid (TDCA), tauroursodeoxycholic acid (TUDCA), taurolithocholic acid (TLCA), glycohyodeoxycholic acid (GHDCA), glycodeoxycholic acid (GDCA), glycoursodeoxycholic acid (GUDCA), glycolithocholic acid (GLCA), omega-muricholic acid (OMCA), deoxycholic acid (DCA), ursodeoxycholic acid (UDCA), lithocholic acid (LCA), hyodeoxycholic acid (HDCA), murocholic acid (MuroCA), iso-ursodeoxycholic acid (IsoUDCA) were measured in plasma samples from all 12 participants by means of ultra-performance liquid chromatography-tandem mass spectrometry (UPLCMS/MS), as previously performed[17]. Briefly, samples (50µl) were extracted with 10 volumes of methanol containing deuterated internal standards (d_4-TCA, d_4-GCA, d_4-GCDCA, d_4-GUDCA, d_4-GLCA, d_4-UDCA, d_4-CDCA, d_4-LCA; 50nM of each). After 10 minutes of vortex and 10 minutes of centrifugation at 20 000g, the supernatant was evaporated under a stream of nitrogen and reconstituted in 200µl methanol:water [1:1]. The samples were injected (5µl) and bile acids were separated on a C18 column (1.7µ, 2.1 x 100mm; Kinetex, Phenomenex, USA) using water with 7.5mM ammonium acetate and 0.019% formic acid (mobile phase A) and acetonitrile with 0.1% formic acid (mobile phase B). The chromatographic separation started with 1 minute isocratic separation at 20%B. The B-phase was then increased to 35% during 4 minutes. During the next 10 minutes the B-phase was increased to 100%. The B-phase was held at 100% for 3.5 minutes before

Koopen A, *et al. Gut* 2022; 71:1577–1587. doi: 10.1136/gutjnl-2020-323297

returning to 20%. The total runtime was 20 minutes. Bile acids were detected using multiple reaction monitoring (MRM) in negative mode on a QTRAP 5500 mass spectrometer (Sciex, Concord, Canada) and quantification was made using external standard curves.

Concentrations of Reg1B in plasma samples and duodenal tissue lysates was assessed by ELISA (Cloud-Clone Corp., Uscn Life Science Kit Inc., Wuhan, China) accordingly to the manufacturer's instructions.

Strain-specific qPCR

The DNA concentrations were determined fluorometrically (Qubit dsDNA HS assay; Invitrogen) and adjusted to 1 ng/µl prior to use as the template in qPCR. Primers targeting 16S rRNA gene of *A. soehngenii* L2-7 EhalF (5'GCGTAGGTGGCAGTGCAA) and EhalR (5'GCACCGRAGCCTATACGG) (Ramirez et al. 2008) were used for quantification. Standard template DNA was prepared from the 16S rRNA gene of *A. soehngenii* L2-7 by amplification with primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3'). Standard curves were prepared with nine standard concentrations of 100 to 10^8 gene copies/µl. PCRs were performed in triplicate with iQ SYBR Green Supermix (Bio-Rad) in a total volume of 10 µl with primers at 500 nM in 384-well plates sealed with optical sealing tape. Amplification was performed with an iCycler (Bio-Rad, USA) with the following protocol: one cycle of 95°C for 10 min; 40 cycles of 95°C for 15 s, 55°C for 20 s, and 72°C for 30 s each; one cycle of 95°C for 1 min, one cycle of 60°C for 1 min, and a stepwise increase of the temperature from 60 to 95°C (at 0.5°C per 5 s) to obtain melt curve data. Data were analysed using the Bio-Rad CFX Manager 3.0.

Fecal 16S rRNA gene amplicon sequencing and bioinformatics

DNA extraction from fecal samples from 11 patients was performed using the repeated bead beating protocol as previously described[18]. DNA was eluted in 50 µl of DNAse- RNAse-free water and its concentration and quality were evaluated using NanoDrop 2000 spectrophotometry. Subsequently, DNA was diluted to reach a concentration of 20 ng/µl which served as template for PCR. The V5-V6 region of 16S ribosomal RNA (rRNA) gene was amplified in duplicate PCR reactions for each sample in a total reaction volume of 50 µl using a master mix containing 1 µl of a unique barcoded primer, 784F-n and 1064R-n (10 µM each per reaction), 1 µl dNTPs mixture, 0.5 µl Phusion Green Hot Start II High-Fidelity DNA Polymerase (2 U/µl; Thermo Scientific, Landsmeer, The Netherlands), 10 µl 5× Phusion Green HF Buffer, and 36.5 µl DNAse- RNAse-free water[19]. The amplification program included 30 seconds (s) of initial denaturation step at 98°C, followed by 25 cycles of denaturation at 98°C for 10 s, annealing at 42°C for 10 s, elongation at 72°C for 10 s, and a final extension step at 72°C for 7 minutes. The PCR product was visualized on 1% agarose gel (~280 bp) and purified with CleanPCR kit (CleanNA, Alphen aan den Rijn, The Netherlands). The concentration of the purified PCR product was measured with Qubit dsDNA BR Assay Kit (Invitrogen, California, USA) and 200 ng of microbial

DNA from each sample were pooled for the creation of the final amplicon library which was sequenced (150 bp, paired-end) on the Illumina HiSeq 2500 platform (GATC Biotech, Constance, Germany).

Raw reads were demultiplexed using the Je software suite (v2.0)[20] allowing no mismatches in the barcodes. After removing the barcodes, linker and primers, reads were mapped against the human genome using bowtie2 (v2.4.1)[21] in order to remove human reads. Surviving microbial forward and reverse reads were pipelined separately using DADA2 (v1.12.1)[22]. Amplicon Sequence Variants (AVSs) inferred from the reverse reads were reverse-complemented and matched against ASVs inferred from the forwards reads. Only non-chimeric forward reads ASVs that matched reverse-complemented reverse reads ASVs were kept. ASV sample counts were inferred from the forward reads. ASV taxonomy was assigned using DADA2 and the SILVA (v132) database. The resulting ASV table and taxonomy assignments were integrated using the phyloseq R package (v1.28.0)[23]. ASVs sequences were aligned using MAFFT (v.7.427)[24] using the auto settings. A phylogenetic tree was constructed from the resulting multiple sequence alignment with FastTree (v.2.1.11 Double Precision)[25] using a generalized time-reversible model ('-gtr'). Biopsy samples were rarefied to 24947 counts per sample, while fecal samples were rarefied to 13229 counts per sample. The vegan R package (v2.5.6)[26] was used to calculate alpha-diversity metrics (Shannon index and ASV richness) and Bray-Curtis dissimilarities. Weighted-Unifrac distances were calculated using the phyloseq package.

Power calculation and statistical analyses

We based our power calculation on the study of Van Baarlen *et al.*[27], in which a striking difference in duodenal mucosal transcriptomic profiling was reported 6 hours after introduction of a single *Lactobacillus* bacterial strain. Based on 60% decrease in duodenal *Fxr* gene expression upon *A. soehngenii* L2-7 administration to db/db mice, compared to placebo[18], with one sample Chi^2 test, the sample size in each group needed to be 12 in order to have a group proportion of 0.5 and with a comparison proportion of 0.1. Mean absolute deviation of glucose (MAD), continuously measured using FreeStyle Libre technology, was calculated using the default 'mad' function from R stats package[28]. Wilcoxon signed rank tests and Mann-Whitney U-tests were used to compare within-group changes of related samples and to compare intervention groups. Student's t-test was used for analyzing differences in groups from *in vitro* experiments. Area under the curve (AUCs) for MMT measurement were calculated using the DescTools package (v0.99.36) [29,30]. Correlation plots were made using the corrplot package (v0.84) [31]. The mixOmics package (v6.12.1) [32] was used to perform multilevel PCA analyses. Principal Coordinate Analyses (PCoA) were performed using the ape package (v5.4)[33]. All other statistical analyses and visualizations were performed in R (v4.0.1)[34] using the tidyverse (v1.3.0)[35] and ggplot2 package (v3.3.1)[30]. P values <0.05 were considered statistically significant.

SI References

1. Gilijamse, P. W. *et al.* Treatment with Anaerobutyricum soehngenii: a pilot study of safety and dose–response effects on glucose metabolism in human subjects with metabolic syndrome. *npj Biofilms Microbiomes* **6**, 1–10 (2020).

2. Andrews, S. FastQC: a quality control tool for high throughput sequence data. (2015).

3. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

4. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).

5. Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat. Methods* **14**, 687–690 (2017).

6. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

7. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

8. Robinson, M., Mccarthy, D., Chen, Y., Lun, A. & Smyth, G. K. edgeR: differential expression analysis of digital gene expression data. (2012).

9. Fernandes, A. D. *et al.* Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome* **2**, 15 (2014).

10. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).

11. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2015).

12. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184 (2009).

13. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 289–300 (1995).

14. Wichmann, A. *et al.* Microbial modulation of energy availability in the colon regulates intestinal transit. *Cell Host Microbe* **14**, 582–590 (2013).

15. Arnon D Lieber, AD. et al. Loss of HDAC6 alters gut microbiota and worsens obesity. FASEB J 33(1), 1098-1109 (2019).

16. Kootte, R. S. *et al.* Improvement of Insulin Sensitivity after Lean Donor Feces in Metabolic Syndrome Is Driven by Baseline Intestinal Microbiota Composition. *Cell Metab.* (2017). doi:10.1016/j.cmet.2017.09.008

17. Tremaroli, V. *et al.* Roux-en-Y Gastric Bypass and Vertical Banded Gastroplasty Induce Long-Term Changes on the Human Gut Microbiome Contributing to Fat Mass Regulation. *Cell Metab.* **22**, 228–238 (2015).

18. Salonen, A. *et al.* Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: Effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J.*

Microbiol. Methods **81**, 127–134 (2010).

19. Ramiro-Garcia, J. *et al.* NG-Tax, a highly accurate and validated pipeline for analysis of 16S rRNA amplicons from complex biomes. *F1000Research* **5**, 1791 (2016).

20. Girardot, C., Scholtalbers, J., Sauer, S., Su, S.-Y. & Furlong, E. E. M. Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics* **17**, 419 (2016).

21. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357 (2012).

22. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Meth* **13**, 581–583 (2016).

23. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**, e61217 (2013).

24. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–66 (2002).

25. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).

26. Oksanen, J. *et al.* vegan: Community Ecology Package. (2017).

27. van Baarlen, P. *et al.* Differential NF- B pathways induction by Lactobacillus plantarum in the duodenum of healthy humans correlating with immune tolerance. *Proc. Natl. Acad. Sci.* **106**, 2371–2376 (2009).

28. Fokkert, M. J. *et al.* Performance of the freestyle libre flash glucose monitoring system in patients with type 1 and 2 diabetes mellitus. *BMJ Open Diabetes Res. Care* **5**, 1–8 (2017).

29. Signorell, A. DescTools: Tools for Descriptive Statistics. (2020).

30. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).

31. Wei, T. & Simko, V. R package 'corrplot': visualization of a correlation matrix (version 0.84).'. *Retrived from https//github. com/taiyun/corrplot* (2017).

32. Cao, K.-A. Le, Rohart, F., Gonzalez, I. & Dejean, S. mixOmics: Omics Data Integration Project, package version 6.1.2 https://CRAN.R-project.org/package=mixOmics. (2017).

33. Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

34. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. (2016).

35. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).

SI Legends of Figures and Tables

Figure S1: Fecal *A. soehngenii* L2-7 levels

Fecal *A. soehngenii* L2-7 levels determined by qPCR at 0, 24 hours, 1 week, 2 weeks after placebo/treatment-interventions. Values indicate gene copies per gr of feces.

Figure S2: Fecal short-chain fatty acids

(A) Concentrations (nmol/mg dried feces weight) of butyrate, (**B**) acetate, and (**C**) propionate in morning stool samples obtained at baseline and 1 day after placebo/treatment-intervention.

Figure S3: Postprandial glucose and insulin
(**A**) Plasma glucose levels (mmol/l) at 0, 20, 30, 120 minutes during mixed meal test (MMT). (**B**) Plasma glucose levels during MMT as total area under the curve (AUC). (**C**) Plasma insulin levels (pmol/l) at 0, 20, 30, 120 minutes during MMT. (**D**) Plasma insulin levels during MMT as total area under the curve (AUC).

Figure S4: Plasma short-chain fatty acids

(A) Concentrations (μM) of butyrate, (**B**) acetate, and (**C**) propionate in plasma at 120 minutes of mixed meal test (MMT).

Figure S5: Duodenal expression of REG genes and Reg1A/1B proteins

Gene expression measured by qPCR in duodenal biopsies at 6 hours post-intervention of (**A**) *REG1B* (different set of forward and reverse primers than in Figure 6B), (**B**) *REG3A*, (**C**) *REG3G*, and (**D**) *REG4*. (**A-D**) Data showing the relative gene expression (to placebo) using the 2ˆ-(delta delta Ct) method. (**E**) Quantification of Reg1A expression levels in duodenal biopsies, Reg1A expression normalized to β-actin (loading control). (**F**) Reg1A expression shown as fold-change treatment versus placebo. (**G**) Westernblot images of duodenal lysates blotted with antibodies against Reg1A and β-actin. (**H**) Immunohistochemical staining of Reg1B (different antibodies used than in Figures 6D,6E) and Reg1A in duodenal biopsies. (**I,K**) Westernblot images showing the specificity of the antibodies against Reg1B and Reg1A; (**I**) Westernblotting for Reg1B: enhanced band intensity with increasing amount of loaded recombinant human (rh)Reg1B and absence of a band when rh Reg1A is loaded; ; (**K**) Westernblotting for Reg1A: stronger band intensity with increasing amount of loaded rh Reg1A and absence of a band when rh Reg1B is loaded. (**J**) Circulating levels (ng/ml) of Reg1B measured by ELISA in plasma samples taken at 8 hours post-intervention. (**L**) Reg1b expression by Caco-2 cells in response to exposure to increasing concentrations of heat-inactivated *A. soehngenii* L2-7 cells.

Figure S6: Assessment of carry-over affect between week 0 and week 4

(**A**) Fecal *A. soehngenii* L2-7 levels determined by qPCR at baseline (week 0 and week 4); values indicate gene copies per gr of feces. (**B**) Concentrations (nmol/mg dried feces weight) of butyrate in morning stool samples obtained at baseline (week 0 and week 4). (**C**) Plasma GLP-1 levels during mixed meal test (MMT) shown as total area under the curve (AUC). (**D**) Median absolute deviation (MAD) of continuous glucose measurements (CGM) over the first 24 hours after placebo/treatment-intervention.

Koopen A, *et al. Gut* 2022; 71:1577–1587. doi: 10.1136/gutjnl-2020-323297

Table S1: Baseline characteristics and safety parameters at both study visits

Data expressed as medians and interquartile ranges. There were no differences after the *A. Soehngenii* L2-7 infusion compared to the placebo infusion. BMI: body mass index, HOMA-IR: homeostatic model assessment of insulin resistance, HbA1c: glycated hemoglobin, HDL: high-density lipoprotein, LDL: low-density lipoprotein; AST: aspartate transaminase, ALT: alanine transaminase, AP: alkaline phosphatase; γgt: gamma-glutamyltransferase; CRP: c-reactive protein.

Table S2

Primer sequences utilized in the analysis of duodenal gene expression. *: primers used in Figure S5A.

Koopen A, *et al. Gut* 2022; 71:1577–1587. doi: 10.1136/gutjnl-2020-323297

Table S1

	Placebo	A. Soehngenii	P - value
Weight (kg)	110.1 [100.0 – 120.2]	110.4 [101.3 – 122.0]	0.812
BMI (kg/m2)	33.4 [32.2 – 38.]	33.7 [31.8 – 37.8]	0.859
Blood pressure: systolic (mmHg)	135 [129 – 147]	144 [131 – 156]	0.258
Blood pressure: diastolic (mmHg)	89 [84 – 91)	91 [79 – 95)	0.917
Fasting glucose (mmol/L)	5.3 [5.0 – 5.6]	5.1 [4.8 – 5.7]	0.430
Insulin (pmol/L)	65 [48 – 83]	56 [34 – 81]	0.099
HOMA - IR	2.3 [1.7 – 2.8]	1.9 [1.1 – 2.8]	0.158
HbA1c (mmol/mol)	37 [36 – 39]	37 [36 – 39]	0.942
Cholesterol: total (mmol/L)	5.29 [4.57 – 6.26]	5.13 [4.71 – 6.08]	0.255
Cholesterol: HDL (mmol/L)	1.17 [0.91 – 1.31]	1.13 [0.97 – 1.32]	0.929
Cholesterol: LDL (mmol/L)	3.05 [2.69 - 3.90]	3.21 [2.79 – 3.83]	0.638
Cholesterol: triglycerides (mmol/L)	1.65 [1.12 – 2.92]	1.66 [1.19 – 2.31]	0.433
Creatinine (umol/L)	87 [78 – 92]	84 [79 – 93]	0.342
AST (U/L)	25 [19 – 31]	26 [23 – 33]	0.109
ALT (U/L)	25 [21 – 36]	27 [22 – 35]	0.124
AP (U/L)	70 [62 – 85]	72 [60 -87]	0.783
yGT (U/L)	39 [26 – 52]	31 [26 – 49]	0.254
CRP (mg/ml)	2.8 [1.8 – 5.0]	3.0 [2.3 – 5.3]	0.477
Leukocytes (10^e9/L)	6.2 [5.6 – 6.7]	6.5 [5.9 – 7.3]	0.130
Caloric intake (kcal/day)	1843 [1607 – 2090]	1888 [1667 – 2041]	0.424
Fat intake (g)	68 [57 – 75]	61 [56 – 81]	0.790
Protein intake (g)	76 [65 – 85]	77 [60 – 91]	0.333
Carbohydrate intake (g)	201 [162 – 233]	228 [173 – 254]	0.241
Fiber intake (g)	18 [14 – 19]	19 [16 – 21]	0.339

Koopen A, *et al. Gut* 2022; 71:1577–1587. doi: 10.1136/gutjnl-2020-323297

Table S2

Gene	Forward primer sequence	Reverse primer sequence
GPR43/FFAR2	TGCTACGAGAACTTCACCGAT	GGAGAGCATGATCCACACAAAAC
TGR5	CACTGTTGTCCCTCCTCTCC	ACACTGCTTTGGCTGCTTG
FXR	TACATGCGAAGAAAGTGTCAAGA	ACTGTCTTCATTCACGGTCTGAT
FGF19	CGGAGGAAGACTGTGCTTTCG	CTCGGATCGGTACACATTGTAG
OSTalpha	CTGGGCTCCATTGCCATCTT	CACGGCATAAAACGAGGTGAT
REG1B	GGTCCCTGGTCTCCTACAAG	TCCATTTCTTGAATCCTGAGCA
REG1A	GGTCCCTGGTCTCCTACAAG	CATTTCTGGAATCCTGTGCTTG
REG1B*	AGTAGTGGGTCCCTGGTCTC	TGAATCCTGAGCATGAAGTCA
REG3A	AGCTACTCATACGTCTGGATTGG	CACCTCAGAAATGCTGTGCTT
REG3G	GGTGAGGAGCATTAGTAACAGC	CCAGGGTTTAAGATGGTGGAGG
REG4	CTGCTCCTATTGCTGAGCTG	GGACTTGTGGTAAAACCATCCAG
ACTB	CCAACCGCGAGAAGATGA	CCAGAGGCGTACAGGGATAG