Online MATERIALS AND METHODS

Dysosmobacter spp. quantification by qPCR.

The qPCR target DNA region was the 16S rRNA gene. To design specific primers, the 16S rRNA gene sequences of *Dysosmobacter welbionis* J115^T and 32 closely related species or strains (highest hits during Blast), cultivated or not, was collected from public databases (NCBI accession numbers are ASM343475v1, ASM309653v1, ASM343476v1, ASM70166v1, ASM62094v1, ASM258694v1, Clos_orbi_1_3_50AFAA_V1, GCA_001405195.1, ASM215986v1, ASM222164v1, ASM145494v1, PRJEB27980, PRJEB20323, ASM216167v1, ASM216123v1, ASM216043v1, ASM216030v1, ASM216027v1, PRJEB18258, PRJEB17348, ASM76523v1, KF447772.2, LT598575.1, ASM347899v1, ASM216123v1, CP048437.1, AB238598.1, JF750939.1, AB040496.1, AB040497.1, AB040495.1, MG963288). Multiple alignment of the sequences was performed using MUSCLE [1], Distances were computed using the maximum composite likelihood method and the phylogenetic trees were reconstructed using maximum-likelihood method in MEGA 7.0 after gaps and unknown bases were eliminated [2]. The tree was reconstructed with 1000 bootstrap replications (Supplemental figure S8). Then, we determined the positions of the bases that differed between *D. welbionis* J115^T 16S rRNA sequence and the 32 other 16S rRNA gene sequences and designed primers on these regions using Primer-Blast [3]. The specificity of the obtained primers was first checked by performing Blast on NCBI Nucleotide collection (nr/nt), then by checking the absence of amplification in qPCR using *Oscillibacter valericigenes* Sjm18-20 and *O. ruminantium* GH1 genomic DNA as template. The fragment between the positions 184 and 239 of the 16S rRNA gene is amplified and the sequences of the primers are ***Dysosmobacter* spp:** Forward: ATGACGCATGACGCATGACC, Reverse:

Le Roy T, *et al. Gut* 2022; 71:534–543. doi: 10.1136/gutjnl-2020-323778

CCAGCGATAAAATCTTTGACATGCC. Genomic DNA was extracted from mouse faeces, mouse caecal content or human stools using the QIAamp DNA Stool Mini Kit (Qiagen, Germany), including a bead-beating step. Quantified standard DNA for *Dysosmobacter* spp. qPCR was obtained by extracting genomic DNA with the same DNA extraction protocol than for the faeces/stools from a *Dysosmobacter welbionis* J115^T culture in exponential growing phase of known concentration in colony forming units (cfu) determined by plating. DNA concentration was determined, and purity (A260/A280) was checked using a NanoDrop2000 (Thermo Fisher Scientific, USA). Samples were diluted to an end concentration of 10 and 0.1 ng/µl in TE buffer pH 8. Total bacteria qPCR was performed on the 0.1 ng/µl dilution and the *Dysosmobacter* spp. qPCR was performed on the 10 ng/µl dilution. Real-time PCRs were performed with the StepOnePlus real-time PCR system and software (Applied Biosystems, Den Ijssel, The Netherlands) using Mesa Fast qPCR sybr green mix (Eurogentec, Seraing, Belgium) for detection according to the manufacturer's instructions. A standard curve was included on each plate by diluting genomic DNA from pure culture. *Dysosmobacter* spp. Standard curve ranged between 3.2 10^3 and 1.0 10^7 cfu per well for *Dysosmobacter* spp. and between 6.4 10^2 and 2.0 10^6 cfu per well for total bacteria. Bacteria quantity in caecal content and faeces was expressed as percentage of total bacteria and as cfu per gram of faeces.

Primer sequences for total bacteria are ***Total Bacteria*** Forward: ACTCCTACGGGAGGCAGCAG Reverse: ATTACCGCGGCTGCTGG.

Culture and preparation of *Dysosmobacter welbionis* for mice experiments. *D. welbionis* J115^T was cultured anaerobically in a modified YCFA medium supplemented with 10 g/L inositol. Cultures were centrifuged at 5000g during 20 min and the supernatant was removed. For fresh bacterial suspension preparation, cells were

resuspended in anaerobic PBS-carbonate buffer and immediately administered to the mice by oral gavage. For frozen bacterial suspension preparation, cells were resuspended in anaerobic PBS-carbonate buffer supplemented with 15 % (vol/vol) trehalose then immediately frozen in anaerobic vials and stored at -80°C. The number of total and cultivable bacteria administered to the mice was calculated by plating the bacterial culture before preparation and the bacterial suspension after preparation for mice administration.

Mice experiments. Sets of 10-week-old C57BL/6J male mice (Janvier Laboratories, Le Genest-Saint-Isle, France) were housed in pairs in SOPF (specific opportunistic and pathogen free) conditions and in a controlled environment (room temperature of 22 ± 2°C, 12h daylight cycle) with free access to sterile food (irradiated) and sterile water. Upon arrival, all mice underwent a 1-week acclimatization period, during which they were fed a control diet (ND) (AIN93Mi, Research Diet, New Brunswick, NJ, USA). During the experiments, food and water intake were recorded once a week. Body composition was assessed by using a 7.5 MHz time domain-nuclear magnetic resonance machine (TD-NMR) (LF50 minispec, Bruker, Rheinstetten, Germany). In the experiment 6, male homozygous *ob/ob* mice (B6.V-Lepob/ob/JRj) were used as genetically obese model (leptin-deficient mice with severe obesity). Mice were purchased at the age of 6 weeks (Janvier Laboratories, Le Genest-Saint-Isle, France) and housed in the same condition as described here above.

Experiment 1. The purpose of this first mice experiment was to assess the impact of freshly prepared, daily cultured *D. welbionis* J115^T on diet-induced obesity. A set of 30 mice was divided in 3 groups of 10 mice. The mice were fed a control diet (AIN93Mi; Research diet, New Brunswick, NJ, USA) or a high-fat diet (60% fat and 20%

Le Roy T, *et al. Gut* 2022; 71:534–543. doi: 10.1136/gutjnl-2020-323778

carbohydrates (kcal/100g), D12492, Research diet, New Brunswick, NJ, USA). One group of HFD-fed mice was treated with an oral administration of daily prepared fresh cultures of *Dysosmobacter welbionis* J115^T by oral gavage at the dose 1.0 10^9 cfu /0.2 ml of cultivable, live bacteria per day and per mice (HFD J115 fresh) and control groups were treated with an oral gavage of an equivalent volume of PBS-carbonate buffer. Treatment continued for 6 weeks. Mice were euthanized after a 6-hour fasting period.

Experiment 2. The aim of this second experiment was to investigate whether heat-killed *D. welbionis* J115^T was still effective on diet-induced obesity. Moreover, we used live *D. welbionis* J115^T not coming from a fresh culture but frozen in anaerobic PBS-carbonate buffer supplemented with 15 % (weight/vol) of the cryoprotectant trehalose. A set of 30 mice was divided in 3 groups of 10 mice. The mice were fed a high-fat diet. One group of HFD-fed mice was treated with live *D. welbionis* J115^T frozen in trehalose by oral gavage at the dose of 1.0 10^9 cfu /0.2 ml of cultivable, live bacteria per day and per mice (HFD Live J115). One group of HFD-fed mice was treated with 1.10^9 cells/0.2 ml of *D. welbionis* J115^T pasteurised by heating procedure (30 min at 70°C) (HFD pasteurised J115), the control HFD groups was treated with an oral gavage of an equivalent volume of PBS-carbonate buffer supplemented with 15 % (weight/vol) trehalose (HFD). Mice have been treated for 10 weeks. Mice were euthanized after a 6-hour fasting period.

Experiment 3. The aim of this third experiment was to confirm the impact of live *D. welbionis* J115^T on diet-induced obesity obtained during experiment 1 and to determine the impact of longer period of treatment (i.e., 13 weeks) using the frozen bacterial preparation. A set of 36 mice was divided in 3 groups of 12 mice. The mice were fed a control diet or a high-fat diet. One group of HFD-fed mice was treated with live *D. welbionis* J115^T frozen in anaerobic PBS-carbonate buffer supplemented with 15 %

trehalose (weight/vol); each mouse received 1.0 10^9 cfu /0.2 ml of cultivable, live bacteria per day (HFD Live J115). HFD and control groups were given an oral gavage of an equivalent volume of PBS-carbonate buffer supplemented with 15 % (weight/vol) trehalose (Control and HFD). At week 10, faeces were collected in the morning, that is to say 16hour after last *D. welbionis* administration. Then, bacteria administration was discontinued for three days, then faeces were collected again, and daily bacteria administration resumed until the end of the experiment. Mice were euthanized after a 6-hour fasting period.

Experiment 4. The purpose of this fourth experiment was to measure the temperature of the mice after a shorter treatment than during experiments 1 and 2. A set of 14 mice was divided in two groups of 7 mice that were housed individually, with free access to food and water. The mice were fed a high-fat diet (HFD). A group of mice was treated by oral gavage with live *D. welbionis* J115^T frozen in anaerobic PBS-carbonate buffer supplemented with 15 % trehalose (weight/vol); each mouse received 1.0 10^9 cfu /0.2 ml of cultivable, live bacteria per day (HFD Live J115). HFD groups was treated with an oral gavage of an equivalent volume of PBS-carbonate buffer supplemented with and 15 % trehalose (HFD). Treatment continued for 3 weeks. On the last day of the experiment, the temperature of the mice was measured using a rodent rectal thermometer. Mice were euthanized in the morning with no fasting period.

Experiment 5. The purpose of this experiment was to assess if gavaging with lower numbers of *D. welbionis was as effective.* A set of 40 mice was divided in 3 groups of mice (n=12 control, n=14 HFD, n=14 HFD Live J115). The mice were fed a control diet or a high-fat diet. One group of HFD-fed mice was treated for 13 weeks with live *D. welbionis* J115^T frozen in anaerobic PBS-carbonate buffer supplemented with 15 % trehalose (weight/vol); each mouse received 1.0 10^8 cfu /0.2 ml of cultivable, live

Le Roy T, *et al. Gut* 2022; 71:534–543. doi: 10.1136/gutjnl-2020-323778

bacteria per day (HFD Live J115). This experiment led us to discover that a low abundance of *D. welbionis* J115^T that is <1.0 10^9 cfu /0.2 ml of cultivable, live bacteria per day is not associated with any improvement of the metabolism (Supplemental figure S2).

Experiment 6. The aim of this sixth experiment was to study the impact of live *D. welbionis* J115^T in a model of severe obesity, therefore, the genetically induced obese *ob/ob* mice (B6.V-Lepob/ob/JRj) have been used. The mice were fed a control diet. One group was treated with live *D. welbionis* J115^T frozen in anaerobic PBS-carbonate buffer supplemented with 15 % trehalose (weight/vol); each mouse received 1.0 10^9 cfu /0.2 ml of cultivable, live bacteria per day (Ob Live J115) (n = 11), the control group was treated with an oral gavage of an equivalent volume of PBS-carbonate buffer supplemented with and 15 % trehalose (Ob CT) (n = 10). The treatment continued for 6 weeks (Supplemental figure S7).

Oral glucose tolerance test (OGTT). One week before the end of experiment 3, the mice were fasted for 6 hours before being given an oral gavage glucose load (2 g glucose per kg body weight). Blood glucose was measured 30 minutes before (time point -30), just prior the oral glucose load (time point 0) and then after 15, 30, 60, 90 and 120 min. Blood glucose was determined with a glucose meter (Accu Check, Roche, Switzerland) on blood samples collected from the tip of the tail vein. Plasma insulin concentration was determined using an ELISA kit (Mercodia, Uppsala, Sweden) according to the manufacturer's instructions.

Tissue sampling. The animals were anesthetized with isoflurane (Forene®, Abbott, Queenborough, Kent, England) and blood was collected from the portal and cava veins. Then mice were immediately euthanized by cervical dislocation. Tissue samples (liver, brown adipose tissue, subcutaneous adipose tissue, mesenteric adipose tissue,

jejunum, ileum and proximal colon, muscles) were dissected, immersed in liquid nitrogen, and stored at –80 °C for further analysis. A part of the adipose tissues and intestines was fixed in 4 % paraformaldehyde in PBS for histological analysis.

Short chain fatty acids assay

Portal plasma (50µL) and caecal content (50mg) were used short chain fatty acids determination by high performance liquid chromatography associated to mass spectrometry (HPLC-MS). Briefly, samples were added to 200 µL of acetonitrile containing valproic acid (used as internal standard). Tubes were mixed vigorously and incubated in an ultrasound water bath (10 min, 4°C) before being incubated 1 h at -20°C to allow protein precipitation. Following centrifugation (10 min, 13000 g) supernatants were transferred in glass tubes for the derivatization reaction. 100 µL of 3-nitrophenylhydrazine solution (0.14 M, in acetonitrile-water (1:1, v/v)) and 200 µL of an EDC solution (0.06 M, in acetonitrile-water (1:1, v/v)) containing 3% (v/v) of pyridine were added to the supernatants. After incubation (1 h, 40°C) under regular mixing, the tubes were placed in ice-cold water, and water (750 µL containing 50 µL of 2 N HCl) and chloroform (1.5 mL) were added to extract the derivatized SCFA. Following removal of the organic layer, the samples were reconstituted in methanol for the HPLC-MS analysis using a LTQ-Orbitrap mass spectrometer coupled to an Accela HPLC (ThermoFischer Scientific). Separation was achieved on a Hypersil GOLD PFP (100x2.1 mm; 1.9 µm) column maintained at 40°C. A gradient (200 µL/min) using water-acetonitrile-acetic acid 94.9:5:0.1 (v/v/v) and acetonitrile-acetic acid 99.9:0.1 (v/v) allowed for the separation of the different SCFA. An APCI source operated in positive mode was used for the MS analysis. For the quantification, calibration curves were obtained in the same conditions using increasing amount of SCFA and a constant

Le Roy T, *et al. Gut* 2022; 71:534–543. doi: 10.1136/gutjnl-2020-323778

amount of internal standard. SCFA levels in the caecal content were normalized to the dry weight.

Caecal microbiota analysis by 16S rRNA gene sequencing. Genomic DNA extracted from the caecal content of the mice (extracted as described in *Dysosmobacter* spp. quantification by qPCR section) was diluted at 20 ng/µl. Sequencing was performed at MR DNA (www.mrdnalab.com, Shallowater, TX) on a MiSeq instrument following the manufacturer's guidelines. The V1-V3 region was amplified from purified DNA with the primers ill27Fmod (AGRGTTTGATCMTGGCTCAG) and ill519Rmod (GTNTTACNGCGGCKGCTG), using 30 amplification cycles with an annealing temperature of 65°C. Because MiSeq sequencing enables paired 300-bp reads, the ends of each read overlap and can be stitched together to generate extremely high-quality, full-length reads covering the entire V1-V3 region. The resulting PCR products were purified and loaded onto the Illumina MiSeq cartridge, according to the manufacturer instructions. The sequences were analysed using FROGS pipeline [4] implemented on a galaxy instance (http://migale.jouy.inra.fr/galaxy/). Briefly, reads were demultiplexed and clustered. After dereplication, the clusterisation tool ran with SWARM [5]. Chimeras were then removed using VSEARCH [6] and sequences were filtered to keep clusters (also called OTUs) accounting for at least 0.005% of all sequences. Taxonomic affiliation was performed with both RDP Classifier and Blastn+ against the SILVA 132 pintail 100 database [7]. The average number of sequences per sample was 40 418 ± 7 203 sequences. Then, we normalized the dataset to the number of sequences of the sample with the lowest sequencing depth, that is to say 26 881 sequences, using Rhea script without random subsampling [8].

Human cohort analyses

To gain insight into the prevalence and relative abundance of *Dysosmobacter welbionis* across a human population, data of the human microbiome project (HMP) was investigated (The Human Microbiome Project Consortium, 2012) [9]. Already trimmed 16S rRNA gene sequence data generated from faeces samples of the healthy human subjects (HHS) cohort was downloaded from the human microbiome project (HMP) data portal (https://portal.hmpdacc.org/). The 16S data set was split based on the amplified region of the 16S gene, the V3-V5 data subset was chosen for further analysis as it contained the most samples (N=216). The USEARCH v11 tool was used for 16S analysis [10, 11]. Reads were globally trimmed to 350 bp and denoised using the unoise3 command with default parameters [12]. Original reads were mapped to the resulting set of ZOTUs (zero-radius operational taxonomic units) to obtain a count table that was converted into the standard BIOM format [13]. The ZOTU table was imported in R version 3.6.3 using the phyloseq 1.30.0 package and further analysed. Samples with sequence library sizes within the first quartile (9057.5) were omitted as these were deemed too low (N=161). Taxonomy of the ZOTU sequences was predicted using SINTAX [11], with a bootstrap cutoff value of 0.8 and the Ribosomal Database Project training set v16 with the reference *Dysosmobacter welbionis* J115^T 16S gene sequence (Genbank accession MG963288.1) added at the database. Two ZOTU sequences with exact matches to the reference *Dysosmobacter welbionis* J115^T 16S gene sequence were retrieved. The relative abundance of these ZOTUs was summed to obtain a better estimate of the relative abundance of *D. welbionis.* For the FGFP the sequence variant was also corresponding to 100% homology to the 16SrRNA sequence on its full length (232 bp) with 0 gaps introduced. The sequencing methods used the V4 region of the 16S rRNA amplified by using the 515F/806R primer pair

Le Roy T, *et al. Gut* 2022; 71:534–543. doi: 10.1136/gutjnl-2020-323778

(GTGYCAGCMGCCGCGGTAA and GGACTACNVGGGTWTCTAAT). The DADA2 pipeline was used to partition rarefaction counts (10,000) into taxonomic units, and as extensively detailed here [14]

For the American Gut Project (AGP), the relative abundance sOTU table obtained by Deblur processing on 16S V4 sequence data, following trimming to 125 bp and bloom removal, was downloaded in the standard BIOM format [15]. The table contained 32,954 sOTUs and 9,511 samples. The sOTU sequences were extracted from the table and BLASTn was performed with the *Dysosmobacter welbionis* J115^T MG963288.1 gene sequence. One sOTU sequence with exact match to the reference *Dysosmobacter welbionis* J115^T 16S gene sequence was retrieved

Histological analyses. Subcutaneous adipose tissue (SAT), mesenteric adipose tissue (MAT), brown adipose tissue (BAT) depots and jejunal tissue were fixed in 4 % paraformaldehyde for 24 h at room temperature. Samples were then dehydrated by immersion in ethanol 100% for 24 h and processed for paraffin embedding. Paraffin sections of 5 µm were stained with haematoxylin and eosin. Images were obtained using a SCN400 slide scanner and Digital Image Hub software (Leica Biosystems, Wetzlar, Germany). Adipocytes size and distribution in white adipose tissues were calculated from 3 fields per sample for mesenteric adipose tissue and from 5 fields per sample for subcutaneous adipose tissue using Fiji and Adiposoft softwares [16]. White area in brown adipose tissue corresponds to the lipid droplets and was quantified from 5 fields per sample using Fiji software. Crypts and villi length in the jejunum were measured using Digital Image Hub from Slidepath (Leica) on an average of 20 crypts/villus (10 crypts/villus on 2 different slices) per sample.

Transit time measurement. Mice were force-fed with 200 µl of carmine red dissolved in drinking water at a concentration of 10 mg/ml. The transit time was the time elapsed between the force-feeding and the emission of the first red pellet.

Energy intestinal absorption. Six weeks after the beginning of the experiment, the faeces of two 7-days periods were collected. During the same time the food intake was monitored (the food was weighed before and after the faeces collection period). The faeces were dried at 60°C for 2 hours and weighted. Total energy of the diet and the faeces was determined by bomb calorimetry (C1, IKA, USA). The net intestinal absorption is calculated based on the ingested and excreted energy and represented the proportion of ingested energy that was not recovered in faeces output.

Fasting concentration of hormones. Circulating leptin, resistin, glucose-dependent insulinotropic polypeptide (GIP) and plasminogen activator inhibitor-1 (PAI-1) concentrations were determined using a multiplex immunoassay kit (Mouse diabetes assay, Bio-Plex Pro, Bio-Rad, Belgium) and measured using Luminex technology (Bioplex, Bio-Rad, Belgium) according to manufacturer's instructions.

Gene expression analysis by real-time qPCR analysis and RNAseq analysis. Total RNA was prepared from tissues using TriPure reagent (Roche). Quantification and integrity analysis of total RNA was performed by running 1 µl of each sample on an Agilent 2100 Bioanalyzer (Agilent RNA 6000 Nano Kit, Agilent).

For qPCR analysis, cDNA was prepared by reverse transcription of 1 µg total RNA using a Reverse Transcription System kit (Promega, Leiden, The Netherlands). Real-time PCRs were performed with the StepOnePlus real-time PCR system and software (Applied Biosystems, Den Ijssel, The Netherlands) using Mesa Fast qPCR sybr green mix (Eurogentec, Seraing, Belgium) for detection according to the manufacturer’s

instructions. RPL19 was chosen as housekeeping gene. All samples were run in duplicate in a single 96-well reaction plate, and data were analysed according to the $2^{-\Delta\Delta Ct}$ method. The identity and purity of the amplified product was checked through analysis of the melting curve carried out at the end of amplification. Primer sequences for the targeted mouse genes are available in Supplemental Table S3.

For RNAseq analysis, the integrity of RNA from brown adipose tissue (BAT) was determined using the Agilent bioanalyzer 2100 system with the RNA 6000 Nano LabChip kit. BAT RNA samples with an RNA integrity number inferior to 8 on a scale ranging from 0 to 10 were eliminated. Then, samples were pooled for each group (HFD pool and HFD-J115 pool) to an end concentration of 50 ng/µl. The samples were then sequenced by Eurofins Genomics, which consisted in purification of poly-A containing mRNA molecules, then mRNA fragmentation, random primed cDNA synthesis (strand specific), adapter ligation and adapter specific PCR amplification and finally paired-end Illumina sequencing with a read length of 2 x 150 bp. Eighty to 90 million read pairs were obtained and analysed on Galaxy server using RNA-STAR and htseq modules.

Citrate synthase activity. Around 10 mg of brown adipose tissue was weighed and lysed in 20 volumes of CelLytic MT Cell Lysis containing 1 % (vol/vol) of protease inhibitor cocktail P8340 (Sigma) by bead-beating. The lysate was centrifuged at 10 000g during 10 min at 4°C two times in order to remove the lipids and the tissue debris. Tissue extract was diluted 1:10 in a 100mM phosphate buffer (pH 7.1) containing 10 mM 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) and 30 mM acetyl-CoA. After the addition of 10 mM oxaloacetate, free coenzyme A produced from the condensation of acetyl-CoA and oxaloacetate was bound to DTNB, and resulting change in light absorbance detected spectrophotometrically at 412 nm was used to determine the activity of citrate synthase (µmol/mg/s).

Statistical analyses. Statistical analyses were performed using GraphPad Prism version 7.00 for Windows (GraphPad Software, San Diego, CA, USA) and R 3.5.1. Comparison between two groups was performed by Mann–Whitney–Wilcoxon test, comparison between three or more groups on one time-point was performed by one-way ANOVA followed by Tukey correction and comparison between 3 or more groups at different time-points was performed by 2-way repeated measures ANOVA. Probability value was referred as p value for uncorrected tests and as value for tests followed by a correction. p or q < 0.05 was considered statistically significant. Principal coordinates analyses (PCoA) were performed using R program and Rhea scripts. The cladogram generator GraPhIAn was used for 16S rRNA gene sequencing data visualization. Data from the Microbes4U cohort deemed not normal using the Shapiro test were log transformed. We used R program and the Hmisc and corrplot packages to produce Pearson correlations matrix.

References

1 Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004;**32**:1792-7.
2 Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 2016;**33**:1870-4.
3 Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC bioinformatics 2012;**13**:134.
4 Escudie F, Auer L, Bernard M, Mariadassou M, Cauquil L, Vidal K*, et al.* FROGS: Find, Rapidly, OTUs with Galaxy Solution. Bioinformatics 2018;**34**:1287-94.
5 Mahe F, Rognes T, Quince C, de Vargas C, Dunthorn M. Swarm: robust and fast clustering method for amplicon-based studies. PeerJ 2014;**2**:e593.
6 Rognes T, Flouri T, Nichols B, Quince C, Mahe F. VSEARCH: a versatile open source tool for metagenomics. PeerJ 2016;**4**:e2584.
7 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P*, et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013;**41**:D590-6.
8 Lagkouvardos I, Fischer S, Kumar N, Clavel T. Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. PeerJ 2017;**5**:e2836.
9 Consortium THMP. Structure, function and diversity of the healthy human microbiome. Nature 2012;**486**:207-14.

10 Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;**26**:2460-1.
11 Edgar RC. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. bioRxiv 2016:074161.
12 Edgar RC. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. bioRxiv 2016:081257.
13 McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D*, et al.* The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. Gigascience 2012;**1**:7.
14 Hughes DA, Bacigalupe R, Wang J, Ruhlemann MC, Tito RY, Falony G*, et al.* Genome-wide associations of human gut microbiome variation and implications for causal inference analyses. Nat Microbiol 2020;**5**:1079-87.
15 McDonald D, Hyde E, Debelius JW, Morton JT, Gonzalez A, Ackermann G*, et al.* American Gut: an Open Platform for Citizen Science Microbiome Research. mSystems 2018;**3**.
16 Galarraga M, Campion J, Munoz-Barrutia A, Boque N, Moreno H, Martinez JA*, et al.* Adiposoft: automated software for the analysis of white adipose tissue cellularity in histological sections. J Lipid Res 2012;**53**:2791-6.