Cell Metabolism, Volume *22*

Supplemental Information

Roux-en-Y Gastric Bypass and Vertical Banded Gastroplasty Induce Long-Term Changes on the Human Gut Microbiome Contributing to Fat Mass Regulation

Valentina Tremaroli, Fredrik Karlsson, Malin Werling, Marcus Ståhlman, Petia Kovatcheva-Datchary, Torsten Olbers, Lars Fändriks, Carel W. le Roux, Jens Nielsen, and Fredrik Bäckhed

Supplemental Experimental Procedures

Metagenomic Data Quality Control

Sequence reads were trimmed from the 3' end with Fastx quality trimmer using quality threshold of 20. Read pairs with either read shorter than 35 bp were removed. Read pairs aligning to the human genome (NCBI version 37) were removed using bowtie with settings "-n 2 -l 35 -e 200 --best -p 8 --chunkmbs 1024 -X 600 –tryhard". The set of high-quality reads was used in further analysis.

Alignment to Reference Genomes and Taxonomical Analysis

A reference catalogue of 2382 reference genomes from the NCBI and HMP databases was collected (Table S2). The reference genomes were split into two Bowtie indexes. High-quality reads were aligned to each index using Bowtie (Langmead et al., 2009) with the following parameters: -n 2 -l 35 -e 300 --best -p 8 --chunkmbs 1024 -X 600 –tryhard. The results from the two indexes were merged by selecting the alignment with fewest mismatches and more than 90% identities; if a read was aligned to a reference genome with the same number of mismatches, each genome was assigned 1/2 read. The relative abundance of each genome was calculated by summing the number of reads aligned to that genome divided by the total number of reads and scaled by the genome size. In each subject, the relative abundance was scaled to sum to one. Taxonomic information for each genome was downloaded from the NCBI taxonomy database. By summing the relative abundance of child ranks, the relative abundance of taxonomic parent ranks was calculated.

Alignment to Reference Gene Catalogue

The human gut microbial gene catalogue generated by MetaHIT (Qin et al., 2010) was downloaded and indexed with Bowtie2 (Langmead and Salzberg, 2012). The metagenomic sequence reads were aligned with Bowtie2 using the default parameters and aligning reads with at least 90% sequence identity were counted. The provided annotation of the genes to KEGG was used and the aligning reads to each KO was summed.

Statistical Analysis

The statistical analysis was performed in the software R (R Core Team, 2012). The relative abundance of species and genera between groups was tested with Wilcoxon rank-sum test. P values were corrected for multiple testing with the method from Benjamini and Hochberg (Benjamini and Hochberg, 1995). Species with a relative abundance above 10^{-5} in any subject were included in the analysis.

For the functional analysis, the R package edgeR was used, which uses the read counts to assess the differential abundance of KOs between groups. The read counts for each KO were loaded into R and the dispersion was calculated with the functions estimateGLMCommonDisp(), estimateGLMTrendedDisp() and estimateGLMTagwiseDisp(). For further analysis, the tagwise dispersion was used in the glmFit() function. The individual comparisons between groups were

performed with the glmLRT() function by pairwise comparing the groups. P values were adjusted with the method from Benjamini and Hochberg (Benjamini and Hochberg, 1995). In the analysis using edgeR, if not stated otherwise, default parameters were used.
To analyze the results from the abundance of KOs in the context of pathways, the KO differential adjusted P values and direction of change were analyzed together with the KO pathway membership using the Reporter Feature algorithm as implemented in the R package Piano (Varemo et al., 2013). The null distribution was used as significance method.
Bile salt hydrolases annotated to the K01442 in the MetaHIT gene catalogue were extracted and annotated to the two types of bile salt hydrolases, Bsh1 and Bsh2, described by Fang et al. (Fang et al., 2009). Sequences for *bsh1* and *bsh2* genes were downloaded from NCBI Genebank with accessions 221062093, 221062123, 221062125, 221062129, 221062139 as representatives for bsh1 and 221062141, 221062143 as representatives for bsh2. MetaHIT genes with annotation to K01442 were aligned to the catalogue of *bsh* genes and the best hits with an E-value score below 10-5 were annotated as either *bsh1* or *bsh2*, for a total of 469 sequences. The pathway for 7α-dehydroxylation of bile acids is responsible for the dehydroxylation of primary bile acids and formation of secondary bile acids by the gut microbiota, and is encoded by the polycistronic bile acid-inducible (*bai*) operon, which contains genes for 8 enzymes (*baiA*, *baiB*, *baiCD*, *baiE*, *baiF*, *baiG*, *baiH* and *baiI*) (Ridlon et al., 2006). This pathway is not completely reconstructed in KEGG and therefore the genes in the MetaHIT gene catalogue were functionally annotated to enzymes in this pathway. 16 protein sequences from two organisms known to carry this pathway, Clostridium hylemonae DSM 15053 (Ridlon et al., 2010) and Clostridium scindens VPI 12708 (Kitahara et al., 2000) were downloaded from the ENA database at EMBL-EBI (http://www.ebi.ac.uk/ena/). The MetaHIT genes were aligned to the 7α-dehydroxylation sequences using UBLAST (Edgar, 2010) with an E-value cutoff of 10^{-5}. The candidates were then aligned to the full version of KEGG including the Bai proteins and only genes with a best hit to Bai proteins in this database were re-annotated. In total 70 MetaHIT genes were re-annotated to *bai* genes. *bsh* and *bai* gene abundances were tested between groups together with KEGG KOs as described above.

Measurement of SCFAs and BCFAs

GC-MS was used for measurement of organic acids in RYGB, VBG and OBS fecal samples. 70-230 mg of frozen fecal material were transferred to glass tubes (16 x 125 mm) fitted with a screw cap, and a volume of 100 µl of internal standards stock solution ([1-13C]acetate and [2H6]propionate 1 M, [13C4]butyrate 0.5 M, [1-13C1]isobutyrate and [1-13C]isovalerate 0.1 M) was added to the tubes. Prior to extraction samples were freeze-dried at -50 ºC for 3 h (yield 25 - 75 mg dry weight). After acidification with 50 µl of 37% HCl, the organic acids were extracted twice in 2 ml of diethyl ether. A 500 µl aliquot of the extracted sample was mixed together with 50 µl of N-tert-butyldimethylsilyl-N-methyltrifluoracetamide (MTBSTFA; Sigma) at room temperature. 1 µl of the resulting derivatized material was injected into a gas chromatograph (Agilent Technologies 7890 A) coupled to a mass spectrometer detector (Agilent Technologies 5975 C). A linear temperature gradient was used: the initial temperature of 65 ºC was held for 6

minutes, increased to 260 ºC (15 ºC*min-1) and then to 280 ºC for 5 minutes. The injector and transfer line temperatures were 250 ºC. Quantitation was completed in selected ion monitoring acquisition mode by comparison to labelled internal standards, and the following m/z ratios were: 117 (acetic acid), 131 (propionic acid), 145 (butyric acid), 146 (isobutyric acid), 159 (isovaleric acid), 121 ([2H2]- and [1-13C]acetate), 136 ([2H5]propionate), 146 ([1-13C1]isobutyrate), 149 ([13C4] butyrate), 160 ([1-13C]isovalerate).

Measurement of TMAO, choline and betaine

TMAO was extracted from 25 µl of plasma with 250 µl of methanol containing 250 nM of d9-TMAO (Cambridge Isotopes Laboratories, Andover, MA, USA) as internal standard. After 10 minutes of vigorous vortexing and 10 minutes of centrifugation at 20000 g, a 50 µl aliquot of the supernatant was diluted with 450 µl of acetonitrile:methanol (3:1). A 5 µl aliquot was then injected onto a UPLC-MS/MS system and TMAO was separated on a BEH HILIC column (2.1 x 100 mm with 1.7 µm particle size) (Waters, Milford, MA, USA), with mobile phases consisting of acetonitrile with 0.2% formic acid (A) and water with 10 mM ammonium acetate and 0.2% formic acid (B). The gradient started with 1 minute of isocratic elution with 15% B. B was increased to 70% over the next 5 minutes, and held at 70% for another 2 minutes. The mobile phase composition was returned to 15% B and the column was equilibrated for 3 minutes to give a total runtime of 10 minutes per sample. The flow rate was 400 µl/min. The detection was made using a QTRAP 5500 instrument (AB Sciex, Toronto Canada) running in positive multiple-reaction monitoring (MRM) mode. The transitions that were monitored were m/z 76 to m/z 58 for the endogenous TMAO and m/z 85 to m/z 66 for the internal standard. TMAO was quantified using an external standard curve with TMAO obtained from Sigma-Aldrich (Stockholm, Sweden).

Betaine and choline were extracted and diluted with the same method used to extract TMAO. A 5 µl aliquot was analyzed using the same instrumental set-up and the same BEH HILIC column as for TMAO but with mobile phases and gradient according to a previously published method (Kirsch et al., 2010). Betaine and choline were quantified using an external standard curve with betaine and choline obtained from Sigma-Aldrich (Stockholm, Sweden).

Measurement of bile acids and FGF19

The extraction and analysis of bile acids were based on a previously published method for the targeted determination of bile acids by UPLC-MS (Garcia-Canaveras et al., 2014). Bile acids were extracted from 50 µl of plasma using protein precipitation with 10 volumes of internal standard-containing methanol. After vortexing and centrifugation, the supernatant was evaporated and reconstituted in 200 µl of methanol:water (1:1). Bile acids (5µl injected) were separated using water with 7.5 mM ammonium acetate and 0.019% formic acid at a pH 4.5 as mobile phase A, and acetonitrile with 0.1% formic acid as phase B. The separation was made using gradient elution on a Kinetex C18 column (2.1x100 mm with 1.7 µm particles) (Phenomenex, Torrance, CA, USA) kept at 60 ºC. The gradient started with 1 minute of isocratic elution with 25% B, which was increased to 35% over the next 4 minutes. During the next 9.5 minutes B increased

from 35% to 95%. After one minute of isocratic elution the gradient was quickly returned to 25% B and the column was equilibrated for 2.5 min to give a total runtime of 18 minutes per sample. The flow rate was 400 µl/min. Detection was made on a QTRAP 5500 instrument (Sciex, Toronto, Canada) with MRM in negative mode. Bile acid standards were obtained from Sigma-Aldrich (Sweden, TCA, TUDCA, TCDCA, TDCA, TLCA, GCA, GCDCA, GDCA, GLCA, CA, UDCA, CDCA, DCA, LCA), CDN isotopes (Quebec, Canada, d4GCDCA, d4GLCA, d4GCA, d4GUDCA, d4CDCA, d4UDCA, d4LCA) and Toronto Research Chemicals (Downsview, Ontario, Canada, d4TCA).
Plasma FGF19 concentrations were measured by quantitative enzyme-linked immunosorbent assay with the human FGF19 Quantikine ELISA kit (R&D Systems, UK) according to the manufacturer's instructions.

Mouse Experiments

For microbiota transplantations approximately 500 mg of frozen stool samples were cut in an anaerobic Coy chamber and suspended in 5 ml of phosphate-buffered saline containing 0.2 g/l Na2S and 0.5 g/l cysteine as reducing agents. Mice were colonized by oral gavage of 200 µl of fecal slurry. Mice that received fecal microbiota from the same individual were housed in the same cage in the isocage system (Tecniplast, Italy).

Supplemental References

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met *57*, 289-300.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460-2461.

Fang, F., Li, Y., Bumann, M., Raftis, E.J., Casey, P.G., Cooney, J.C., Walsh, M.A., and O'Toole, P.W. (2009). Allelic variation of bile salt hydrolase genes in Lactobacillus salivarius does not determine bile resistance levels. Journal of bacteriology *191*, 5743-5757.

Garcia-Canaveras, J.C., Donato, M.T., and Lahoz, A. (2014). Ultra-performance liquid chromatography-mass spectrometry targeted profiling of bile acids: application to serum, liver tissue, and cultured cells of different species. Methods in molecular biology *1198*, 233-247.

Kirsch, S.H., Herrmann, W., Rabagny, Y., and Obeid, R. (2010). Quantification of acetylcholine, choline, betaine, and dimethylglycine in human plasma and urine using stable-isotope dilution ultra performance liquid chromatography-tandem mass spectrometry. Journal of chromatography B, Analytical technologies in the biomedical and life sciences *878*, 3338-3344.

Kitahara, M., Takamine, F., Imamura, T., and Benno, Y. (2000). Assignment of Eubacterium sp. VPI 12708 and related strains with high bile acid 7alpha-dehydroxylating activity to Clostridium scindens and proposal of Clostridium hylemonae sp. nov., isolated from human faeces. International journal of systematic and evolutionary microbiology *50 Pt 3*, 971-978.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357-359.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology *10*, R25.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T.*, et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59-65.

R Core Team (2012). R: A Language and Environment for Statistical Computing (Vienna, Austria).

Ridlon, J.M., Kang, D.J., and Hylemon, P.B. (2006). Bile salt biotransformations by human intestinal bacteria. Journal of lipid research *47*, 241-259.

Ridlon, J.M., Kang, D.J., and Hylemon, P.B. (2010). Isolation and characterization of a bile acid inducible 7alpha-dehydroxylating operon in Clostridium hylemonae TN271. Anaerobe *16*, 137-146.

Varemo, L., Nielsen, J., and Nookaew, I. (2013). Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. Nucleic Acids Res *41*, 4378-4391.

Supplemental Figures and Legends

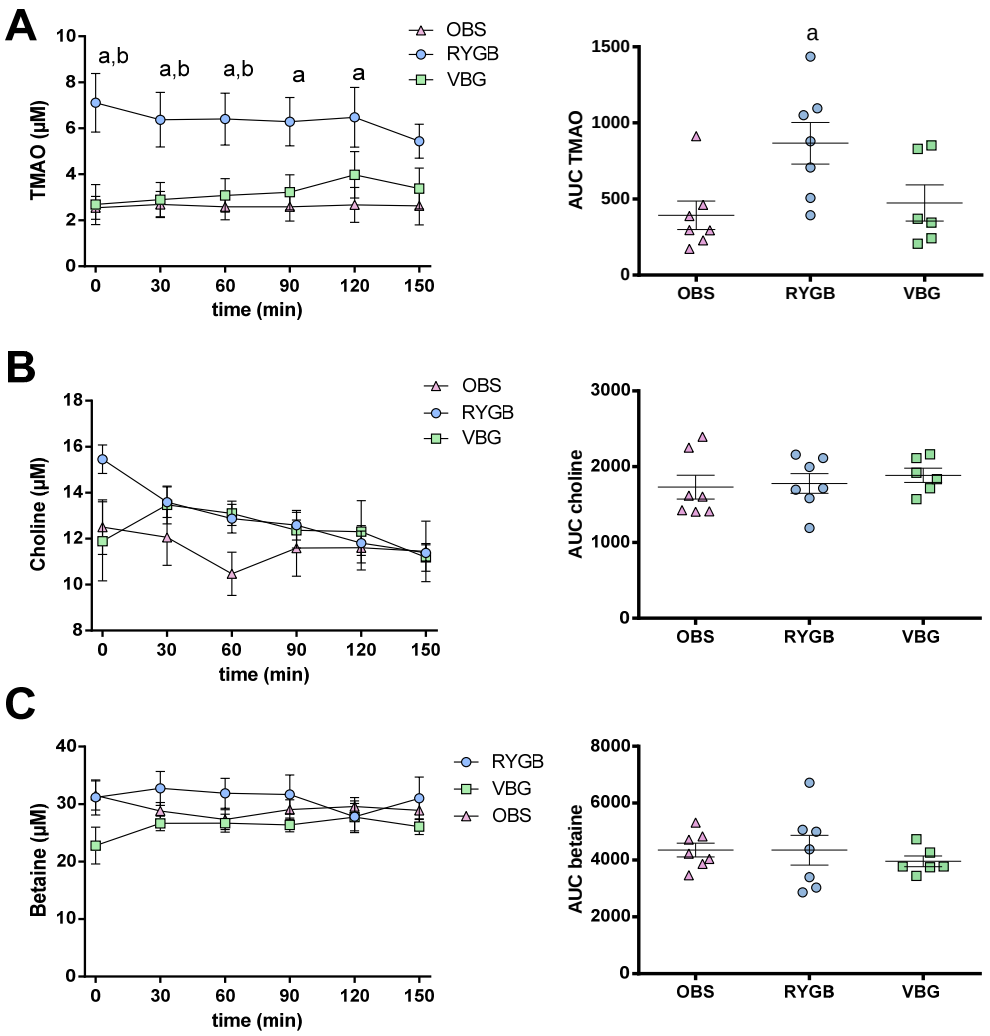


Figure S1. Measurement of TMAO, Choline and Betaine in the Plasma of RYGB, VBG and OBS Women. Related to Figure 2 and Table 1.

(A) Circulating levels of TMAO, (B) choline and (C) betaine. The concentrations of TMAO, choline and betaine were measured during fasting and every 30 minutes for 2.5 hours after a standard meal. Plasma samples for one woman in the VBG group could not be obtained, so the results represent the mean ± standard error of the mean for 7 RYGB, 6 VBG and 7 OBS women. [a] $P<0.05$ for RYGB compared to OBS and [b] $P<0.05$ for RYGB compared to VBG according to one-way ANOVA with Tukey's correction for multiple comparisons.

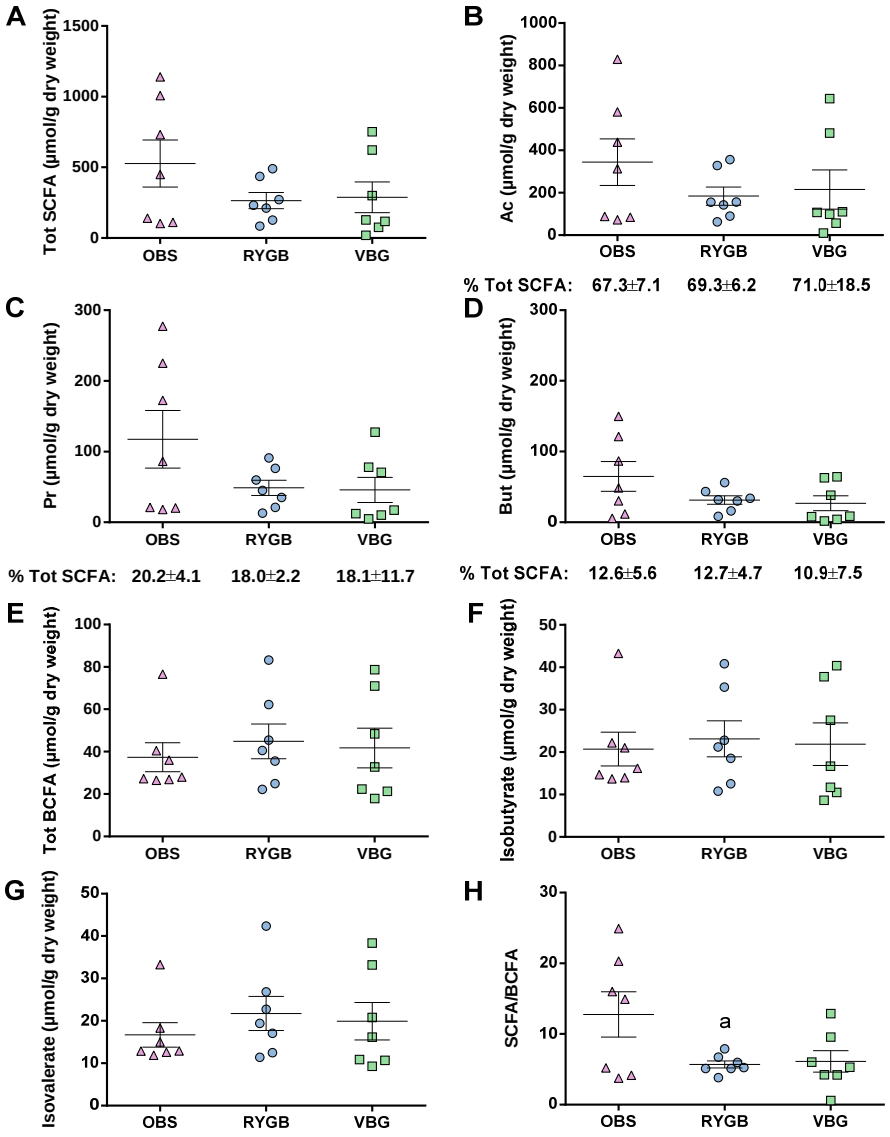
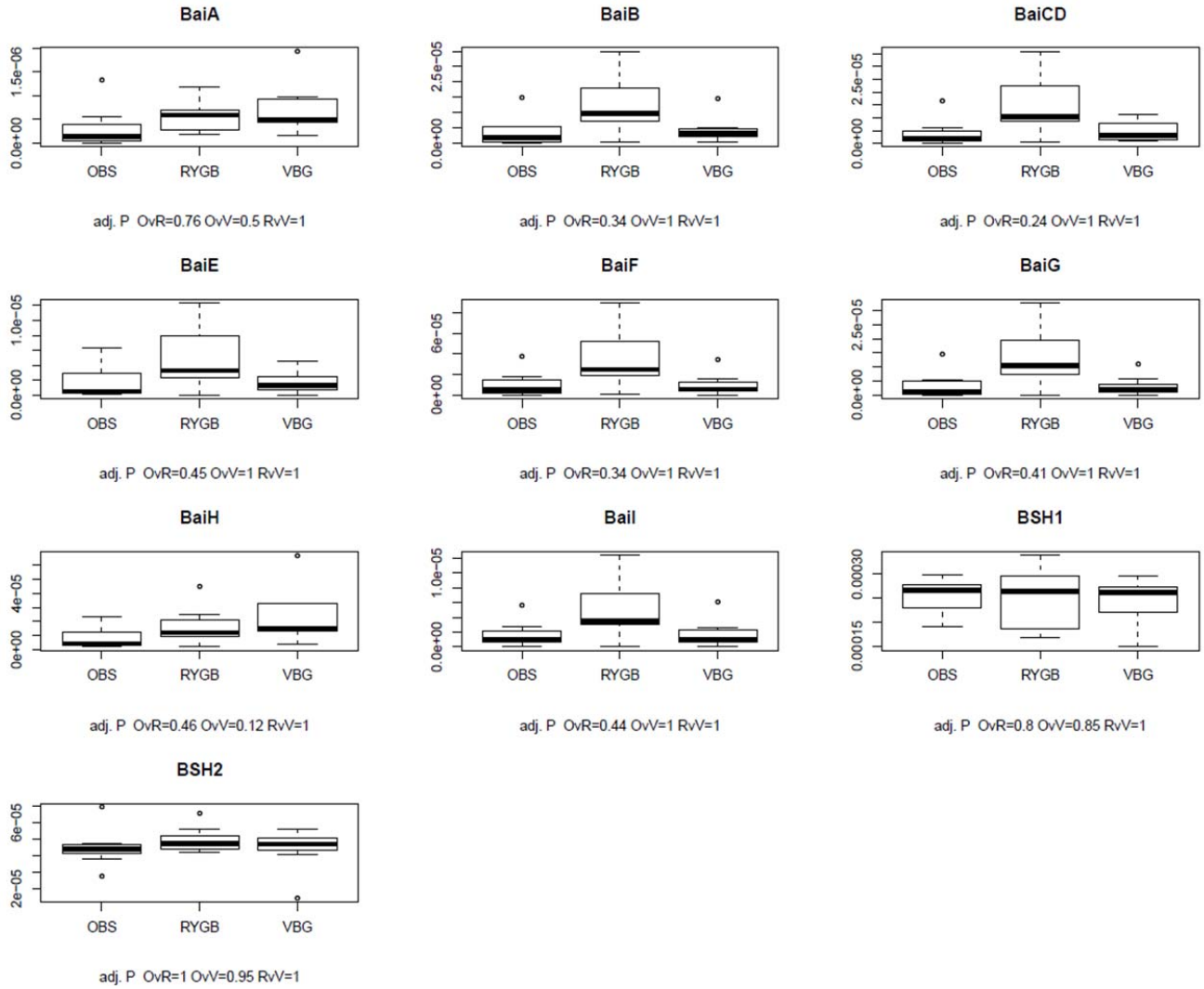


Figure S2. Measurement of Short-Chain Fatty Acids (SCFAs) and Branched-Chain Fatty Acids (BCFAs) in Stools from the RYGB, VBG and OBS Women. Related to Figure 2.

(A) Total SCFAs measured as the sum of the concentrations of acetate, propionate and butyrate. (B) Concentration of acetate (Ac), (C) propionate (Pr) and (D) butyrate (But). The relative proportions of acetate, propionate and butyrate in the total SCFA pools (% Tot SCFA) are indicated under each plot. (E) Total BCFAs measured as the sum of the concentrations of isobutyrate and isovalerate. (F) Concentration of isobutyrate and (G) isovalerate. (H) Ratio of SCFA/BCFA as a measure of microbial fermentation of carbohydrates versus proteins.

Concentrations were normalized to the dry weight of the extracted sample. Results are presented as mean ± standard error of the mean. [a] $P<0.05$ for RYGB compared to OBS according to one-way ANOVA with Tukey's correction for multiple comparisons.

A


BaiA
0.0e+00
1.5e−06
OBS
RYGB
VBG
adj. P OvR=0.76 OvV=0.5 RvV=1
BaiB
0.0e+00
2.5e−05
OBS
RYGB
VBG
adj. P OvR=0.34 OvV=1 RvV=1
BaiCD
0.0e+00
2.5e−05
OBS
RYGB
VBG
adj. P OvR=0.24 OvV=1 RvV=1
BaiE
0.0e+00
1.0e−05
OBS
RYGB
VBG
adj. P OvR=0.45 OvV=1 RvV=1
BaiF
0e+00
6e−05
OBS
RYGB
VBG
adj. P OvR=0.34 OvV=1 RvV=1
BaiG
0.0e+00
2.5e−05
OBS
RYGB
VBG
adj. P OvR=0.41 OvV=1 RvV=1
BaiH
0e+00
4e−05
OBS
RYGB
VBG
adj. P OvR=0.46 OvV=0.12 RvV=1
BaiI
0.0e+00
1.0e−05
OBS
RYGB
VBG
adj. P OvR=0.44 OvV=1 RvV=1
BSH1
0.00015
0.00030
OBS
RYGB
VBG
adj. P OvR=0.8 OvV=0.85 RvV=1
BSH2
2e−05
6e−05
OBS
RYGB
VBG
adj. P OvR=1 OvV=0.95 RvV=1

B

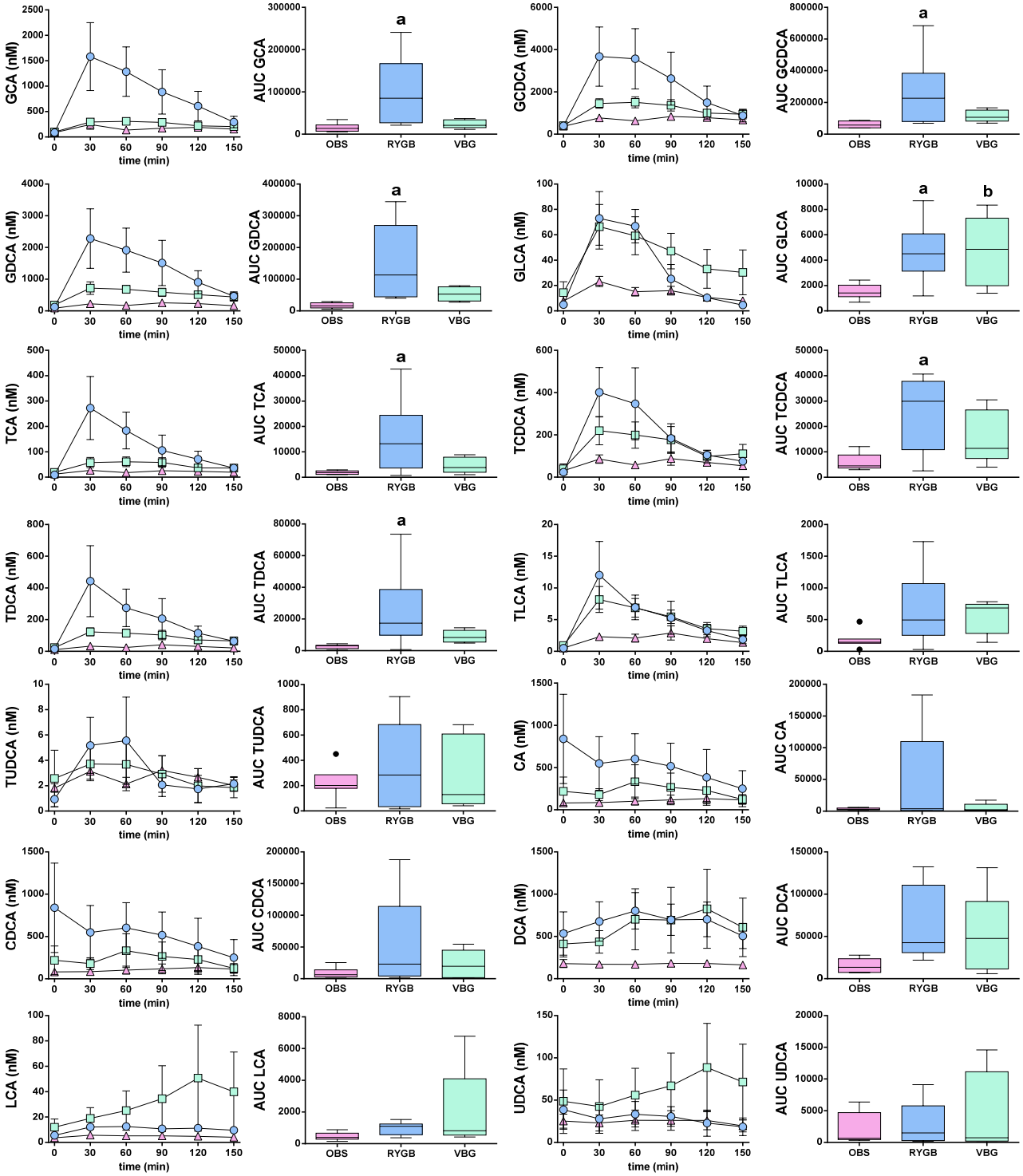

GCA (nM)
AUC GCA
GCDCA (nM)
AUC GCDCA
GDCA (nM)
AUC GDCA
GLCA (nM)
AUC GLCA
TCA (nM)
AUC TCA
TCDCA (nM)
AUC TCDCA
TDCA (nM)
AUC TDCA
TLCA (nM)
AUC TLCA
TUDCA (nM)
AUC TUDCA
CA (nM)
AUC CA
CDCA (nM)
AUC CDCA
DCA (nM)
AUC DCA
LCA (nM)
AUC LCA
UDCA (nM)
AUC UDCA
time (min)
OBS
RYGB
VBG
a
b

Figure S3. Identification of Microbial Genes for Bile Acid Metabolism and Quantification of Circulating Bile Acids after Bariatric Surgery. Related to Figure 3.

(A) Abundance of putative genes for bile acid metabolism by the gut microbiota. The details about the annotation of *bai* and *bsh* genes are described in Supplemental Experimental Procedures. Boxes denote the interquartile range (IQR) between the first and third quartiles and the line within denotes the median; whiskers denote the lowest and highest values within 1.5 times IQR from the first and third quartiles, respectively. Circles denote data points beyond the whiskers. Adjusted P values for pairwise comparisons between groups are provided under each figure using the initial letter for each group, e.g OvR for OBS vs RYGB. Bai, bile acid-inducible genes; BSH, bile salt hydrolases.

(B) Circulating bile acid concentrations in RYGB, VBG and OBS women. Plasma bile acids concentrations were measured during fasting and every 30 minutes for 2.5 hours after a standard meal. Plasma samples from colecystectomized patients (one in the RYGB and one in the VBG group) were excluded from the analysis, and plasma samples for one VBG woman could not be obtained, so the results represent the mean ± standard error of the mean for 6 RYGB, 5 VBG and 7 OBS women. The Tukey box-plots show the area under the curve (AUC) as a measure of the total postprandial response of each bile acid. [a] $P<0.05$ for RYGB compared to OBS and [b] $P<0.05$ for VBG compared to OBS according to one-way ANOVA with Tukey's correction for multiple comparisons. GCA, glycocholic acid; GCDCA, glycochenodeoxycholic acid; GDCA, glycodeoxycholic acid; GLCA, glycolithocholic acid; TCA, taurocholic acid; TCDCA, taurochenodeoxycholic acid; TDCA, taurodeoxycholic acid; TLCA, taurolithocholic acid; TUDCA, tauroursodeoxycholic acid; CA, cholic acid; CDCA, chenodeoxycholic acid; DCA, deoxycholic acid; LCA, lithocholic acid; UDCA, ursodeoxycholic acid.

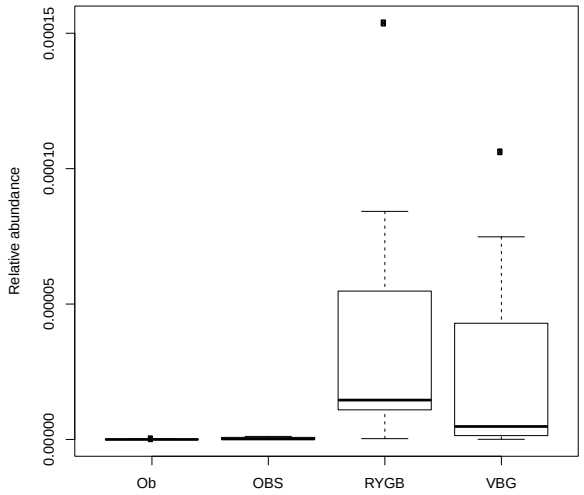


Figure S4. Enrichment of the KO Coding for Trimethylamine *N*-Oxide (TMAO) Reductase. Related to Figure 4 and Table S5.

Relative abundance of the KO coding for the TMAO reductase (K07811) in RYGB, VBG, OBS and Ob women. Boxes denote the interquartile range (IQR) between the first and third quartiles and the line within denotes the median; whiskers denote the lowest and highest values within 1.5 times IQR from the first and third quartiles, respectively. Filled squares denote data points beyond the whiskers.

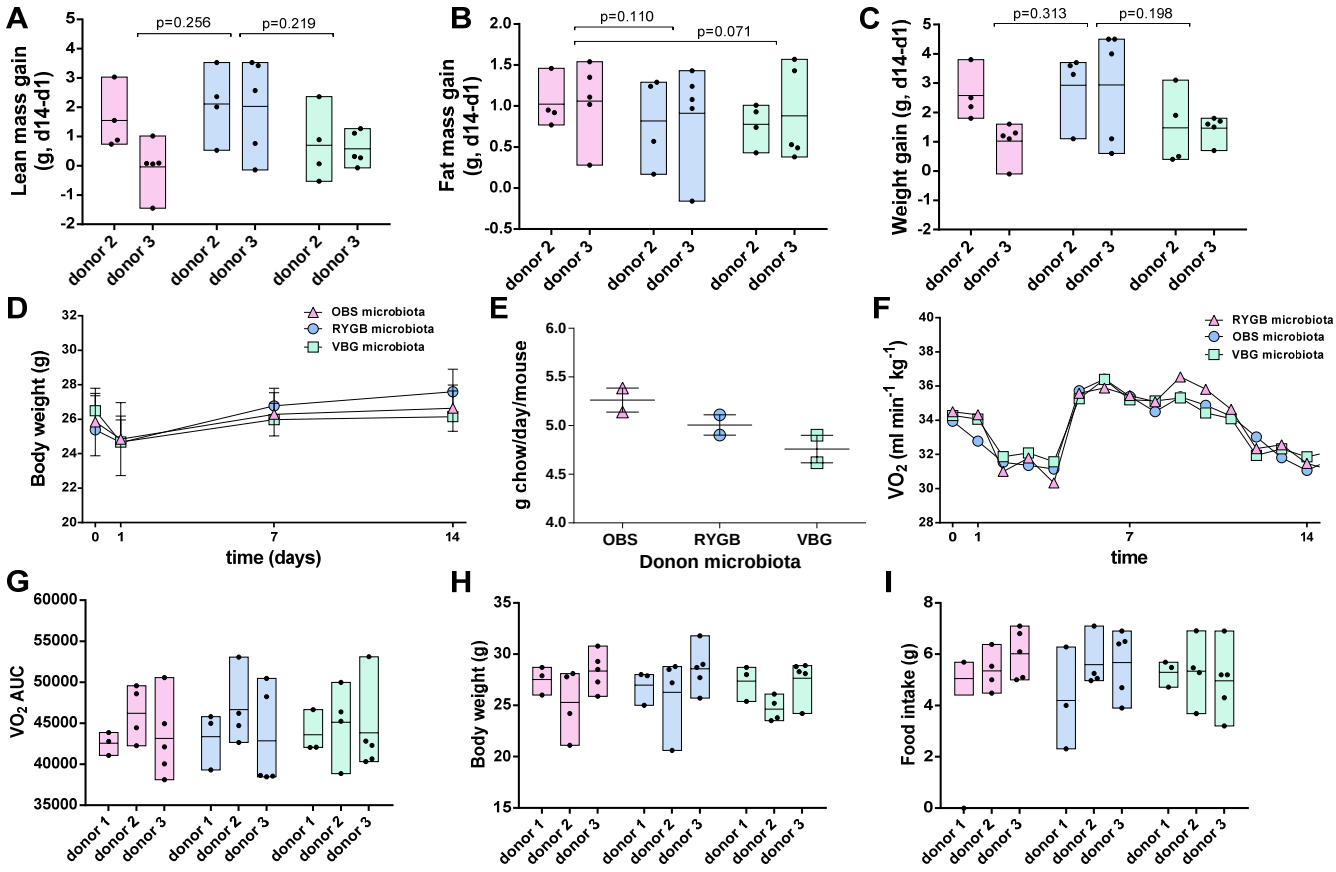


Figure S5. Weight Gain and Metabolism of the Colonized Mice During the Two-Week Colonization Period. Related to Figure 5.

(A) Absolute lean mass gain, (B) absolute fat mass gain and (C) absolute weight gain for mice colonized with fecal microbiota from RYGB, VBG or OBS women. Lean mass, fat mass and weight gains were calculated as the difference between values measured at the end of the experiment (d14) and 1 day after colonization (d1). (D) Body weight and (E) food consumption for mice colonized with fecal microbiota from RYGB, VBG or OBS women. 4-5 mice were colonized *per* donor microbiota for a total of 2 donors in each of the RYGB, VBG and OBS groups.

(F) Oxygen consumption (VO_2) over the 22 h period in the Somedic Chamber and (G) area under the curve (AUC) as a measure of total VO_2. (H) Body weight and (I) food intake for the colonized mice while in the Somedic chamber. Food intake was measured for the total 22 h period in the Somedic chamber. 3-5 mice were colonized *per* donor microbiota for a total of 3 donors in each of the RYGB, VBG and OBS groups.

The lines in the bars indicate mean values of measured parameters in mice colonized with the same donor microbiota. For the curves in panel F, only mean values are shown. The statistical significance of differences between the means was tested by one-way ANOVA with Tukey's correction for multiple comparisons (Panels G-I, Adj. P>0.7, values not shown).