research articles Journal of proteome research

Top-Down Systems Biology Modeling of Host Metabotype−Microbiome Associations in Obese Rodents

Alison Waldram,[†] Elaine Holmes,[†] Yulan Wang,[†,+] Mattias Rantalainen,[†] Ian D. Wilson,[‡] Kieran M. Tuohy,[§] Anne L. McCartney,[§] Glenn R. Gibson,[*,§] and Jeremy K. Nicholson[*,†]

Department of Biomolecular Medicine, SORA Division, Faculty of Medicine, Sir Alexander Fleming Building, Imperial College London, South Kensington SW7 2AZ, United Kingdom, AstraZeneca, Alderley Park, Mereside, Macclesfield, Cheshire SK10 4TF, United Kingdom, and Department of Food Biosciences, Whiteknights, The University of Reading, P.O. Box 226 Reading RG6 6AP, United Kingdom

Received November 14, 2008

Covariation in the structural composition of the gut microbiome and the spectroscopically derived metabolic phenotype (metabotype) of a rodent model for obesity were investigated using a range of multivariate statistical tools. Urine and plasma samples from three strains of 10-week-old male Zucker rats (obese (*fa*/*fa*, $n = 8$), lean (*fa*/−, $n = 8$) and lean (−/−, $n = 8$)) were characterized *via* high-resolution ^{1}H NMR spectroscopy, and in parallel, the fecal microbial composition was investigated using fluorescence *in situ* hydridization (FISH) and denaturing gradient gel electrophoresis (DGGE) methods. All three Zucker strains had different relative abundances of the dominant members of their intestinal microbiota (FISH), with the novel observation of a *Halomonas* and a *Sphingomonas* species being present in the (*fa*/*fa*) obese strain on the basis of DGGE data. The two functionally and phenotypically normal Zucker strains (*fa*/− and −/−) were readily distinguished from the (*fa*/*fa*) obese rats on the basis of their metabotypes with relatively lower urinary hippurate and creatinine, relatively higher levels of urinary isoleucine, leucine and acetate and higher plasma LDL and VLDL levels typifying the (*fa*/*fa*) obese strain. Collectively, these data suggest a conditional host genetic involvement in selection of the microbial species in each host strain, and that both lean and obese animals could have specific metabolic phenotypes that are linked to their individual microbiomes.

Keywords: Zucker rat • FISH • metabonomics • metabotype • microbiota • obesity • OPLS-DA • O2PLS • ^{1}H NMR • metagenomics

Introduction

All coelomate animals have symbiotic gut microorganisms providing an extended genome consortium (microbiome) that closely interacts with and modulates the metabolism, immune systems and general health of the host.[1−4] These microbiome−mammalian metabolic interactions result in the formation of a diverse range of co-metabolites *via* enzyme systems contributed by multiple interacting genomes.[4,5] Such 'transgenomic' interactions result in enriched complexity of the host's metabotype, and statistical links between microbiome variation and specific metabolic patterns have been reported,[6] which may be modulated in diverse disorders such as parasitic diseases,[7−9] in fat-fed models of nonalcoholic fatty liver disease,[10] type II diabetes[11] and obesity.[12] Gut microbial consortia can also be transplanted from species to species[13] and modulated using probiotics with resulting effects on host metabolism of short-chain fatty acids, lipoproteins and bile acids.[14] Large-scale metabolic variation between phenotypes of human populations are also partly attributable to variations in gut microbial activities[15,16] and the modulating effects of dietary choice in individuals on both the host and the microbiome.[17,18] Gut microbes have a high metabolic capacity and strongly affect drug toxicity; therefore, they may be important in population differences in drug metabolism and toxicity.[4,19] However, the role of gut microbes in the etiology of the various pathologies they have been associated with and the interacting responses of the gut microbes to the developing pathology are difficult to unravel. Top-down systems biology approaches have proved useful in understanding gut microbial−mammalian metabolic network interactions[18] and methods for connecting microbial population signatures with metabolic profiles have recently been described.[6] To investigate the potential for such transgenomic "keystone interactions", that is, the degree to which the host genome determines the microbiome and its modulation in disorder, we have chosen to apply a combined metabonomic and microbiomic[20] approach to investigate the co-

* To whom correspondence should be addressed. J.K.N.: e-mail, j.k.nicholson@imperial.ac.uk; tel, +44 (0)20 7594 3195; fax, +44 (0)20 7594 3226. G.R.G.: e-mail, g.r.gibson@reading.ac.uk; tel, +44 (0)118 378 8715, fax, +44 (0)118 931 0080.

† Imperial College London.

+ Current address: State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Centre for Magnetic Resonance, Wuhan Institute of Physics and Mathematics, The Chinese Academy of Sciences, Wuhan, 430071, PR China.

‡ AstraZeneca.

§ The University of Reading.

10.1021/pr8009885 CCC: $40.75

Published on Web 03/10/2009

variation between the metabolic phenotype and the metagenome of the Zucker rat which is an established and important model of obesity and type II diabetes.

Certain strains of the Zucker rat have a mutation of the *fa* gene[21] which leads to a defect in the leptin receptor,[22] resulting in the animal becoming overweight and glucose-intolerant. In particular, the homozygous Zucker (*fa/fa*) obese rat, widely used as a type II diabetes model, readily becomes obese and insulin-resistant typically by 12 weeks.[23,24] In contrast, the heterozygous (*fa*/−) and homozygous (−/−) Zucker rats remain lean as they age and do not become insulin-resistant.[21]

Metabonomic and metabolomic studies rely mainly on a range of NMR- and MS-based techniques to generate multiparametic metabolic profiles of physiological and pathological variation.[25−29] These techniques have been variously applied to the analysis of urine and plasma from the Zucker rat and shown to be suitable for characterizing metabolic phenotypes (metabotypes).[24,30−32] To investigate metabolic interactions between host and gut microbiota in this animal model, urine and plasma from all three Zucker strains were "metabotyped" *via* high-resolution ^{1}H NMR spectroscopy and correlated with strain-related differences in the microbiome measured using semiqualitative and semiquantitative molecular-based techniques, polymerase chain reaction (PCR) followed by denaturing gradient gel electrophoresis (DGGE)[33] and fluorescence *in situ* hybridization (FISH),[34] respectively. These resulting data were thus analyzed using various multivariate statistical methods in order to discover any underlying associations between host and gut microbiota. This pilot study presents an integrated characterization of the microbiome−host phenotype of a disease model and represents a first step toward generating hypotheses regarding potential contributions of transgenomic metabolic interactions to the development of obesity.

Materials and Methods

Animals and Sample Collection. Three strains of male rat, Zucker (*fa/fa*) obese $n = 8$, Zucker (*fa*/−) lean $n = 8$ and Zucker (−/−) lean $n = 8$ were used in this study (animals bred on site, Alderley-Park, AstraZeneca). The animals were produced by mating (*fa*/−) male and female Zucker rats to produce the required crosses. The pups were reared together until 6 weeks of age, at which point they were separated into three strains after genotyping. Thereafter, the animals were gang-housed in polycarbonate cages, with all three groups maintained in the same animal room and rack. Food (standard rat and mouse diet, R&M No 1 modified irradiated diet, Special Diet Services Ltd., Essex, U.K.) and water were available *ad libitum* throughout. At 10 weeks of age, the animals were euthanized, and urine, feces and plasma were collected. Urine and feces were stored frozen at −20 °C until analysis. Blood was collected into lithium heparin vials and centrifuged at ∼2400*g* for 10 min, and the plasma was then removed and stored at −20 °C until analysis.

^{1}H NMR Spectroscopy. For urine, an aliquot of 460 μL of each sample was mixed with 240 μL of phosphate buffer (0.2 M in 10% D_2O; pH 7.4) containing sodium 3-(trimethylsilyl) propionate-2,2,3,3-d_4 (TSP) (0.5 mg/mL) as a chemical shift reference. The samples were centrifuged at ∼16 000*g* for 10 min and 600 μL of the solution was transferred to a 5 mm outer diameter Precision NMR tube (GPE Scientific Ltd.) before being analyzed by ^{1}H NMR spectroscopy. For plasma, 300 μL of each sample was diluted with 600 μL of a 0.9% (w/v) saline solution in 10% D_2O. The samples were centrifuged at ∼16 000*g* for 10 min and 600 μL of solution was transferred to a 5 mm outer diameter Precision NMR tube (GPE Scientific Ltd.) before being analyzed by ^{1}H NMR spectroscopy, as with the urine samples. The analysis was carried out using a Bruker Avance 600 NMR spectrometer (Bruker, Germany) operating at a 600.44 MHz ^{1}H resonance frequency with a BBI probe. For the urine samples, a standard one-dimensional pulse sequence was used with a presaturation water suppression pulse sequence: 90°-t_1-90°-t_2-90° acquire FID (Free Induction Decay) [$t_1 = 3$ μs, $t_2 = 100$ ms].[35]

For each spectrum, 128 transients were collected into 32k data points. The temperature remained constant at 298 K and the 90° pulse was set to 13 μs for the urine spectra and 11.5 μs for the plasma spectra. For the plasma samples, three types of ^{1}H NMR experiment were performed, that is, the standard one-dimensional pulse sequence used with water suppression as described above, the water-suppressed Carr−Purcell−Meiboom−Gill (CPMG) spin−echo spectrum: RD-90°−(t-180°-t)$_n$− acquire FID.[36,37]

This pulse sequence attenuates signals from fast relaxing species, such as broad peaks from high molecular weight compounds, and so forth. Finally, the one-dimensional diffusion-edited ^{1}H NMR spectrum which attenuates signals from small molecules with fast translational diffusion rates[38,39] (generally small molecules with a low Stokes' radius): RD-90°-t_1-180°-t_1-90°-t_2-90°-t_1-180°-t_1-90°-LED-90° acquire FID.

Absolute Glucose Measurements. Each plasma sample was analyzed for glucose using a MediSense Precision QID monitor and MediSense blood glucose test strips (Abbott, Kent) with one drop of plasma.

Microbiome Analyses. 1. Microbial DNA Extraction. Each fecal sample was thawed on ice; for each, a 1 in 10 dilution (w/w) was made in phosphate-buffered saline (Oxoid). An aliquot (1 mL) of sample was centrifuged at ∼13 800*g* (rotor no. 3753, Biofuge 28RS; Heraeus SEPATECH) and 4 °C for 5 min in a weighed sterile 2 mL Safe-Lock tube. The supernatant was discarded and the pellet weighed. Additional sample was added to the tube and centrifuged to give a pellet of ∼500 mg. For each sample, DNA was extracted from the pellet by using the FastDNA SPIN Kit (for soil) (QBiogene) following the manufacturer's instructions. The quantity of DNA in each sample was assessed by running 5 μL aliquots of the samples with a MassRuler-DNA Ladder Mix (Helena Biosciences Europe) on 1.5% (w/v) ultraPURE agarose (GibcoBRL) gels containing ethidium bromide (0.4 mg/mL), and visualizing the bands using a UV light. For each sample, DNA was diluted in sterile distilled water to give a concentration of ∼5 ng/μL.

2. Polymerase Chain Reaction (PCR). Fragments derived from the variable V3 region of the 16S rRNA gene were amplified from samples by using the universal primers p2 (5′-ATTACCGCGGCTGCTGG-3′; synthesized by MWG Biotech) and p3 (5′-CGCCCGCCGCGCGCGGCGGGCGGGGCGGGGGCACG-GGGGGCCTACGGGAGGCAGCAG-3′; synthesized by MWG Biotech).[33] PCR mixtures comprised 5 μL of 10× $MgCl_2$-free buffer (Promega), 5 μL of dNTPs (12.5 mM each; Promega), 1 μL of primer p2 (20 pmol), 1 μL of primer p3 (20 pmol), 6 μL of $MgCl_2$ (25 mM; Promega), 30 μL of H_2O, 1 μL of *Taq* (1.25 U DNA polymerase; Promega) and 1 μL of template DNA. The PCR was carried out using a MJ Research PTC-200 Peltier Thermal Cycler (GRI) machine with the following parameters under 'block' control. Samples were heated at 94 °C for 5 min; following the hot start, two cycles consisting of 1 min at 94 °C, 1 min at 65 °C and 1 min at 72 °C were run; the annealing temperature was subsequently decreased by 1 °C every second cycle until a

'touchdown' at 55 °C, at which temperature five additional cycles were carried out (27 cycles in total).[33] Amplification products were examined using agarose gel electrophoresis; products were stored at −20 °C until required.

3. Denaturing Gradient Gel Electrophoresis (DGGE). This was carried out using the INGENYphorU2 system (GRI). PCR samples were applied directly onto polyacrylamide gels in 0.5× TAE buffer (diluted from a 50× concentrated solution; Fisher Scientific) with gradients (30−60%) that were formed with 8% (w/v) acrylamide stock solutions [40% acrylamide/bis solution, 37.5:1 (2.6% C); Bio-Rad] containing 2% (v/v) glycerol (BDH), and which contained 0 and 100% denaturant [(7 M PlusOne urea; Pharmacia Biotech) and 40% (w/v) PlusOne formamide (Amersham Biosciences)]. Electrophoresis was run at a constant voltage of 150 V and a temperature of 60 °C for 16 h. Following electrophoresis, the gels were silver-stained according to the method of Sanguinetti et al.[40] with minor modifications. Gels were scanned at 600 dpi and the images analyzed using GeneTools and GeneDirectory (Syngene).

4. Cloning and Sequencing of Bands. Bands of interest were excised, crushed in 20 μL of sterile distilled H_2O and left at 4 °C overnight. Samples were pulse-centrifuged and aliquots (3 μL) of the supernatants from the crushed bands were used in PCRs with primer p2 and primer p3 as described above. Each resulting PCR product was cloned using the StrataClone PCR Cloning Kit (Stratagene) according to the manufacturer's instructions. Three clones for each PCR product were selected randomly from LB/ampicillin agar plates and subcultured. After incubation overnight at 37 °C, plasmids were extracted from the clones by using the QIAprep Spin Miniprep kit (Qiagen). Sequencing reactions were set up containing the following: 4 μL of BigDye Terminator v3.1 (Applied Biosystems), 1 μL of primer T7 (5′-TAATACGACTCACTATAGGG-3′), 3 μL of PCR product and 2 μL of sterile distilled H_2O. The reactions were run on an MJ Research PTC-200 Peltier Thermal Cycler machine with the following parameters under 'block' control: 25 cycles of 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 min. Sequencing of PCR products was done *via* the University of Reading's Biocenter.

5. Analysis of Sequence Data. All sequences were edited *via* 4Peaks version 1.7.2 (by A. Griekspoor and Tom Groothuis, http://www.mekentosj.com) and proofread. WU-Blast2 searches were conducted *via* the European Bioinformatics Institute Web site (http://www.ebi.ac.uk/databases). Sequence results are presented as similarity (%) to nearest relative, as absolute identification of species represented by bands cannot be made on the basis of an ∼190 nt fragment.

6. Fluorescence *in Situ* Hybridization (FISH). Frozen fecal samples were diluted 1 in 10 (w/v) with phosphate-buffered saline (pH 7.2; Oxoid) and homogenized by pipetting through a 1 mL tip for 2 min. Aliquots (375 μL) of homogenized fecal samples were fixed overnight in 3 vol of cold 4% paraformaldehyde. Cells were harvested by centrifugation and washed twice in 1 mL of phosphate buffered saline, then resuspended in 300 μL of phosphate-buffered saline/ethanol (1:1, v/v) and stored at −20 °C until required.

Differences in fecal bacterial populations were assessed through the use of FISH with DNA probes targeting specific diagnostic regions of the 16S rRNA molecule. To provide up-to-date information on probe specificity, probe sequences were checked against database sequences *via* Probe Match at the RDP (http://rdp.cme.msu.edu/index.jsp) using the following search options: 0 mismatches; strains, type and nontype; source, isolates; size ≥1200 nt; quality good.

The probes used were Bif164 (targeting most bifidobacteria and *Parascardovia denticolens*[41]), Bac303 (most *Bacteroides* and *Prevotella* spp., *Barnesiella* spp. and *Odoribacter splanchnicus*[42]), Chis150 (most members of *Clostridium* cluster I and all members of *Clostridium* cluster II[43]), Erec482 (most members of *Clostridium* cluster XIVa[43]), Lab158 (most *Lactobacillus*, *Leuconostoc* and *Weissella* spp., all *Enterococcus*, *Vagococcus*, *Melisococcus*, *Tetragenococcus*, *Paralactobacillus*, *Pediococcus*, *Oenococcus* and *Catellicoccus* spp. and *Lactococcus lactis*[44]), and Ato291 (all *Cryptobacterium*, *Collinsella*, *Atopobium*, *Eggerthella* and *Olsenella* spp.[45]). Probes were commercially synthesized and 5′-labeled with the fluorescent dye Cy3 (Sigma Aldrich). Hybridization and washing conditions were as described in the literature. The nucleic acid stain DAPI (4′6-diamidino-2-phenylindole) was used for total bacterial counts. Washed samples were collected on vacuum-mounted 0.2 μm polycarbonate filters (Millipore), which were transferred to slides. A drop of SlowFade Light Antifade Kit (Invitrogen Ltd.) and a coverslip were applied to each slide. Slides were stored in the dark at 4 °C until counted or for a maximum of 3 days. Slides were viewed under a Nikon E400 Eclipse microscope. DAPI slides were visualized with the aid of a DM 400 filter; probe slides were visualized with the aid of a DM 575 filter. Cells (between 15 and 50 per field of view) were counted for 15 fields of view, and the numbers of specific bacteria and total bacteria were determined by using the following equation;

$$\frac{DF \times ACC \times 14873.73 \times 1000}{VH}$$

Where the DF (dilution factor) was calculated by taking into account the concentration of the original sample (375 μL to 300 μL = 0.8×, taken to 8× to account for the 1/10 dilution made of the original fecal sample), dilution of the sample in hybridization buffer and the ratio of total hybridization volume to volume of sample added. ACC (average cell count) was determined by counting 15 fields of view and assumes that a normal distribution was observed for the counts. The figure 14873.73 refers to the area of the filter divided by the area of the field of view. *VH* refers to the volume of hybridization mixture added to the wash buffer. Results from the FISH data are given as means ± standard deviations (SD). Statistically significant differences in bacterial groups between the different groups of animals were determined using the unpaired Student's *t*-test ($P < 0.05$).

7. Data Analysis. XWIN NMR software was used to manually phase, baseline correct and calibrate the NMR spectra. For urine, the TSP peak was calibrated to δ0.0; for plasma, the chemical shift of the anomeric proton from alpha-glucose was calibrated to δ5.223. The spectra were then exported into the MATLAB (MathWorks) program where, for the urine spectra, the spectral regions containing water and urea signals was removed (δ4.6−6.0) and spectra normalized to the total area. Orthogonal Projection to Latent Structures Discriminant Analysis (OPLS-DA) models were constructed using unit variance scaled and mean-centered NMR data as the descriptor matrix and class information as the response variable.[46−48] The plasma CPMG spectra were also imported into MATLAB and the water peak (δ4.6−5.0) was removed. Again an OPLS-DA model was created, as for the urine data. As the spectra were affected by shifts relating to small pH changes, both models were recreated

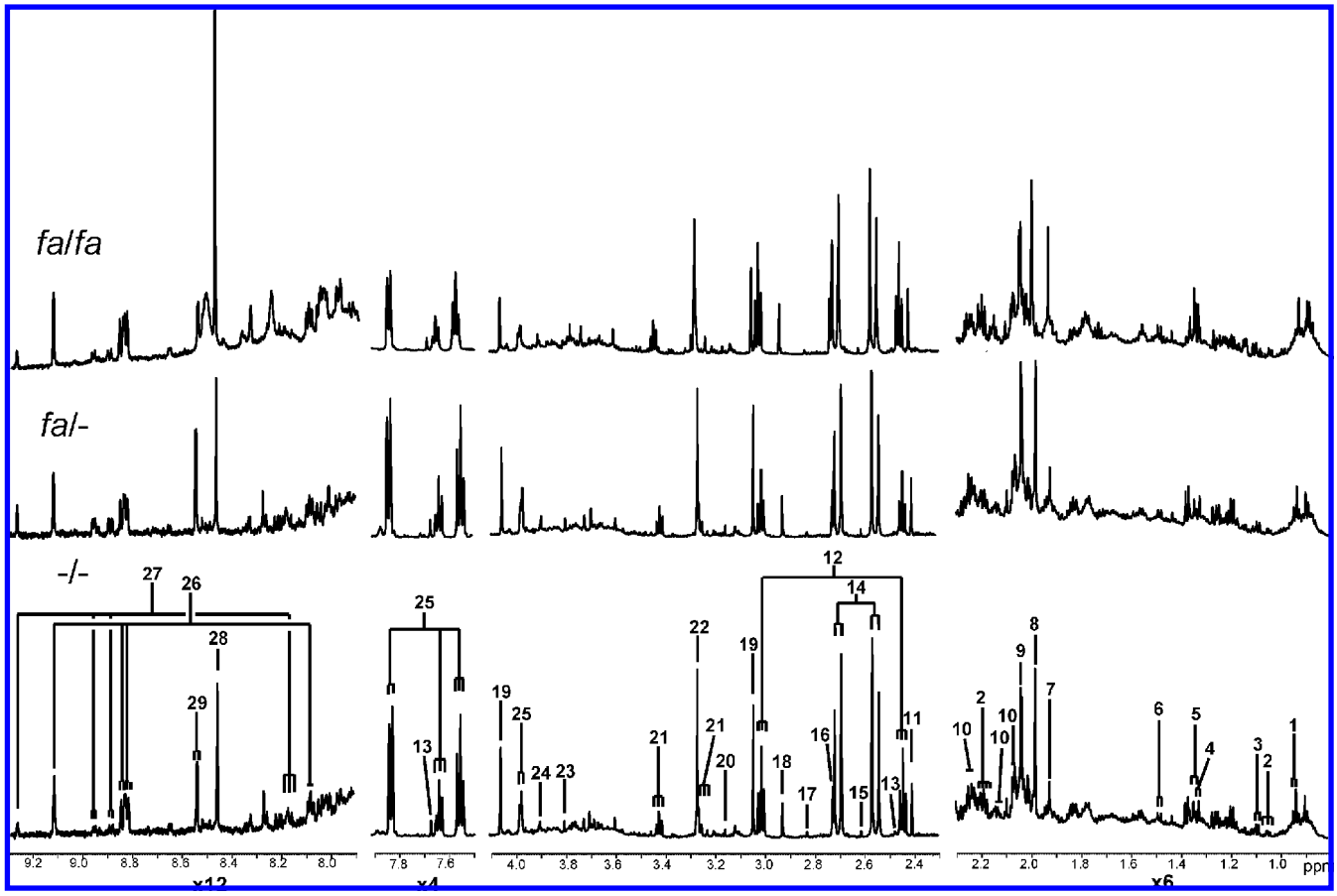

Figure 1. Typical urinary ^{1}H NMR spectra for each Zucker strain. **1**, 3-methylglutarate; **2**, Propionate; **3**, 3-hydroxyisobutyrate; **4**, Lactate; **5**, Threonine; **6**, Alanine; **7**, Acetate; **8**, **9**, Glycoprotein fragments (*N*-acetyl); **10**, Formiminoglutamic acid; **11**, Succinate; **12**, 2-oxoglutarate; **13**, Pyridoxine; **14**, Citrate; **15**, Methylamine; **16**, Dimethylamine; **17**, Trimethylamine; **18**, Dimethylglycine; **19**, Creatinine; **20**, *cis* Aconitate; **21**, Taurine; **22**, Trimethylamine-*N*-oxide; **23**, Guanidoacetate; **24**, Betaine; **25**, Hippurate; **26**, unknown; **27**, *N*-methylnicotinamide; **28**, Formate; **29**, unknown.

with the use of a peak alignment protocol.[49] Both the microbiology and NMR data were analyzed by OPLS-DA to assess the difference between the animal genotypes as indicated by the OPLS-DA predictive cross-validated scores. The OPLS-DA model was interpreted by means of the OPLS coefficients. Integrated modeling of microbiology and NMR data was accomplished using the OPLS and O2PLS methods[50] using back projection of the variables to the covariance matrix according to the method of Cloarec et al.[47] In the O2PLS analysis, the DGGE data were used as the response matrix with the urine spectra as the descriptor matrix. Prediction performance was assessed by the Q^2 parameter,[51] indicating how well each variable is predicted over cross-validation. Individual DGGE bands were used as response variables with the urine ^{1}H NMR data as the descriptor matrix to complete OPLS analysis.

Results

^{1}H NMR Spectroscopy of Urine. Selected expansions of typical urinary ^{1}H NMR spectra of each Zucker strain are shown in Figure 1. For clarity, the spectra are displayed in four frequency ranges and were broadly similar in composition to published rat urine spectra from previous studies with metabolites such as citrate, 2-oxoglutarate, trimethylamine *N*-oxide (TMAO), hippurate and creatinine dominating the profile.[24,52,53] Visual examination of the spectra indicated that the (*fa*/*fa*) obese strain of the Zucker rats contained relatively higher levels of acetate (δ1.93 singlet (s)) and lower levels of hippurate (δ3.97 doublet (d), δ7.56 triplet (t), δ7.65 t, δ7.84 d) and creatinine (δ3.05 s, δ4.05 s) compared to the other two strains.

The rat strains were easily discriminated using an Orthogonal Projection to Latent Structure-Discriminant Analysis (OPLS-DA) model as indicated in the cross-validated OPLS scores plot (Figure 2A). The obese (*fa*/*fa*) strain is completely separated from the lean strains, but the two phenotypically lean strain clusters overlap. To systematically assess detailed metabolic differences between the three strains, an OPLS-DA model of ^{1}H NMR data obtained from urine was used to establish pairwise models between all the strains. Successful discrimination between the (−/−) and the (*fa*/*fa*) strains and (*fa*/−) and (*fa*/*fa*) strains was achieved based on a Q^2 or goodness-of-prediction value of 0.86 and 0.82, respectively, whereas the comparison for the (−/−) and (*fa*/−) showed no strongly significant differences. The OPLS coefficient plots generated for each model provided a means for interpretation of the relative influence of the original variables (chemical shifts) on the OPLS-DA models. The coefficient plots are color-coded according to the strength of the OPLS coefficients, which are calculated from the correlation matrix. Orientation of the peak represents its positive correlation with a given class, and the magnitude of the signals in the plot is back-scaled to the covariance matrix. This process results in a plot which is similar

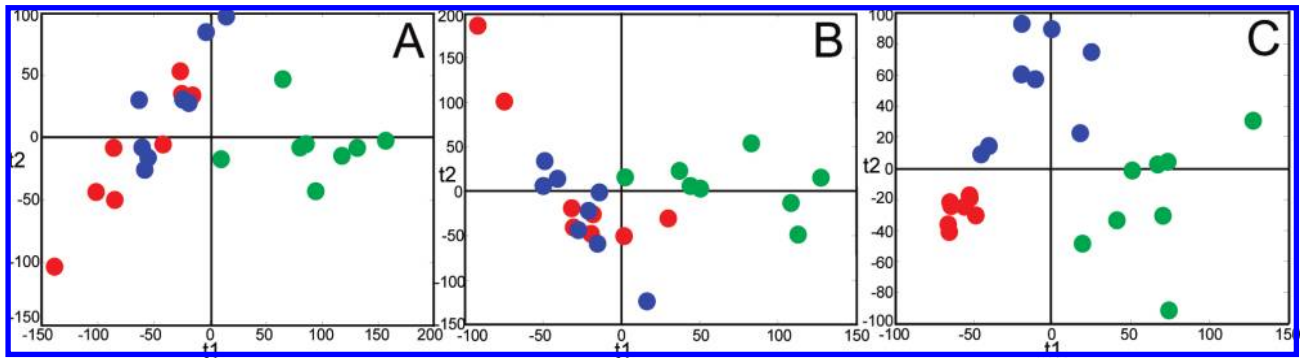

Figure 2. Fully cross-validated scores plot from the OPLS-DA model for the three Zucker strains. (A) Scores plot from the urinary ^{1}H NMR data. (B) Scores plot from the plasma CPMG ^{1}H NMR data. (C) Scores plot from the DGGE data (red circle, (−/−) lean strain; blue circle, (*fa*/−) lean strain; green circle, (*fa*/*fa*) obese strain).

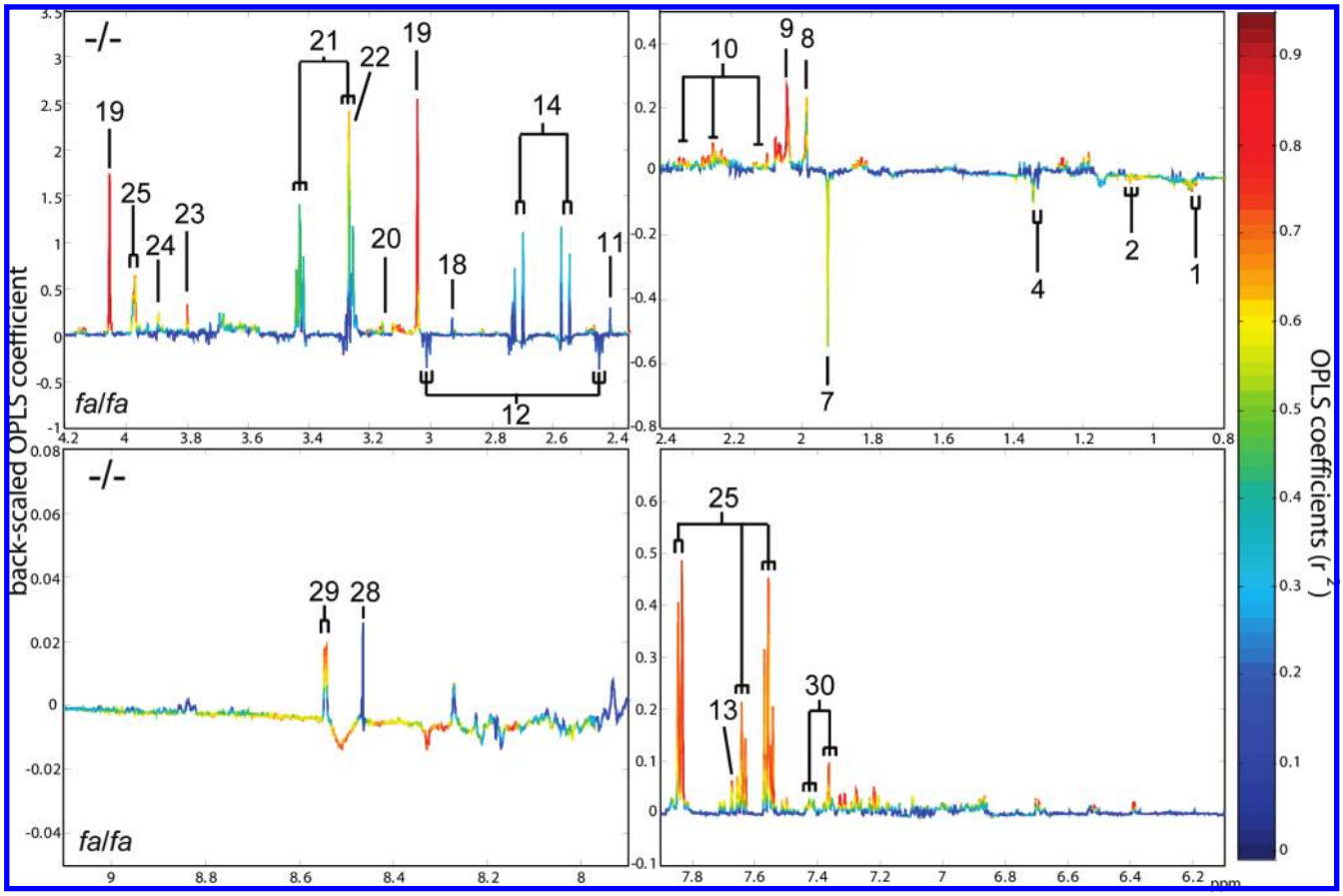

Figure 3. An OPLS-DA model discriminating between the lean (−/−) strain and the (*fa*/*fa*) obese strain from the urinary ^{1}H NMR profiles. $R^2 = 0.97$; $Q^2 = 0.86$. **1**, 3-Methylglutarate; **2**, Propionate; **4**, Lactate; **6**, Alanine; **7**, Acetate; **8**, **9**, Glycoprotein fragments (*N*-acetyl); **10**, Formiminoglutamic acid; **11**, Succinate; **12**, 2-Oxoglutarate; **13**, Pyridoxine; **14**, Citrate; **18**, Dimethylglycine; **19**, Creatinine; **20**, *cis* aconitate; **21**, Taurine; **22**, Trimethylamine-*N*-oxide; **23**, Guanidoacetate; **24**, Betaine; **25**, Hippurate; **28**, Formate; **29**, unknown; **30**, Phenylacetylglycine.

to the original NMR spectra thereby aiding the interpretation of the discriminating peaks.[47] Thus, a red colored peak pointing upright, for example, creatinine (19) in Figure 3, is positively correlated with the (−/−) lean strain (the lean strain is represented at the top of each pairwise coefficient plot). It also signifies that the (−/−) lean strain has a relatively higher concentration of a particular metabolite than the comparator strain. To confer equal influence of all variables on the model, unit variance scaling of mean-centered data was applied to the descriptor matrix such that even metabolites with low concentrations can influence the overall model if the direction and nature of differentiation is consistent throughout all animals in a given strain. Selected sections of the OPLS coefficient plot illustrating the differences between the (−/−) lean and the (*fa*/*fa*) obese strains are shown in Figure 3. The (−/−) and the (*fa*/−) strains produced similar metabolite profiles, resulting in a poor prediction model when these strains were compared. This was also consistent with the model comparing the (*fa*/−) against the (*fa*/*fa*) strain being highly similar to the model comparing (−/−) and (*fa*/*fa*) strains (data not shown).

The major differences in metabolite concentrations between the (*fa*/*fa*) obese and the (−/−) lean Zucker strains included

Table 1. Relative Change in Concentration between the Lean (−/−) Strain and the (*fa*/*fa*) Obese Strain from the Urinary and Plasma CPMG ^{1}H NMR Data[a]

URINE		PLASMA		
Metabolite	δ	(*fa*/*fa*) > (-/-)		
(*fa*/*fa*) > (-/-)		Metabolite		δ
3-methylglutarate	0.93 (d), 2.23 (m)	LDL	$C=CCH_2C=C$	2.73 (bs)
Propionate	1.08 (t), 2.19 (q)	LDL	$C=CCH_2C=C$	2.77 (bs)
Lactate	1.33 (d), 4.11 (q)	Unsaturated LDL	$=CHCH_2$	5.30 (bs)
Acetate	1.92 (s)	VLDL	$(CH_2)_n$	1.28 (bs)
(*fa*/*fa*) < (-/-)		LDL	CH_2-C=C	2.00 (bs)
Creatinine	3.04 (s), 4.05(s)	VLDL	CH_2CH_2CO	1.56 (bs)
Forminimoglutamic acid	2.04 (m), 2.16 (m), 2.22 (m) 4.14 (t)	VLDL	$-CH_2CO$	2.22 (bs)
Pyridoxine	2.47 (s), 7.67 (s)	LDL	CH_3	0.87 (bs)
Glycoprotein (*N*-acetyl)	1.98 (s), 2.05 (s)	LDL	$C=CCH_2C=C$	2.81 (bs)
Guanidoacetate	3.80 (s)	Choline	$N(CH_3)_3$	3.20 (bs)
Hippurate	3.97 (d), 7.56 (t), 7.64 (t), 7.84 (d)	Lactate		1.33 (d), 4.11 (q)
Phenylacetylglycine	3.67 (s), 3.77 (d), 7.35 (t), 7.41 (t)	Glyceryl of lipids	CHOCOR	5.20 (bs)
Betaine	3.27 (s), 3.90 (s)	Isoleucine		0.94 (t), 1.01 (d), 1.26 (m), 1.48 (m), 1.98 (m), 3.68 (d)
Trimethylamine-*N*-oxide	3.27 (s)	Valine		0.99 (d), 1.04 (d), 2.28 (m) 3.62 (d)
cis Aconitate	3.16 (s)	Acetoacetate		2.22 (s), 3.43 (s)
Taurine	3.27 (t), 3.43 (t)	(*fa*/*fa*) < (-/-)		
		Metabolite		δ
		Glycine		3.56 (s)
		Glutamate		2.43 (m)

[a] Listed in order of magnitude, greatest first. (s), singlet; (d), doublet; (t), triplet; (q), quartet; (m), multiplet; (bs), broad singlet.

lower urinary concentrations of creatinine, hippurate and TMAO in the (*fa*/*fa*) obese strain, with creatinine being the most markedly different between the two groups. The (*fa*/*fa*) strain had higher levels of urinary acetate than (−/−) and (*fa*/−). From the OPLS coefficient plots, a relative measure of concentration differences could be observed and they also showed which of the rat strains had the dominant amount of any given metabolite. The relative differences in concentration for the most important metabolites are listed in Table 1.

^{1}H NMR Spectroscopy of Plasma. Each plasma sample was analyzed by three different NMR experiments producing three separate spectral data sets for each sample, containing information on molecular diffusion and spin relaxation. An example of each spectrum type can be viewed in the Supporting Information, Figure 1. Spin−echo spectroscopy (CPMG) of plasma allows signals from small molecules to be observed more easily because of the attenuation of broad protein and lipoprotein signals *via* T_2 relaxation.[25,54] If the overall lipoprotein concentrations are high, then signals from lipoproteins, particularly the more mobile components of the lipoproteins, may still be observed; consequently, the CPMG spectra have been chosen to create the statistical models to represent the plasma samples and the other spectra were consulted for verification of several of the lipoprotein peaks. Selected expansions of typical plasma CPMG spin−echo ^{1}H NMR spectra of each strain, displayed over three different NMR frequency ranges for clarity, are supplied in the Supporting Information Figure 2. The glucose signals (δ3.2−3.9) appeared slightly lower in intensity in the plasma of the (*fa*/*fa*) obese animals compared to those of the other strains. This is in agreement with previous ^{1}H NMR investigations,[32] but was not found to be an important discriminator in the OPLS-DA model (Figure 4), suggesting that glucose levels were still under control and that the (*fa*/*fa*) strain had not yet developed insulin resistance. Conventional measurement of plasma glucose levels supported this finding (lower plasma glucose levels for the obese (*fa*/*fa*) strain), but likewise, these data were not found to be statistically significant (−/− 16.4 mmol/L, *fa*/− 17.8 mmol/L, *fa*/*fa* 15.1 mmol/L). The spectra obtained for the plasma of the (*fa*/−) and (−/−) lean strains were similar and did not result in discriminatory models, as was the case for the urine spectral data (data not shown). OPLS-DA modeling enabled discrimination between the obese (*fa*/*fa*) strain and the two lean strains, which were overlapping in the cross-validated OPLS scores plot (Figure 2B). Table 1 shows the relative concentration differences for metabolites demonstrating the highest correlation with strain. A key metabolic difference separating the (−/−) and the (*fa*/*fa*) strains in the plasma spectra is acetoacetate (δ2.22, singlet), which is higher in the (*fa*/*fa*) obese strain (Figure 4). Another major difference was that the plasma of the (*fa*/*fa*) obese strain contained higher levels of low density lipoprotein (LDL) and very low density lipoprotein (VLDL) concentrations relative to the other strains. This was also confirmed by viewing the relative lipoprotein concentrations in an OPLS-DA analysis of the diffusion-edited spectra (Figure 3, Supporting Information).

Microbiology. For a quantitative measure of fecal bacterial numbers, FISH was used to estimate the relative population sizes of bacterial groups in the feces of the three different Zucker strains. While FISH analysis has inherent biases in relation to cell-membrane permeability and accessibility of probes, it gives a more accurate representation of the numbers of bacteria belonging to different bacterial groups (at the phylum, family or genus level) than PCR-based studies. Differences in bacterial population levels within the gut microbiota of the different Zucker strains may indicate a direct impact of the obese state on the composition and functioning of the intestinal microbiota.

Numbers of total bacteria, based on DAPI counts, were significantly lower in the obese (*fa*/*fa*) rats compared to the nonobese (*fa*/− and −/−) rats (Figure 5A). The most striking difference between the microbiota profiles of the (*fa*/*fa*) rats

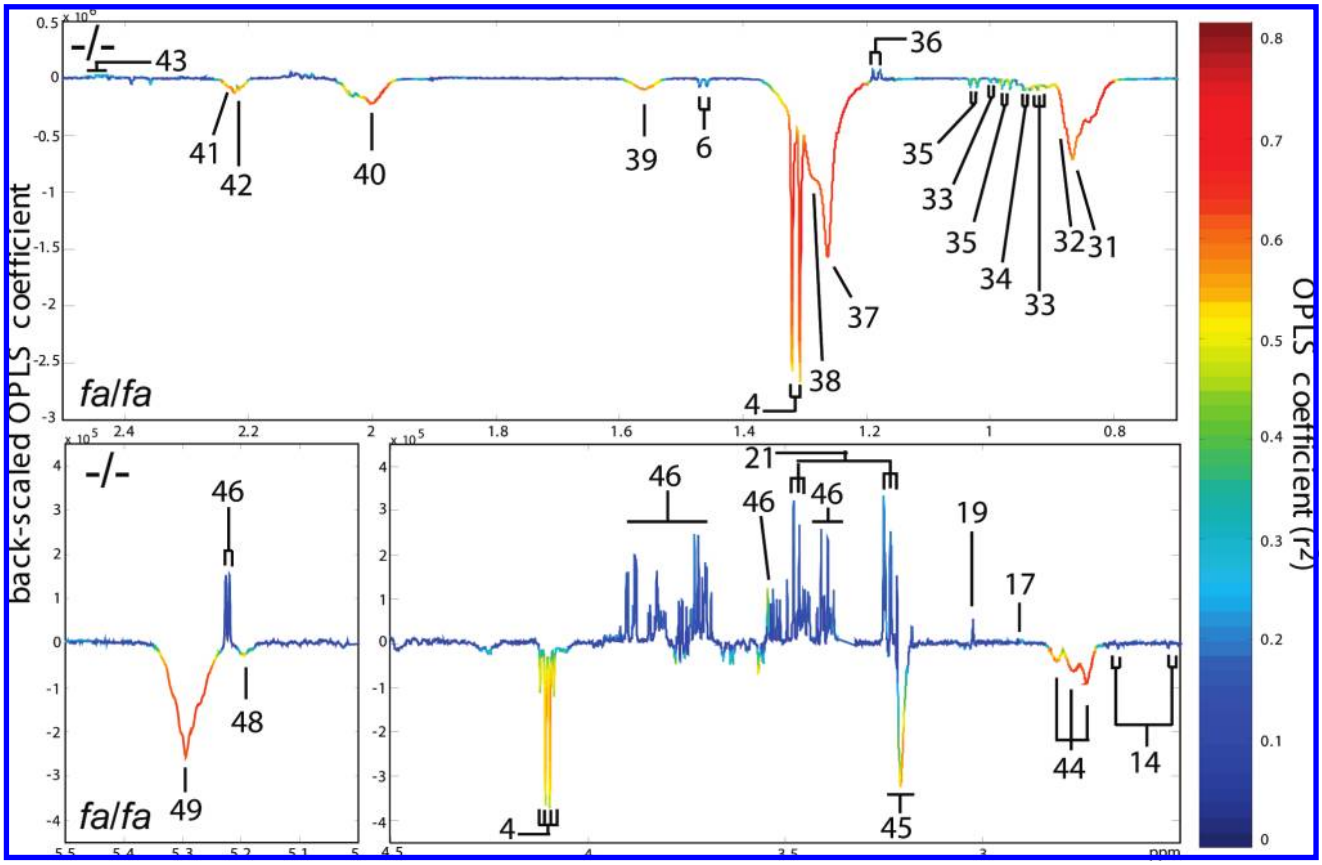

Figure 4. An OPLS-DA model discriminating between the lean (−/−) strain and the (*fa*/*fa*) obese strain from the plasma CPMG ^{1}H NMR profiles; $R^2 = 0.82$; $Q^2 = 0.69$. **4**, Lactate; **6**, Alanine; **7**, Acetate; **14**, Citrate; **17**, Trimethylamine; **19**, Creatinine; **21**, Taurine; **31**, LDL ($C\underline{H}_3$); **32**, VLDL ($C\underline{H}_3$); **33**, Isoleucine; **34**, Leucine; **35**, Valine; **36**, D-3-Hydroxybutyrate; **37**, LDL ($C\underline{H}_2)_n$; **38**, VLDL ($C\underline{H}_2)_n$; **39**, VLDL ($C\underline{H}_2CH_2CO$); **40**, LDL ($C\underline{H}_2C{=}C$); **41**, LDL ($C\underline{H}_2CO$); **42**, Acetoacetate; **43**, Glutamate; **44**, LDL ($C{=}CC\underline{H}_2C{=}C$); **45**, Phosphocholine; **46**, Glucose; **47**, Glycine; **48**, Glyceryl of lipids ($C\underline{H}OCOR$); **49**, Unsaturated LDL ($={=}C\underline{H}CH_2CH_2$).

and the nonobese phenotypes was a significantly lower Bif164 count in the obese rats. Erec482 and Lab158 counts were significantly higher in the (*fa*/*fa*) rats compared to the (−/−) and (*fa*/−) rats, respectively. Ato291 counts were significantly lower in the (*fa*/*fa*) rats compared to the (*fa*/−) rats. In addition, significant differences were observed between the microbiota profiles of the (*fa*/−) and (−/−) strains. The (−/−) rats had significantly lower Erec482 and Lab158 counts compared to the (*fa*/−) rats. Similar Chis150 counts were observed for all three Zucker strains, while much variation and no significant strain-related differences were observed in the Bac303 counts between the different strains. A cross-validated scores plot from the OPLS-DA analysis using the FISH count information is shown in Figure 5B, which describes differences in the gut microbiota profiles. With the use of the bacterial count information, the three Zucker strains are each occupying their own space in the plot, although there was one outlier from the (−/−) strain which was closer to the (*fa*/−) strain (due to higher Bif164, Ato291, Erec482 and Bac303 counts). The plot shows that the obese (*fa*/*fa*) rats were clearly separated from the other animals due to lower Bif164 and higher Erec482 and Lab158 counts.

The DGGE profiles obtained from the feces of each individual rat are shown in Figure 6. These profiles provide a semiqualitative fingerprint of the bacteria present in the feces of the animals. It is well-known that qualitative methods such as cloning, and by extension DGGE, provide only a 'snapshot' of a bacterial community, with "*...each physical, chemical and biological step involved in the molecular analysis of an environment...a source of bias which will lead to a distorted view of the 'real world'*".[55] The advantage of DGGE over other methodologies is that it allows a relatively high throughput of large numbers of samples for a modest cost. Its disadvantage is that phylogenetically useful information cannot be retrieved from sequences of <200 nt, as is the case with the primers used herein. This is also a disadvantage with novel high-throughput sequencing (HTS) methods, from which reads of 25−250 bases are produced.[56] DGGE does, however, allow genus-level classifications to be made. While HTS methods are widely used in analyses of microbial diversity, they are generally applied to small numbers of samples, which precludes characterization of patterns of microbial diversity across space and time.[57] Tagging approaches may increase the robustness of HTS methodologies, but they still do not get around the fact that many thousands of sequences have to be collated and analyzed from several separate libraries of clones. In a study such as the one described herein, analyzing thousands of clones from each of the animals studied would be impractical. In addition, while approaches such as 454 pyrosequencing can determine the sequences of huge numbers of different DNA strands at one time, to date, it seems that these new approaches are not as yet much cheaper than electrophoresis technologies.[56] While HTS methods are ideal for sequencing whole genomes relatively cheaply, their use in ecological studies is still limited. Only one study has used fecal samples (from calves) with a HTS method (pyrosequencing).[58] Fingerprinting techniques are better suited toward ecological studies, and allow us to monitor shifts in

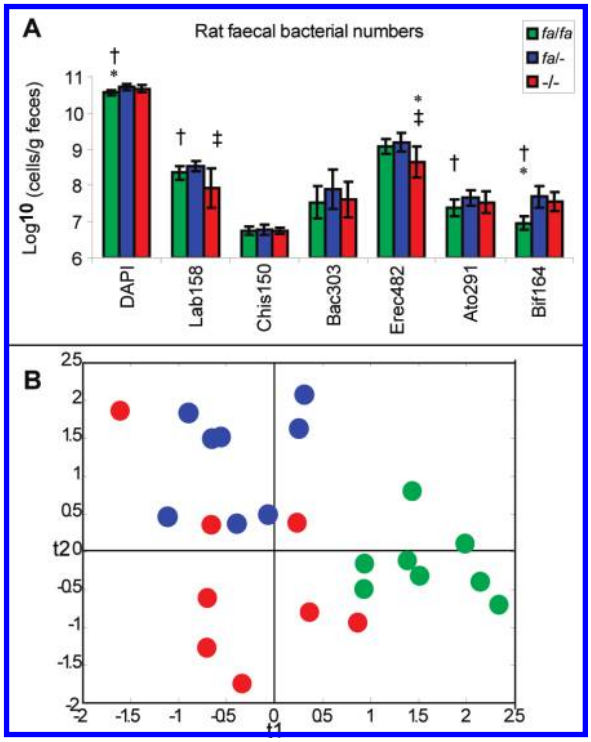

Figure 5. (A) Fecal bacterial population levels determined *via* FISH analysis for the three rat strains examined in this study. Data (n = 8 per rat strain) are expressed as mean ± standard deviation. Details for the bacterial groups covered by the probes are given in Materials and Methods. Significant differences ($P < 0.05$) in the counts between strains were determined by unpaired Student's t-test and are represented by the following; †, (*fa*/*fa*) and (*fa*/−); *, (*fa*/*fa*) and (−/−); ‡, (*fa*/−) and (−/−). (*fa*/*fa*) green square; (*fa*/-) blue square; and (−/−) red square. (B) Cross-validated scores plot from the OPLS-DA model showing clear separation of the Zucker genotypes according to fecal bacterial population levels as determined by FISH. Red circle, (−/−) lean; blue circle, (*fa*/−) lean, green circle, (*fa*/*fa*) obese.

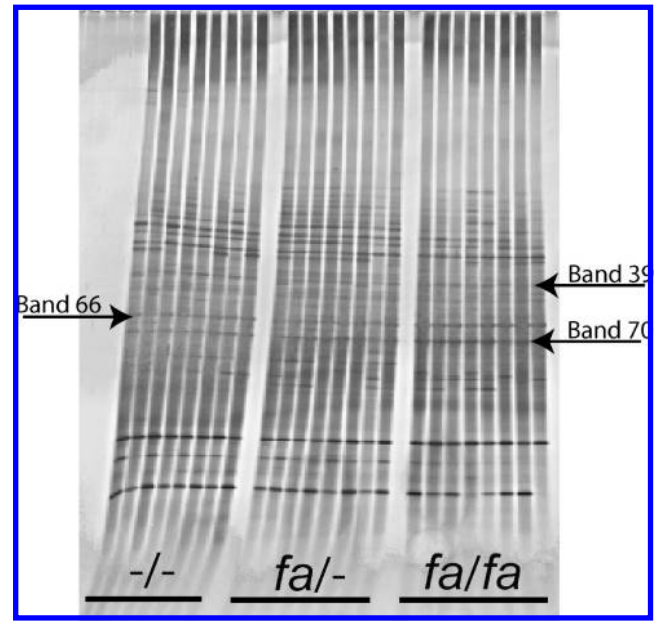

Figure 6. DGGE gel showing the bacterial profiles generated from the fecal microbiota of the three Zucker strains examined in this study. Arrows indicate bands that were shown to be of significance in relation to the different genotypes. Band 39 was unique to the (*fa*/*fa*) strain and its sequences were derived from a *Halomonas* sp. and a member of the Bacteroidetes; band 66 was most intense in the (*fa*/*fa*) strain but the two lean (*fa*/− and −/−) strains had equally less intense bands, and its sequence was derived from a *Sphingomonas* sp.; band 70 showed a gradual decrease in intensity from the (*fa*/*fa*) to (*fa*/−) and finally the lean (−/−) strain, and its sequence, was derived from a *Halomonas* sp. (99.5% sequence similarity to band 39).

bacterial populations over time and between different animals; hence, DGGE was also applied in this study.

The DGGE gel was analyzed using Syngene GeneTools, which was used to extract the bands in multiple lanes providing a semiquantitative measure of area under the curve (band density). A total of 102 bands were detected in the DGGE gel, with some bands only present in one animal and others ubiquitous. These data were used to create OPLS coefficient plots in the same way as for the ^{1}H NMR data (OPLS-DA analysis). To allow the best model, the data were filtered and only bands that appeared in at least 50% of the animals in at least one strain were retained; this resulted in 68 of the bands remaining. The cross-validated scores plot from the OPLS-DA analysis of the Zucker strains using the DGGE data showed excellent separation and good predictive ability for all three strains (Figure 2C). Several bands were seen to be causing the separation of the lean (−/−) and obese (*fa*/*fa*) strains (Supporting Information, Figure 4) and this was also true for each of the pairwise OPLS coefficient plots. OPLS-DA models generated using DGGE data were significant for all the pairwise comparisons.

To decide which bands should be examined further by cloning and 16S rRNA gene fragment sequencing, the OPLS coefficients were used to indicate the relative influence of each band on the OPLS-DA models. Three bands were excised, cloned and sequenced to give a preliminary view of the microbial species which caused the separation for the obese animals. All three bands were positively correlated with the obese (*fa*/*fa*) strain. Band 39 only appeared in the (*fa*/*fa*) strain; band 66 was intense for the obese (*fa*/*fa*) strain, but the two lean (*fa*/− and −/−) strains had equally less intense bands, and band 70 showed a gradual decrease in intensity from the (*fa*/*fa*) to (*fa*/−) and finally the lean (−/−) strain. Assessment of the intensity of the bands was based on the area under the curve for each peak. However, it should be noted that assuming that the appearance and disappearance of a band, or indeed its intensity, can be attributed to the increase or decrease in numbers of a specific type of bacterium is a conclusion that can only be reached with considerable caution, particularly when using universal primer sets to examine complex ecosystems.[59]

Multiple clones, from the bands mentioned above, were processed and each replicate sequence was subject to an WU-Blast2 search for sequence identification. Although one of the replicates of band 39 was 100% similar to *Halomonas* sp. 10003 (GenBank accession no. EU432573; a member of the class Gammaproteobacteria), another replicate was 100% similar to an uncultured member of the phylum Bacteroidetes (clone D748, GenBank accession no. AY986347). More refined sequence analyses revealed that the uncultured Bacteroidetes clone was most closely related to *Bacteroides caccae* and *Bacteroides coprosius* (~5% sequence divergence, based on a comparison of 189 nt). The finding of one band representing

more than one species of bacteria in DGGE profiles is not unique, but it is a phenomenon that is often overlooked or not reported by researchers.[59,60] It is also important to remember that sequences that are different may still result in similar denaturing properties, causing them to comigrate on a gel.[59] For example, the *Halomonas* sp. and *Bacteroides* sp. identified in band 39 were 194 and 189 nt long, respectively, and differed in their G + C contents (56.7 and 47.6 mol%, respectively), yet they migrated to the same point of the gel. Band 70 was also closely related to *Halomonas* sp. 10003 (99.5%). Comparison of the *Halomonas* sequences from bands 39 and 70 showed them to display 99.5% sequence similarity (i.e., 1 mismatch in 194 nt). It is possible that both bands are from the same bacterium, with the organism having multiple copies of the 16S rRNA gene but with sequence heterogeneities. The problems of variable copy numbers and sequence heterogeneities in relation to cloning and DGGE studies based on 16S rRNA gene sequences have been highlighted by a number of authors,[61−63] although it has been acknowledged that targeting the V3 region of the 16S rRNA gene for the DGGE studies of gastrointestinal microbiomes generates the best profiles in terms of species richness and resolution.[64] Band 66 was 100% similar to *Sphingomonas* sp. BF14 (GenBank accession no. Z23157; a member of the class Alphaproteobacteria). The sequences determined from the bands discussed in this study have been deposited with GenBank under accession numbers AM950263−AM950266.

Integration of Metabotype and Microbiome Data. To model covariation patterns between metabolic data and microbiome data, and investigate potential relationships between the urinary ^{1}H NMR metabolite data and the individual bacterial counts from FISH, OPLS models using the ^{1}H NMR data to predict the relevant bacteria were established. As an example, the ^{1}H NMR urinary data were used as the descriptor matrix (**X**) and the bacterial count for Bif164 was used as the response variable (**Y**). The bacteria targeted by this probe were chosen as an example as they had connections to the "healthier" strain and showed the most important difference for the (*fa*/*fa*) strain compared to the other two strains. The OPLS model shows the correlation between the urinary ^{1}H NMR data (descriptor matrix) and the Bif164 count from FISH (response variable) as seen in the OPLS inner relation plot (Figure 7A). The plot not only showed the correlation but also that the obese (*fa*/*fa*) animals had consistently lower Bif164 counts. The corresponding OPLS coefficient plot indicates which metabolites are correlated and which are anticorrelated with the particular bacterial group (Figure 7B). The metabolites that correlated most strongly with the Bif164 population were creatinine, hippurate and phenylacetylglycine (PAG), whereas acetate was anticorrelated.

Another method of achieving an integrated model is to use a single DGGE band as the response variable to show interactions between metabolites and bacteria. Band 70 was selected from the DGGE gel and regressed against the ^{1}H NMR urinary metabolite profiles. This band was selected as it had the highest positive correlation with the (*fa*/*fa*) strain and was successfully identified (*Halomonas* sp.). The OPLS inner relation plot is shown in Figure 8A. The corresponding OPLS coefficient plot indicates which metabolites are correlated and which are anticorrelated with band 70 (Figure 8B).

While band 70 (*Halomonas* sp.) was positively correlated to the (*fa*/*fa*) obese strain, the OPLS coefficient plot showed that there were only a few specific metabolites that were positively correlated to this band, acetate being the main covarying metabolite. However, there were many more metabolites that were anticorrelated to this band, including hippurate, phenylacetylglycine and TMAO, which are microbial cometabolites.

Since the model predicting individual DGGE bands from the urinary ^{1}H NMR data was strong, that is, Q^2 values of ∼0.55, using OPLS modeling, an O2PLS model was established to further examine if a joint model including all DGGE variables could be used for prediction modeling between the two data sets, which would indicate that consistent multivariate covariation patterns between the metabolites and bacterial strains were present. The urinary ^{1}H NMR data were used as the descriptor matrix and the filtered DGGE data used as the response matrix. The models that were created are shown in Figure 9, where the color indicates the Q^2 values (prediction performance) for each variable, peaks or bars which appear red have high predictive capability.

The ^{1}H NMR peaks which were statistically well-predicted by the DGGE data were citrate, hippurate and three sets of peaks at δ2.03, δ2.16 and δ2.26 (all multiplets) which have been tentatively identified as formiminoglutamic acid (FIGLU) based on their concordance with the literature.[65] FIGLU, citrate and hippurate had Q^2 values of 0.69, 0.40, and 0.31, respectively. Several of the DGGE bands were well-predicted by the ^{1}H NMR data. Two of the bands that appeared in red (highest correlation) had already been identified from the OPLS-DA models of the strain comparison using the DGGE data; band 66 (*Sphingomonas* sp.) had a predictive Q^2 value of 0.58 and band 70 (*Halomonas* sp.) had a predictive Q^2 value of 0.54. To further test the robustness of this method, several peaks/bands from the ^{1}H NMR and the DGGE gel were selected, with high and low Q^2 values, and both Pearson and Spearman correlations were completed. The results are consistent with the O2PLS model and are displayed in the Supporting Information in Table 1.

Discussion

Urinary Metabolic Differences Corresponding to Genotype Variation. Significant metabolic differences were established between the disease-prone Zucker (*fa*/*fa*) obese rats and the other two lean strains even at 10 weeks of age, that is, prior to the development of insulin resistance in the obese strain, but after the onset of obesity (from weaning at 4 weeks). The ^{1}H NMR spectroscopic data obtained from the analysis of urine showed major differences between strains including creatinine, TMAO and hippurate which were all lower in the urine of the (*fa*/*fa*) strain; conversely, acetate was relatively higher in the (*fa*/*fa*) obese strain. Creatinine is linked to muscle mass[66] and the lower levels observed could be a result of their excessive weight and relative lack of muscle due to low physical activity. TMAO derives both from the diet and also from microbial metabolism of choline. Hippurate is a gut microbial mammalian co-metabolite of benzoic acid which can be generated by a range of gut microbes from other low molecular weight aromatic compounds and polyphenolics in the gut and is subsequently glycine conjugated in the mitochondria and excreted in the urine.[4] Given that the food type and supply was constant, the observed strain differences in hippurate and TMAO excretion is more likely an indication of functional differences in the microbiome metabolic activity of the (*fa*/*fa*) animals.[24] Hippurate has also been linked inversely to blood pressure, suggesting a further connection with diet and obesity[16] and has been reported to be present in higher concentrations in the urine of caloric-restricted dogs in comparison with

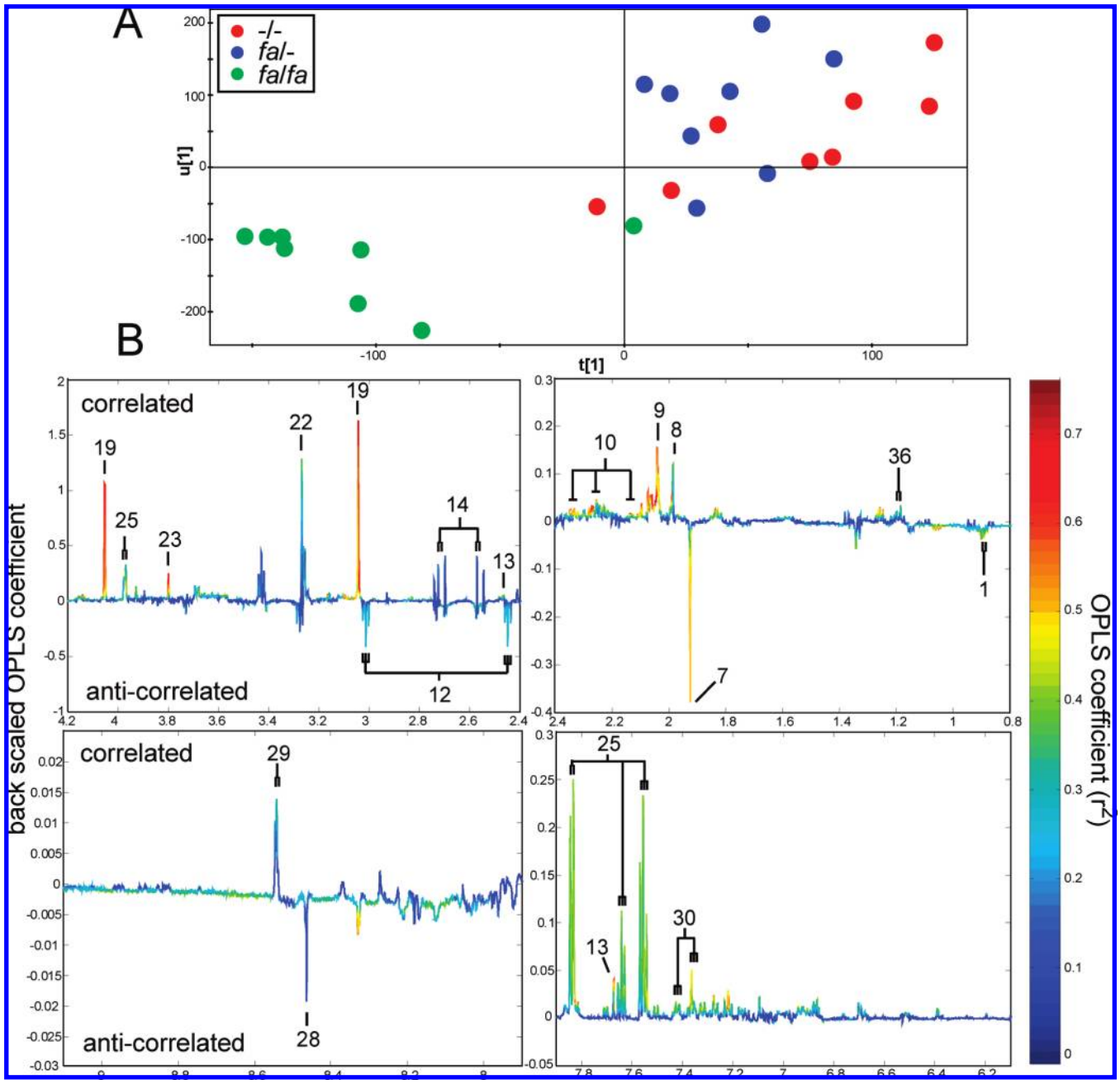

Figure 7. (A) OPLS inner relation plot showing the correlation between the urinary ^{1}H NMR spectroscopic data (**X**) and the Bifidobacteria counts (**Y**). (B) OPLS coefficient plot showing urinary ^{1}H NMR metabolites that correlate and anticorrelate to Bifidobacteria. $R^2 = 0.86$; $Q^2 = 0.55$. **1**, 3-Methylglutarate; **7**, Acetate; **8**, **9**, Glycoprotein fragments (*N*-acetyl); **10**, Formiminoglutamic acid; **12**, 2-Oxoglutarate; **13**, Pyridoxine; **14**, Citrate; **19**, Creatinine; **22**, Trimethylamine-*N*-oxide; **23**, Guanidoacetate; **25**, Hippurate; **30**, Phenylacetylglycine; **28**, Formate; **29**, unknown; **36**, D-3-Hydroxybutyrate.

their nonrestricted counterparts.[67] Therefore, the consensus from all of these studies is that urinary hippurate is positively associated with a lean phenotype.

Plasma Metabolic Differences Corresponding to Genotype Variation. For the plasma CPMG spin−echo and diffusion-edited ^{1}H NMR data, the LDL and VLDL lipid profiles were significantly different between the obese and lean animals, with these lipoproteins relatively increased in the obese animals. Here, we illustrate this using CPMG pulse sequence data which demonstrate residual contributions from the smaller, more mobile lipoprotein components. This is further supported by the OPLS-DA model of the diffusion-edited data in Figure 3 in the Supporting Information. Other strain-related variations included acetoacetate, isoleucine and leucine, all of which were relatively higher in the obese (*fa*/*fa*) animals. Acetoacetate is a ketone body formed from D-3-hydroxybutyrate, which would appear in both plasma and urine if the animal has a reduced ability to respond to insulin; elevated ketone levels are an early indicator of insulin resistance. Ketone bodies are formed by oxidizing nonesterified fatty acids (NEFAs) in the liver, the production of NEFAs are regulated by insulin.[68] These ketones are detrimental to the health of the animals if present in high concentrations. Here, isoleucine and leucine (ketogenic amino acids[54]) were observed at higher levels in the plasma in the

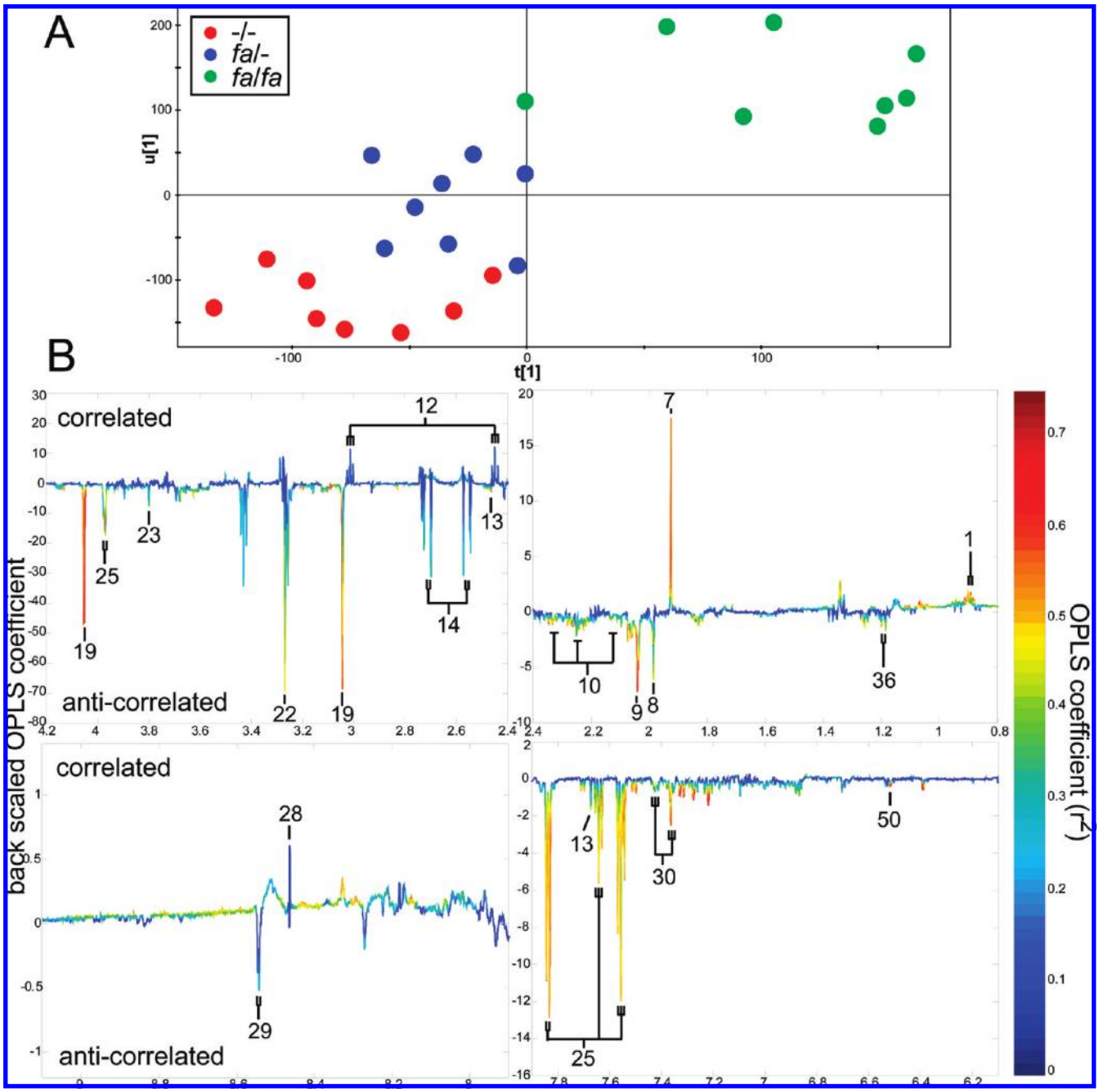

Figure 8. (A) OPLS inner relation plot showing the correlation between the urinary ^{1}H NMR spectroscopic data (**X**) and the *Halomonas* sp. (band 70) (**Y**). (B) OPLS coefficient plot showing urinary ^{1}H NMR metabolites that correlate and anticorrelate to *Halomonas* sp. (band 70). $R^2 = 0.62$; $Q^2 = 0.50$. **1**, 3-methylglutarate; **7**, Acetate; **8**, **9**, Glycoprotein fragments (*N*-acetyl); **10**, Formiminoglutamic acid; **12**, 2-oxoglutarate; **13**, Pyridoxine; **14**, Citrate; **19**, Creatinine; **22**, Trimethylamine-*N*-oxide; **23**, Guanidoacetate; **25**, Hippurate; **30**, Phenylacetylglycine; **28**, Formate; **29**, unknown; **36**, D-3-Hydroxybutyrate; **50**, Fumarate.

obese (*fa*/*fa*) rats compared to the lean (−/−) animals, and have also been shown to be high in uncontrolled human diabetics.[54] The relatively higher concentration of the ketogenic amino acids with acetoacetate in the obese animals suggests the beginning of a disruption of glucose regulation.

Microbial Differences Corresponding to Host Genotype Variation. The three Zucker strains could be clearly distinguished from each other through differences within the bacterial population levels of their dominant gut microbiota based on FISH counts and *via* DGGE profiling of the fecal microbiota. All animals had a common biological lineage and were subject to the same diet (i.e., same batches of food) and exposure to laboratory environment. Although they were housed according to phenotype after 6 weeks of age, they were kept in the same environment. The initial intestinal microbiota of infants is obtained from maternal inoculum at birth and develops as a consequence of diet and genetics.[69] The production of these rat strains by crossing (*fa*/−) parents should ensure that a similar fecal microbiota was initially conferred on all animals. Consequently, we conclude that, although the differences observed between the obese and lean animals may well be a result of differing caloric intake, the variations among the three strains may arise from host biochemistry and genetics having a role in the composition of the fecal microbiota.

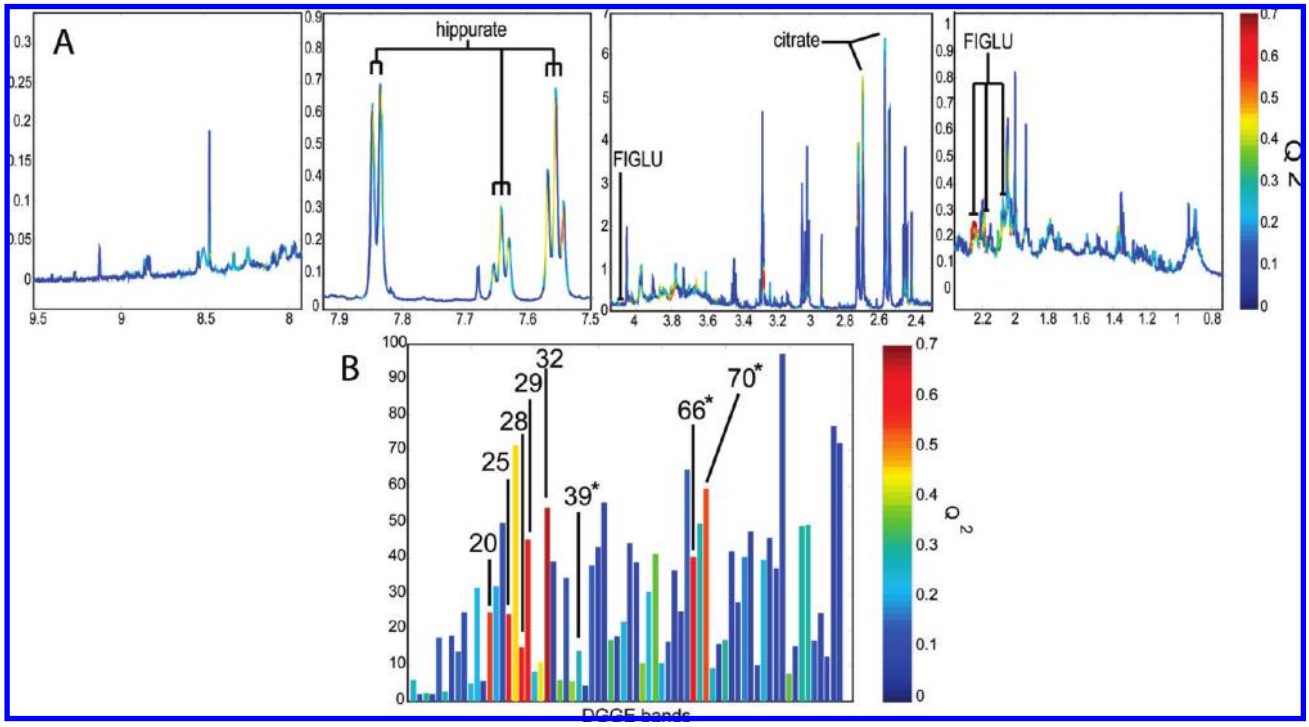

Figure 9. (A) ^{1}H NMR prediction results from the O2PLS model using the urine data. FIGLU, Formiminoglutamic acid. (B) DGGE prediction results from the O2PLS model using the bacterial band information (*, identified bands). Color corresponds to the predictive ability of the model (Q^2).

The (*fa*/*fa*) obese strain had significantly lower DAPI (total bacteria) and Bif164 counts when compared to the lean (*fa*/− and −/−) strains. Differences in bacterial populations were also observed between the two lean phenotypes, with the (−/−) rats having significantly lower Lab158 and Erec482 counts compared to the (*fa*/−) rats. The physiological status of the rats' gastrointestinal tracts is clearly genotype-dependent, and the differences observed in the fecal microbiota of the three strains studied here support previous findings,[70] that is, the physiological status of the gastrointestinal tract and not just diet has a major role in the regulation of important groups of the gastrointestinal microbiota. The finding of lower Bif164 counts in the obese rats compared with the lean rats provides the first report demonstrating a direct and inverse association between bifidobacteria and an obese phenotype. It also provides support, in relation to obesity-associated conditions, for the conclusions of Cani et al.,[71] who suggested a role for bifidobacteria in the regulation of endotoxaemia in C57b16/J mice.

The lower Bif164 counts in the (*fa*/*fa*) strain were parallel to the decrease in urinary hippurate observed in this strain, suggesting that bifidobacteria are important in distinguishing between the obese and lean strains. *Bifidobacterium* species are known to produce acetate and lactate as principal end-products of glucose fermentation, are recognized as being important members of the beneficial gut microbiota, are commonly used probiotics and are the means through which most prebiotic functional foods are proposed to mediate their health-promoting activities.[72] There are currently no published reports on the ability of bifidobacteria to produce hippurate from benzoate *via* glycine conjugation or an associated role for them influencing systemic hippurate levels and hippurate is not observed in the fecal water of any strain of animal. Comparison of the benzoic acid-producing capabilities of strains of bifidobacteria with other intestinal bacteria would provide information on those members of the fecal microbiota responsible for benzoate products and/or its metabolic intermediaries.

While recent studies have shown the importance of Bacteroidetes/Firmicutes ratio in obesity,[73] our data broadly support these findings, and additionally, we provide analyses at a lower taxonomic level (genus) indicating that bifidobacteria may also play a role in protection against obesogenesis. The bifidobacteria are rarely considered in such global studies because they are not usually recovered in metagenomic PCR-based clone libraries employing "universal bacterial" 16S rRNA gene primers. Various reasons have been provided for this observation, including selection of PCR primers (in relation to efficiency and target) and cycle conditions.[74,75] Genome size and the copy number of the 16S rRNA gene also vary between species and even strains of the same species, confounding the issue further.[55,76] It is important to consider the DGGE profile, as well as the FISH data, as it attempts to map the entire colonization found in feces, although the same biasing against bifidobacteria is encountered as for clone libraries. The DGGE profile indicates that there are differences in the other bacteria present in the gut. The relative intensities of the bands from the DGGE profiles are also capable of completely separating each strain, providing a useful tool to microbiotype the three strains of animals.

Three bands (39, 66 and 70) selected from the OPLS-DA analyses of DGGE data were cloned. Two bands were successfully identified as a *Halomonas* sp. (band 70) and a *Sphingomonas* sp. (band 66). Band 70 (*Halomonas* sp.), which was present in the DGGE profiles of all animals, was the most significantly different when comparing the lean (−/−) and the obese (*fa*/*fa*) strains, with a relatively denser band in the obese animals. *Halomonas aquamarina* has been detected *via* a cloning study in pigs[77] and '*Halomonas phocaeensis*' has been the cause of an outbreak of

bacteraemia in a neonatal intensive care unit.[78] *Halomonas magadiensis* has been shown to have antagonistic activity against lipopolysaccharide (LPS) produced from *Escherichia coli*, with this LPS having been shown to be associated with endotoxic shock in the human enteric system.[79] Recent work identified *Halomonas* spp. as important components of acetate-fed sludge reactors.[80] Thus, the relative acetate levels of the different rat strains may be responsible for the observed distinction in the density of band 70, with the obese strain associated with higher acetate levels. The current study is the first to suggest an association between *Halomonas* spp. and obesity. With regard to band 66, the genus *Sphingomonas* includes important xenobiotic-degrading bacteria that are sometimes associated with human infections[81] and immune responses.[82] *Sphingomonas* spp. have also been isolated from the oral cavity of mice[83] and, together with *Bacteroides thetaiotaomicron* and other bacterial species, are considered to be efficient carbohydrate scavengers (*via* TonB-dependent receptors).[84]

Metagenomic−Metabolic Data Integration. An OPLS analysis of the urinary ^{1}H NMR data and a single band (70) revealed that several microbiota co-metabolites, hippurate, phenylacetylglycine, TMAO, are anticorrelated to the detected *Halomonas* sp. The O2PLS analysis integrating the urinary ^{1}H NMR data and the filtered DGGE data provided insights into which metabolites were associated to the microbiota. Unsurprisingly, it revealed that the hippurate concentration in the urine, as represented by the ^{1}H NMR data, was well-predicted by the microbial data, especially bands 28, 29, 32, and 66 (*Sphingomonas* sp.). These results could be further tested by introducing inoculums obtained from the three different Zucker strains into germfree animals.

An OPLS model assessing the multivariate associations of the ^{1}H NMR urinary data and the Bif164 counts (from FISH) revealed a good relationship. It was also illustrated that the obese (*fa*/*fa*) strain had lower Bif164 counts, together with a relatively lower hippurate concentration and a higher acetate concentration. The concentration of acetate is likely to be attributed to other bacteria present in the obese animals, as bifidobacteria are acetate producers.[85]

The O2PLS model using the urinary ^{1}H NMR data and the filtered DGGE density data allowed visualization of several possible associations between the two data sets. This model supported many of the results already discussed, that is, the association of urinary hippurate with several microbial species and that *Halomonas* sp. (band 70) and *Sphingomonas* sp. (band 66) are closely associated with the urinary ^{1}H NMR data.

Conclusions

Classification of the metabolic phenotypes for lean (−/−) and obese (*fa*/*fa*) strains was readily achieved using ^{1}H NMR spectra. However, the two lean strains (−/− and *fa*/−) could not be easily discriminated. Interestingly, despite being metabolically similar, the (*fa*/−) and the (−/−) leans rats differed in their microbiomes to the extent that the DGGE profile could be used to robustly predict a separation, suggesting that the microbial differences did not significantly impact on the host end-point metabolic status. Fecal bacterial counting using FISH also achieved good separation between all three strains using seven FISH probes. The (*fa*/*fa*) obese rat was unlike the other animals in this study, separating in each of the three methods, that is, ^{1}H NMR, FISH and DGGE, from the other two strains of Zucker rat. Modeling the metabolic and microbial covariation using the O2PLS method identified correlations between specific DGGE bands and several urinary variables, which gave insight into possible interactions between the host and the microbiome. It is not clear how the differences in microbiotype between the (*fa*/*fa*) obese animals and the lean rats relate to host genotype variation. However, both the metabolic and metagenomic data sets show obvious associations with lean and obese phenotypes, consistent with previous observations on obese (*ob*/*ob*) mouse models.[86] It now appears that even phenotypically silent gene or SNP variations (e.g., (*fa*/−) and (−/−)) may have occult host phenotype physiological influences that give rise to microbiome population selection and/or activity variations which may have wider implications when considering interspecies and interstrain variations in disease development and responses to therapies and their relevance to problems in personalized healthcare.[4,5,87,88] It also opens up new discovery biology opportunities for finding novel functional members of the microbiome that connect to host metabolism in disease states and as we have recently suggested the possibility of treating or selectively drugging the microbiome with its large number of targets to achieve personalized healthcare benefits.[19]

Acknowledgment. The authors wish to thank the staff at the ABU in Alderley-Park, Nestlé for funding Yulan Wang and AstraZeneca for its continued financial support. Lesley Hoyles is thanked for processing and running the samples for DGGE and for sequencing and analyzing the bands.

Supporting Information Available: Table showing both Pearson correlations and Spearman correlations for important metabolites and bands from the O2PLS analysis; examples of ^{1}H NMR spectra produced from each pulse program from the plasma samples; typical plasma CPMG ^{1}H NMR for each Zucker strain; OPLS coefficient plot comparing the (−/−) lean strain and the (*fa*/*fa*) obese strain from the diffusion-edited plasma ^{1}H NMR data; OPLS coefficient plot comparing the (−/−) lean strain and the (*fa*/*fa*) obese strain from the DGGE profile. This material is available free of charge *via* the Internet at http://pubs.acs.org.

References

(1) Gill, S. R.; Pop, M.; DeBoy, R. T.; Eckburg, P. B.; Turnbaugh, P. J.; Samuel, B. S.; Gordon, J. I.; Relman, D. A.; Fraser-Liggett, C. M.; Nelson, K. E. Metagenomic analysis of the human distal gut microbiome. *Science* **2006**, *312* (5778), 1355–1359.
(2) Kirjavainen, P. V.; Gibson, G. R. Healthy gut microflora and allergy: factors influencing development of the microbiota. *Ann. Med.* **1999**, *31* (4), 288–292.
(3) Lederberg, J. Infectious history. *Science* **2000**, *288* (5464), 287–293.
(4) Nicholson, J. K.; Holmes, E.; Wilson, I. D. Gut microorganisms, mammalian metabolism and personalized health care. *Nat. Rev. Microbiol.* **2005**, *3* (5), 431–438.
(5) Nicholson, J. K.; Wilson, I. D. Understanding 'global' systems biology: Metabonomics and the continuum of metabolism. *Nat. Rev. Drug Discovery* **2003**, *2* (8), 668–676.
(6) Li, M.; Wang, B.; Zhang, M.; Rantalainen, M.; Wang, S.; Zhou, H.; Zhang, Y.; Shen, J.; Pang, X.; Zhang, M.; Wei, H.; Chen, Y.; Lu, H.; Zuo, J.; Su, M.; Qiu, Y.; Jia, W.; Xiao, C.; Smith, L. M.; Yang, S.; Holmes, E.; Tang, H.; Zhao, G.; Nicholson, J. K.; Li, L.; Zhao, L. Symbiotic gut microbes modulate human metabolic phenotypes. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (6), 2117–2122.
(7) Martin, F. P. J.; Verdu, E. F.; Wang, Y. L.; Dumas, M. E.; Yap, I. K. S.; Cloarec, O.; Bergonzelli, G. E.; Corthesy-Theulaz, I.; Kochhar, S.; Holmes, E.; Lindon, J. C.; Collins, S. M.; Nicholson, J. K. Transgenomic metabolic interactions in a mouse disease model: Interactions of Trichinella spiralis infection with dietary Lactobacillus paracasei supplementation. *J. Proteome Res.* **2006**, *5* (9), 2185–2193.
(8) Wang, Y. L.; Holmes, E.; Nicholson, J. K.; Cloarec, O.; Chollet, J.; Tanner, M.; Singer, B. H.; Utzinger, J. Metabonomic investigations in mice infected with *Schistosoma mansoni*: An approach for biom-

arker identification. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (34), 12676–12681.
(9) Wang, Y.; Utzinger, J.; Saric, J.; Li, J. V.; Burckhardt, J.; Dirnhofer, S.; Nicholson, J. K.; Singer, B. H.; Brun, R.; Holmes, E. Global metabolic responses of mice to *Trypanosoma brucei brucei* infection. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (16), 6127–6132.
(10) Dumas, M. E.; Barton, R. H.; Toye, A.; Cloarec, O.; Blancher, C.; Rothwell, A.; Fearnside, J.; Tatoud, R.; Blanc, V.; Lindon, J. C.; Mitchell, S. C.; Holmes, E.; McCarthy, M. I.; Scott, J.; Gauguier, D.; Nicholson, J. K. Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (33), 12511–12516.
(11) Dumas, M. E.; Wilder, S. P.; Bihoreau, M. T.; Barton, R. H.; Fearnside, J. F.; Argoud, K.; D'Amato, L.; Wallis, R. H.; Blancher, C.; Keun, H. C.; Baunsgaard, D.; Scott, J.; Sidelmann, U. G.; Nicholson, J. K.; Gauguier, D. Direct quantitative trait locus mapping of mammalian metabolic phenotypes in diabetic and normoglycemic rat models. *Nat. Genet.* **2007**, *39* (5), 666–672.
(12) Ley, R. E.; Turnbaugh, P. J.; Klein, S.; Gordon, J. I. Microbial ecology—Human gut microbes associated with obesity. *Nature* **2006**, *444* (7122), 1022–1023.
(13) Martin, F. P. J.; Dumas, M. E.; Wang, Y. L.; Legido-Quigley, C.; Yap, I. K. S.; Tang, H. R.; Zirah, S.; Murphy, G. M.; Cloarec, O.; Lindon, J. C.; Sprenger, N.; Fay, L. B.; Kochhar, S.; van Bladeren, P.; Holmes, E.; Nicholson, J. K. A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model. *Mol. Syst. Biol.* **2007**, *3*, 112.
(14) Martin, F. P. J.; Wang, Y. L.; Sprenger, N.; Holmes, E.; Lindon, J. C.; Kochhar, S.; Nicholson, J. K. Effects of probiotic *Lactobacillus paracasei* treatment on the host gut tissue metabolic profiles probed via magic-angle-spinning NMR spectroscopy. *J. Proteome Res.* **2007**, *6* (4), 1471–1481.
(15) Dumas, M. E.; Maibaum, E. C.; Teague, C.; Ueshima, H.; Zhou, B. F.; Lindon, J. C.; Nicholson, J. K.; Stamler, J.; Elliott, P.; Chan, Q.; Holmes, E. Assessment of analytical reproducibility of 1H NMR spectroscopy based metabonomics for large-scale epidemiological research: the INTERMAP study. *Anal. Chem.* **2006**, *78* (7), 2199–2208.
(16) Holmes, E.; Loo, R. L.; Stamler, J.; Bictash, M.; Yap, I. K. S.; Chan, Q.; Ebbels, T.; De Iorio, M.; Brown, I. J.; Veselkov, K.; Daviglus, M.; Kestleloot, H.; Ueshima, H.; Zhao, L.; Nicholson, J. K.; Elliot, P. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* **2008**, *453* (7193), 396–400.
(17) Rezzi, S.; Ramadan, Z.; Fay, L. B.; Kochhar, S. Nutritional metabonomics: Applications and perspectives. *J. Proteome Res.* **2007**, *6* (2), 513–525.
(18) Rezzi, S.; Ramadan, Z.; Martin, F. P. J.; Fay, L. B.; van Bladeren, P.; Lindon, J. C.; Nicholson, J. K.; Kochhar, S. Human metabolic phenotypes link directly to specific dietary preferences in healthy individuals. *J. Proteome Res.* **2007**, *6*, 4469–4477.
(19) Wei, J.; Houkai, L.; Liping, Z.; Nicholson, J. K. Gut microbiota: a potential new territory for drug targeting. *Nat. Rev. Drug Discovery* **2008**, *7*, 123–129.
(20) Manichanh, C.; Rigottier-Gois, L.; Bonnaud, E.; Gloux, K.; Pelletier, E.; Frangeul, L.; Nalin, R.; Jarrin, C.; Chardon, P.; Marteau, P.; Roca, J.; Dore, J. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* **2006**, *55* (2), 205–211.
(21) Kava, R.; Greenwood, M. R. C.; Johnson, P. R. Zucker (*fa/fa*) rat. *ILAR J.* **1990**, *32*, 3.
(22) Friedman, J. M.; Halaas, J. L. Leptin and the regulation of body weight in mammals. *Nature.* **1998**, *395* (22), 763–770.
(23) Burguera, B.; Couce, M. E.; Curran, G. L.; Jensen, M. D.; Lloyd, R. V.; Cleary, M. P.; Poduslo, J. F. Obesity is associated with a decreased leptin transport across the blood-brain barrier in rats. *Diabetes* **2000**, *49* (7), 1219–1223.
(24) Williams, R. E.; Lenz, E. M.; Rantalainen, M.; Willson, I. D. The comparative metabonomics of age-related changes in the urinary composition of male Wistar-derived and Zucker (fa/fa) obese rats. *Mol. BioSyst.* **2006**, *2* (3−4), 193–202.
(25) Nicholson, J. K.; Timbrell, J. A.; Sadler, P. J. Proton NMR spectra of urine as indicators of renal damage: Mercury-induced nephrotoxicity in rats. *Mol. Pharmacol.* **1985**, *27* (6), 644–651.
(26) Nicholson, J. K.; Wilson, I. D. High resolution proton NMR spectroscopy of biological fluids. *Prog. Nucl. Magn. Reson. Spectrosc.* **1989**, *29*, 449–501.
(27) Nicholson, J. K.; Lindon, J. C.; Holmes, E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **1999**, *29* (11), 1181–1189.
(28) Wilson, I. D.; Nicholson, J. K.; Castro-Perez, J.; Granger, J. H.; Johnson, K. A.; Smith, B. W.; Plumb, R. S. High resolution "Ultra performance" liquid chromatography coupled to oa-TOF mass spectrometry as a tool for differential metabolic pathway profiling in functional genomic studies. *J. Proteome Res.* **2005**, *4* (2), 591–598.
(29) Zhang, Q.; Wang, G. J.; Du, Y.; Zhu, L. L.; Jiye, A. GC/MS analysis of the rat urine for metabonomic research. *J. Chromatogr., B:Anal. Technol. Biomed. Life Sci.* **2007**, *854* (1−2), 20–25.
(30) Granger, J. H.; Williams, R.; Lenz, E. M.; Plumb, R. S.; Stumpf, C. L.; Wilson, I. D. A metabonomic study of strain- and age-related differences in the Zucker rat. *Rapid Commun. Mass Spectrom.* **2007**, *21* (13), 2039–2045.
(31) Gavaghan, C. L.; Holmes, E.; Lenz, E.; Wilson, I. D.; Nicholson, J. K. An NMR-based metabonomic approach to investigate the biochemical consequences of genetic strain differences: application to the C57BL10J and Alpk: ApfCD mouse. *FEBS Lett.* **2000**, *484* (3), 169–174.
(32) Williams, R. E.; Lenz, E. A.; Wilson, A. J.; Granger, J.; Wilson, I. D.; Major, H.; Stumpf, C. L.; Plumb, R. A multi-analytical platform approach to the metabonomic analysis of plasma from normal and zucker (fa/fa) obese rats. *Mol. BioSyst.* **2006**, *2* (1), 174–183.
(33) Muyzer, G.; de Waal, E. C.; Uitterlinden, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **1993**, *59* (3), 695–700.
(34) Harmsen, H. J. M.; Gibson, G. R.; Elfferich, P.; Raangs, G. C.; Wildeboer-Veloo, A. C. M.; Argaiz, A.; Roberfroid, M. B.; Welling, G. W. Comparison of viable cell counts and fluorescence *in situ* hybridization using specific rRNA-based probes for the quantification of human fecal bacteria. *FEMS Microbiol. Lett.* **2000**, *183* (1), 125–129.
(35) Neuhaus, D.; Ismail, I. M.; Chung, C. W. "FLIPSY"—A new solvent-suppression sequence for nonexchanging solutes offering improved integral accuracy relative to 1D NOESY. *J. Magn. Reson.* **1996**, *118* (2), 256–263.
(36) Carr, H. Y.; Purcell, E. M. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* **1954**, *94* (3), 630–638.
(37) Meiboom, S.; Gill, D. Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* **1958**, *29*, 688–691.
(38) Liu, M. L.; Nicholson, J. K.; London, J. C. High-resolution diffusion and relaxation edited one- and two-dimensional 1H NMR spectroscopy of biological fluids. *Anal. Chem.* **1996**, *68* (19), 3370–3376.
(39) Liu, M. L.; Nicholson, J. K.; Parkinson, J. A.; Lindon, J. C. Measurement of biomolecular diffusion coefficients in blood plasma using two-dimensional ^{1}H-^{1}H diffusion-edited total-correlation NMR spectroscopy. *Anal. Chem.* **1997**, *69* (8), 1504–1509.
(40) Sanguinetti, C. J.; Dias Neto, E.; Simpson, A. J. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *BioTechniques* **1994**, *17* (5), 914–21.
(41) Langendijk, P. S.; Schut, F.; Jansen, G. J.; Raangs, G. C.; Kamphuis, G. R.; Wilkinson, M. H. F.; Welling, G. W. Quantitative fluorescence *in-situ* hybridization of *Bifidobacterium* spp. with genus-specific 16s ribosomal-RNA-targeted probes and its application in fecal samples. *Appl. Environ. Microbiol.* **1995**, *61* (8), 3069–3075.
(42) Manz, W.; Amann, R.; Ludwig, W.; Vancanneyt, M.; Schleifer, K. H. Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. *Microbiology (Reading, U.K.)* **1996**, *142*, 1097–1106.
(43) Franks, A. H.; Harmsen, H. J. M.; Raangs, G. C.; Jansen, G. J.; Schut, F.; Welling, G. W. Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-Targeted oligonucleotide probes. *Appl. Environ. Microbiol.* **1998**, *64* (9), 3336–3345.
(44) Harmsen, H. J. M.; Elfferich, P.; Schut, F.; Welling, G. W. A 16S rRNA-targeted probe for detection of lactobacilli and enterococci in faecal samples by fluorescent *in situ* hybridization. *Microb. Ecol. Health Dis.* **1999**, *11*, 3–12.
(45) Harmsen, H. J. M.; Wildeboer-Veloo, A. C. M.; Grijpstra, J.; Knol, J.; Degener, J. E.; Welling, G. W. Development of 16S rRNA-based probes for the Coriobacterium group and the Atopobium cluster and their application for enumeration of Coriobacteriaceae in human feces from volunteers of different age groups. *Appl. Environ. Microbiol.* **2000**, *66* (10), 4523–4527.
(46) Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* **2002**, *16* (3), 119–128.
(47) Cloarec, O.; Dumas, M. E.; Trygg, J.; Craig, A.; Barton, R. H.; Lindon, J. C.; Nicholson, J. K.; Holmes, E. Evaluation of the orthogonal projection on latent structure model limitations caused by chemical shift variability and improved visualization of biomarker changes in ^{1}H NMR spectroscopic metabonomic studies. *Anal. Chem.* **2005**, *77* (2), 517–526.

(48) Bylesjo, M.; Rantalainen, M.; Cloarec, O.; Nicholson, J. K.; Holmes, E.; Trygg, J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemom.* **2006**, *20* (8−10), 341–351.
(49) Veselkov, K. A.; Lindon, J. C.; Ebbels, T. M.; Crockford, D.; Volynkin, V. V.; Holmes, E.; Davies, D. B.; Nicholson, J. K. Recursive segment-wise peak alignment of biological ^{1}H NMR spectra for improved metabolic biomarker recovery. *Anal. Chem.* **2009**, *81* (1), 56–66.
(50) Trygg, J. O2-PLS for qualitative and quantitative analysis in multivariate calibration. *J. Chemom.* **2002**, *16* (6), 283–293.
(51) Eriksson, L.; Johanaaon, E.; Kettaneh-Wold, N.; Wold, S. *Multi- and Megavariate Data Analysis Principles and Applications*; Umetrics: Umea, 2001.
(52) Williams, R. E.; Lenz, E. A.; Evans, J. A.; Wilson, I. D.; Granger, J. H.; Plumb, R. S.; Stumpf, C. L. A combined 1H NMR and HPLC-MS-based metabonomic study of urine from obese (*fa/fa*) Zucker and normal Wistar-derived rats. *J. Pharm. Biomed. Anal.* **2005**, *38* (3), 465–471.
(53) Williams, R. E.; Lenz, E. M.; Lowden, J. S.; Rantalainen, M.; Wilson, I. D. The metabonomics of aging and development in the rat: an investigation into the effect of age on the profile of endogenous metabolites in the urine of male rats using 1H NMR and HPLC-TOF MS. *Mol. BioSyst.* **2005**, *1* (2), 166–175.
(54) Nicholson, J. K.; O'Flynn, M. P.; Sadler, P. J.; Macleod, A. F.; Juul, S. M.; Sonksen, P. H. Proton NMR studies of serum, plasma and urine from fasting normal and diabetic subjects. *Biochem. J.* **1984**, *217* (2), 365–375.
(55) von Wintzingerode, F.; Gobel, U. B.; Stackebrandt, E. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* **1997**, *21* (3), 213–229.
(56) Hertl, D. G.; Fredlake, C. P.; Barron, A. E. Advantages and limitations of next-generation sequencing technologies: a comparison of electrophoresis and non-electrophoresis methods. *Electrophoresis* **2008**, *29*, 4618–4626.
(57) Taylor, D. L.; Booth, M. G.; Mcfarland, J. W.; Herriott, I. C.; Lennon, N. J.; Nusbaum, C.; Marr, T. G. Increasing ecological inference from high throughput sequencing of fungi in the environment through a tagging approach. *Mol. Ecol. Resour.* **2008**, *8* (4), 742–752.
(58) Dowd, S. E.; Callaway, T. R.; Wolcott, R. D.; Sun, Y.; McKeehan, T.; Hagevoort, R. G.; Edrington, T. S. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiol.* **2008**, *8*, 125.
(59) Abecia, L.; Fondevila, M.; Balcells, J.; Edwards, J. E.; Newbold, C. J.; McEwan, N. R. Effect of antibiotics on the bacterial population of the rabbit caecum. *FEMS Microbiol. Lett.* **2007**, *272* (2), 144–153.
(60) Kocherginskaya, S. A.; Aminov, R. I.; White, B. A. Analysis of the rumen bacterial diversity under two different diet conditions using denaturing gradient gel electrophoresis, random sequencing, and statistical ecology approaches. *Anaerobe* **2001**, *7* (3), 119–134.
(61) Case, R. J.; Boucher, Y.; Dahllof, I.; Holmstrom, C.; Doolittle, W. F.; Kjelleberg, S. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.* **2007**, *73* (1), 278–288.
(62) Crosby, L. D.; Criddle, C. S. Understanding bias in microbial community analysis techniques due to rrn operon copy number heterogeneity. *BioTechniques* **2003**, *34* (4), 790+.
(63) Dahllof, I.; Baillie, H.; Kjelleberg, S. rpoB-based microbial community analysis avoids limitations inherent in 16S rRNA gene intraspecies heterogeneity. *Appl. Environ. Microbiol.* **2000**, *66* (8), 3376–3380.
(64) Yu, Z. T.; Morrison, M. Comparisons of different hypervariable regions of rrs genes for use in fingerprinting of microbial communities by PCR-denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* **2004**, *70* (8), 4800–4806.
(65) Perry, T. L.; Applegarth, D. A.; Evans, M. E.; Hansen, S.; Jellum, E. Metabolic studies of a family with massive formiminoglutamic aciduria. *Pediatr. Res.* **1975**, *9* (3), 117–122.
(66) Rikimaru, T.; Oozeki, T.; Ichikawa, M.; Ebisawa, H.; Fujita, Y. Comparisons of urinary creatinine, skeletal-muscle mass, and indexes of muscle protein catabolism in rats fed ad-libitum, with restricted food-intake, and deprived of food. *J. Nutr. Sci. Vitaminol.* **1989**, *35* (3), 199–209.
(67) Wang, Y. L.; Lawler, D.; Larson, B.; Ramadan, Z.; Kochhar, S.; Holmes, E.; Nicholson, J. K. Metabonomic investigations of aging and caloric restriction in a life-long dog study. *J. Proteome Res.* **2007**, *6* (5), 1846–1854.
(68) Holt, R. I. G.; Hanley, N. A., *Essential Endocrinology and Diabetes*, 5th ed.; Blackwell Publishing Ltd: Malden, MA, 2007.
(69) Roger, L. C.; McCartney, A. L., Prebiotics and the infant microbiota. In *Probiotics & Prebiotics: Scientific Aspects*; Tannock, G. W., Ed.; Caister Academic Press: Norfolk, U.K., 2005; pp 195−216.
(70) Lesniewska, V.; Rowland, I.; Cani, P. D.; Neyrinck, A. M.; Delzenne, N. M.; Naughton, P. J. Effect on components of the intestinal microflora and plasma neuropeptide levels of feeding *Lactobacillus delbrueckii*, *Bifidobacterium lactis*, and inulin to adult and elderly rats. *Appl. Environ. Microbiol.* **2006**, *72* (10), 6533–6538.
(71) Cani, P. D.; Neyrinck, A. M.; Fava, F.; Knauf, C.; Burcelin, R. G.; Tuohy, K. M.; Gibson, G. R.; Delzenne, N. M. Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. *Diabetologia* **2007**, *50* (11), 2374–2383.
(72) Tuohy, K. M.; Probert, H. M.; Smejkal, C. W.; Gibson, G. R. Using probiotics and prebiotics to improve gut health. *Drug Discovery Today* **2003**, *8* (15), 692–700.
(73) Turnbaugh, P. J.; Ley, R. E.; Mahowald, M. A.; Magrini, V.; Mardis, E. R.; Gordon, J. I. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **2006**, *444* (7122), 1027–1031.
(74) Bonnet, R.; Suau, A.; Dore, J.; Gibson, G. R.; Collins, M. D. Differences in rDNA libraries of faecal bacteria derived from 10- and 25-cycle PCRs. *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 757–763.
(75) Suau, A.; Bonnet, R.; Sutren, M.; Godon, J. J.; Gibson, G. R.; Collins, M. D.; Dore, J. Direct analyses of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl. Environ. Microbiol.* **1999**, *65*, 4799–4807.
(76) Nubel, U.; Engelen, B.; Felske, A.; Snaidr, J.; Wieshuber, A. I.; Amann, R.; Ludwig, W.; Backhaus, H. Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. *J. Bacteriol.* **1996**, *178*, 5636–5643.
(77) Leser, T. D.; Amenuvor, J. Z.; Jensen, T. K.; Lindecrona, R. H.; Boye, M.; Moller, K. Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. *Appl. Environ. Microbiol.* **2002**, *68* (2), 673–690.
(78) Berger, P.; Barguellil, F.; Raoult, D.; Drancourt, M. An outbreak of *Halomonas phocaeensis* sp nov bacteraemia in a neonatal intensive care unit. *J. Hosp. Infect.* **2007**, *67* (1), 79–85.
(79) Ialenti, A.; Di Meglio, P.; Grassia, G.; Maffia, P.; Di Rosa, M.; Lanzetta, R.; Molinaro, A.; Silipo, A.; Grant, W.; Ianaro, A. A novel lipid A from Halomonas magadiensis inhibits enteric LPS-induced human monocyte activation. *Eur. J. Immunol.* **2006**, *36* (2), 354–360.
(80) Osaka, T.; Shirotani, K.; Yoshie, S.; Tsuneda, S. Effects of carbon source on denitrification efficiency and microbial community structure in a saline wastewater treatment process. *Water Res.* **2008**, *42* (14), 3709–3718.
(81) Hsueh, P. R.; Teng, L. J.; Yang, P. C.; Chen, Y. C.; Pan, H. J.; Ho, S. W.; Luh, K. T. Nosocomial infections caused by *Sphingomonas paucimobilis*: Clinical features and microbiological characteristics. *Clin. Infect. Dis.* **1998**, *26* (3), 676–681.
(82) Kinjo, Y.; Wu, D.; Kim, G. S.; Xing, G. W.; Poles, M. A.; Ho, D. D.; Tsuji, M.; Kawahara, K.; Wong, C. H.; Kronenberg, M. Recognition of bacterial glycosphingolipids by natural killer T cells. *Nature* **2005**, *434* (7032), 520–525.
(83) Lee, B.; Bowden, G. H.; Myal, Y. Identification of mouse submaxillary gland protein in mouse saliva and its binding to mouse oral bacteria. *Arch. Oral Biol.* **2002**, *47* (4), 327–332.
(84) Blanvillain, S.; Meyer, D.; Boulanger, A.; Lautier, M.; Guynet, C.; Denance, N.; Vasse, J.; Lauber, E.; Arlat, M. Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. *PLoS ONE* **2007**, *2*, e224.
(85) Gibson, G. R.; Wang, X. Regulatory effects of bifidobacteria on the growth of other colonic bacteria. *J. Appl. Bacteriol.* **1994**, *77* (4), 412–420.
(86) Ley, R. E.; Backhed, F.; Turnbaugh, P.; Lozupone, C. A.; Knight, R. D.; Gordon, J. I. Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (31), 11070–11075.
(87) Clayton, T. A.; Lindon, J. C.; Cloarec, O.; Antti, H.; Charuel, C.; Hanton, G.; Provost, J. P.; Le Net, J. L.; Baker, D.; Walley, R. J.; Everett, J. R.; Nicholson, J. K. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature* **2006**, *440* (7087), 1073–1077.
(88) Nicholson, J. K.; Holmes, E.; Lindon, J. C.; Wilson, I. D. The challenges of modeling mammalian biocomplexity. *Nat. Biotechnol.* **2004**, *22* (10), 1268–1274.

PR8009885