

## Human Metabolic Phenotypes Link Directly to Specific Dietary Preferences in Healthy Individuals

Serge Rezzi,<sup>†</sup> Ziad Ramadan,<sup>†</sup> François-Pierre J. Martin,<sup>†</sup> Laurent B. Fay,<sup>†</sup> Peter van Bladeren,<sup>†</sup> John C. Lindon,<sup>‡</sup> Jeremy K. Nicholson,<sup>‡</sup> and Sunil Kochhar<sup>\*,†</sup>

*BioAnalytical Science, Metabonomics & Biomarkers, Nestlé Research Center, P.O. Box 44, CH-1000 Lausanne 26, Switzerland, and Department of Biomolecular Medicine, Division of Surgery, Oncology, Reproductive Biology and Anaesthetics, Faculty of Medicine, Sir Alexander Fleming Building, Imperial College London, London SW7 2AZ, U.K.*

Received July 12, 2007

Individual human health is determined by a complex interplay between genes, environment, diet, lifestyle, and symbiotic gut microbial activity. Here, we demonstrate a new “nutrimetabonomic” approach in which spectroscopically generated metabolic phenotypes are correlated with behavioral/psychological dietary preference, namely, “chocolate desiring” or “chocolate indifferent”. Urinary and plasma metabolic phenotypes are characterized by differential metabolic biomarkers, measured using <sup>1</sup>H NMR spectroscopy, including the postprandial lipoprotein profile and gut microbial co-metabolism. These data suggest that specific dietary preferences can influence basal metabolic state and gut microbiome activity that in turn may have long-term health consequences to the host. Nutrimetabonomics appears as a promising approach for the classification of dietary responses in populations and personalized nutritional management.

**Keywords:** chemometrics • gut microbial cometabolism • <sup>1</sup>H NMR spectroscopy • lipoproteins • metabolic imprinting • nutritional metabonomics

### Introduction

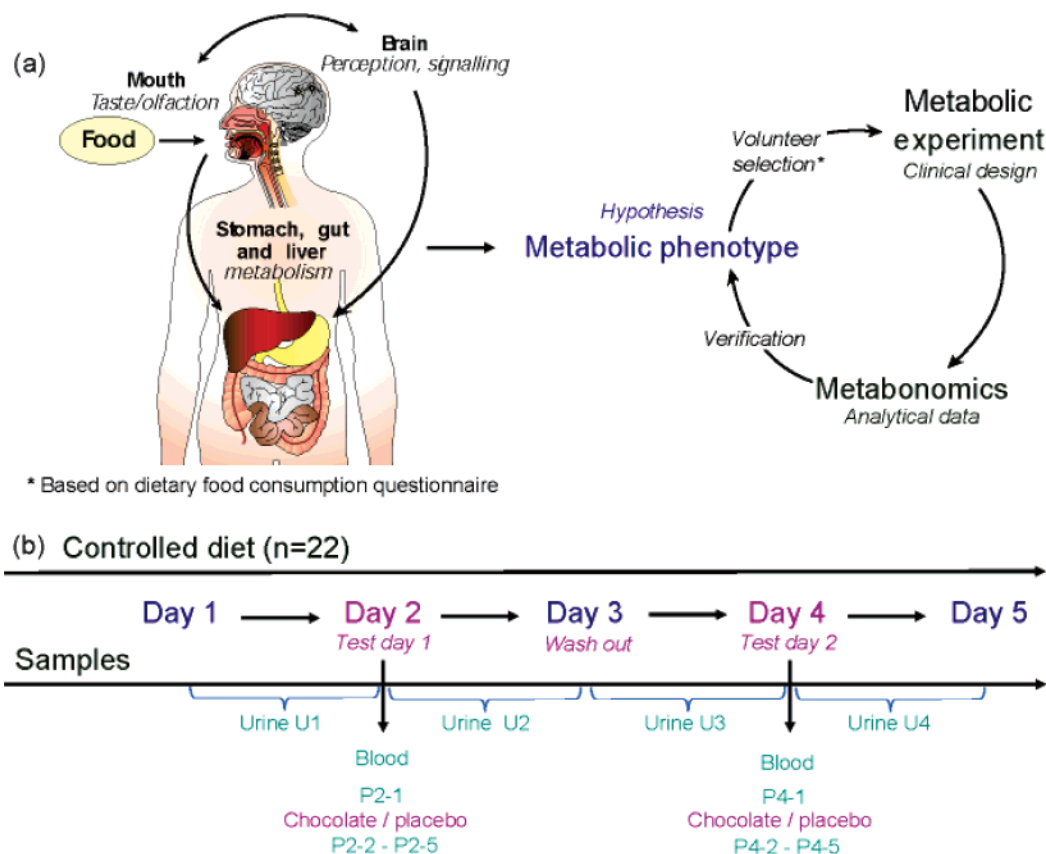
Dietary preferences and habits, which are predominantly cultural in origin, affect the health of both individuals and populations.<sup>1</sup> Food selection by an individual results from a decision process which integrates multiple biological, socio-economic, psychological, and behavioral determinants. Hunger, satiety, palatability, and organoleptic and sensory aspects represent the biological processes that guide a person toward specific food preferences and food choices. In particular, genetically mediated taste response to specific foods or ingredients is ultimately reflected in various dietary preferences.<sup>2,3</sup> Furthermore, genetic heritability has also been reported as a significant contributor in eating habits.<sup>4</sup> So far, no metabolic phenotype (metabotype) has ever been ascribed to an individual human dietary preference group. Variations in the basal metabotype have, however, previously been shown to relate to drug metabolism outcomes, as well as to influence the effects of high fat diets in experimental animals.<sup>5,6</sup> Thus, the ability to classify an individual’s metabotype according to specific dietary preferences, that in turn are part of complex behavioral traits, would be of great potential value in health and well-being assessment, personalized healthcare, and in molecular epidemiological studies. Dietary composition impacts on the metabolism of an individual, and dynamic responses of the metabotype to dietary modulation are well-documented.<sup>1,6,7</sup>

Long-term dietary preferences associated with, for example, high levels of saturated fat or carbohydrate consumption are strongly associated with obesity and heart disease.<sup>8,9</sup> There are also clear differences between human metabolic phenotypes associated with alternating vegetarian and omnivorous diets.<sup>7</sup> Finally, there are strong interactions between dietary composition and gut microbial metabolic activities that might contribute to the development of insulin resistance and nonalcoholic fatty liver disease.<sup>6</sup> Recent studies have also highlighted the importance of altered gut microbiota as possible major contributors to calorific harvest<sup>10</sup> and to obese phenotypes in animal models<sup>11</sup> and man.<sup>12</sup> However, any stable basal metabolic phenotype, that is, one that is not directly stimulated by the dietary preference at the time of sampling, in man relating to specific dietary habits would have to be regarded as imprinted, and this could then be linked to longer-term health risks and outcomes at the individual level. We have recently shown that pharmacometabonomics, that is, individual metabotype variations, in experimental animals can be linked to post-dose drug intervention outcomes such as toxicity and metabolism.<sup>5</sup> Here, we extend this principle to human dietary intervention studies, broadened to include metabotypes that may be indicative of the dietary habits of the individuals; we term this approach “nutrimetabonomics”. In this study, we have used well-validated NMR spectroscopy-based metabotyping methods<sup>13,14</sup> coupled with multivariate statistics to identify specific metabolic subclasses from plasma and urine samples taken from biochemically unscreened adult males with a questionnaire-determined dietary preference and classified as

\* To whom correspondence should be addressed. E-mail: sunil.kochhar@rdls.nestle.com. Phone: + 41 21 785 9336.

<sup>†</sup> Nestlé Research Center.

<sup>‡</sup> Imperial College London.



**Figure 1.** Nutrimetabonomic hypothesis. (a) Outline of the hypothesis linking dietary preference to metabolic phenotype. (b) Schematic representation of the human clinical trial design. The total study period was 5 days, one test day with chocolate consumption and one test day with bread taken as a placebo, with a washout period of 1 day between each test day. On the first day of the study, a portion of a commercially available chocolate (50 g) was given to half the participants of each group to be eaten in the afternoon, and the other half received an equivalent amount of bread. A controlled diet was provided to the subjects a day before and during the study period to minimize variability. After the washout period, the participants who received chocolate first had bread and vice versa. Five samples of blood on test days (P2 and P4) were collected at 0 (P2-1 and P4-1), 5 (P2-2 and P4-2), 15 (P2-3 and P4-3), 30 (P2-4 and P4-4), and 60 (P2-5 and P4-5) min after chocolate/bread intake. Twenty-four hour urine samples (U1–U4) were collected on each day of the clinical trial.

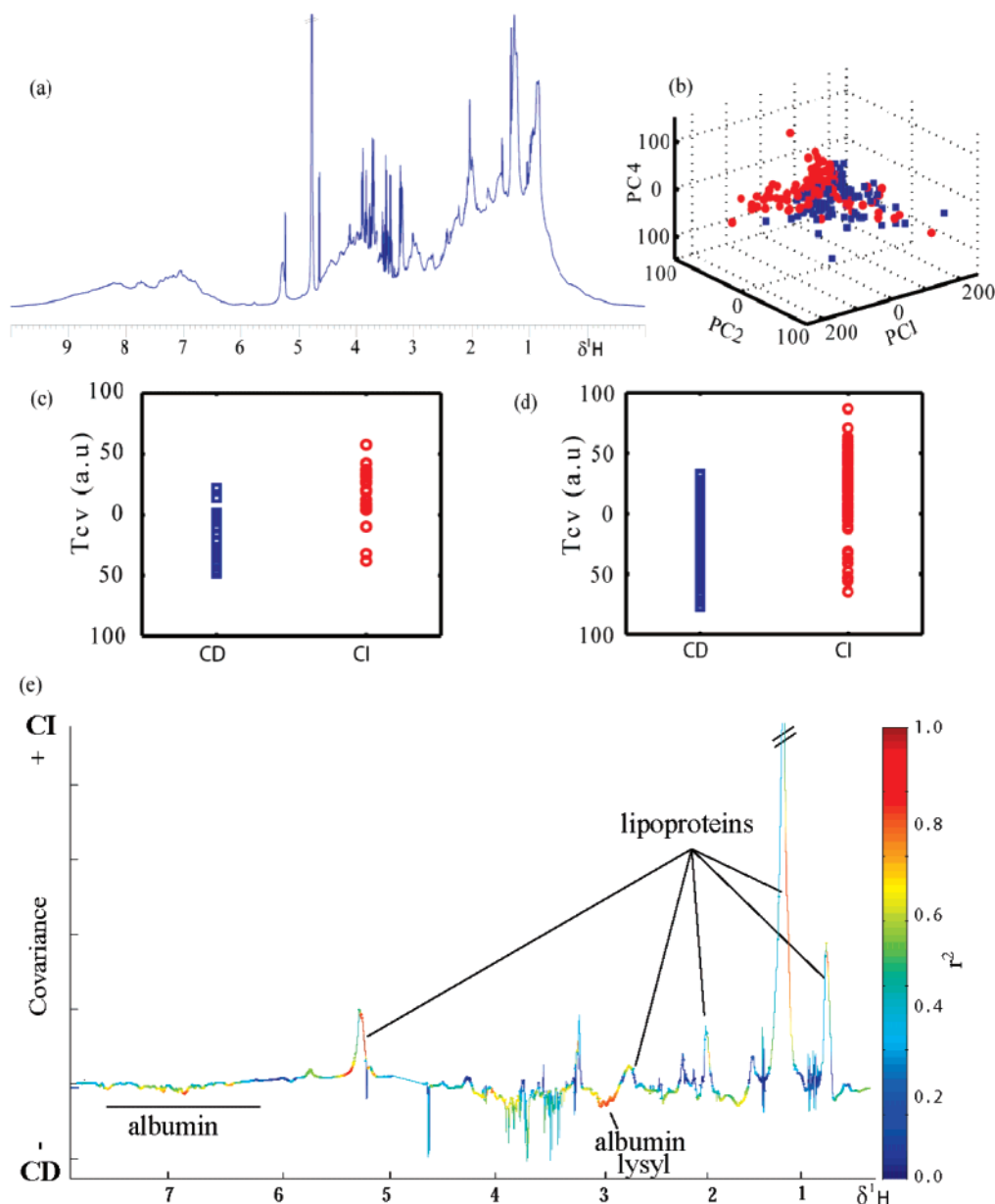
having strong “chocolate desiring” preference and “chocolate indifferent” behavioral phenotype.

## Materials and Methods

**Experimental Design.** This study was conducted in accordance with the ethical principles of Good Clinical Practice and the Declaration of Helsinki, approved by the Nestlé ethical committee, and all subjects gave written informed consent. The hypothesis and the design of the human subject trial are given in Figure 1. A total of 75 healthy men participated in the study, and as a result of the scored chocolate eating habits questionnaire given in the Supporting Information, these were scored between 6 and 30 as chocolate desiring (CD) or chocolate indifferent (CI). Of these, 22 subjects were selected for biofluid collection and data analysis, comprising 11 at each end of the score range. Subjects were also matched according to age and body mass index (BMI), and were selected according to medical evaluation of a confidential health questionnaire. Exclusion criteria included any strong food preference other than chocolate, smoking, alcohol addiction, strong sport activity, medical treatments or investigations, and chronic and temporary illness. Women subjects were not selected in this exploratory study to avoid confounding effects of the menstrual cycle on metabolic profiles. Background information on the subjects participating

in the study is reported in Supplementary Table 1 in Supporting Information. The experiment was designed as a double cross-over clinical trial, during which a standardized diet was provided to the subjects a day before and during the study period to minimize biochemical variability of biological fluids. During the study, chocolate, coffee, and their derived products, tea, soft drinks, energy drinks, spicy foods, alcoholic beverages, and any dietary supplement, were avoided. The total study covered a period of 5 days, one test day with chocolate consumption and one test day with bread taken as a placebo, with a washout period of 1 day between each test day (Figure 1). On the first test day, a portion of a commercially available chocolate (50 g) was given to half the participants of each group to be eaten in the afternoon, and the other half received an isocaloric amount of placebo (bread). Five samples of blood were collected on both test days (P2 and P4) at 0 (P2-1 and P4-1), 5 (P2-2 and P4-2), 15 (P2-3 and P4-3), 30 (P2-4 and P4-4), and 60 (P2-5 and P4-5) min after chocolate/placebo intake. Additionally, samples of the 24 h urine were taken at each day of the clinical trial (U1–U4).

**Sample Preparation and  $^1\text{H}$  NMR Spectroscopic Analysis.** Samples were codified with barcodes and randomly measured using  $^1\text{H}$  NMR spectroscopy under automation using the Bruker laboratory automation and management system (SampleTrack,



**Figure 2.** Multivariate data analysis of <sup>1</sup>H NMR plasma metabolic profiles. (a) A 600 MHz <sup>1</sup>H NMR spectrum of human plasma. (b) 3D PCA scores plot for data from standard <sup>1</sup>H NMR spectra from all plasma samples collected while under dietary control; blue square, “chocolate desiring”; red circle, “chocolate indifferent”; PC1, PC2, and PC4, 55%, 12%, and 6% of the total variance, respectively. (c) O-PLS-DA cross-validated scores (Tcv) plot ( $Q^2 = 0.15$  (7-fold cross-validation) of samples before chocolate/placebo intake (P2-1 and P4-1). CD and CI indicate “chocolate desiring” and “chocolate indifferent” subjects, respectively. (d) O-PLS-DA cross-validated scores plot ( $Q^2 = 0.36$  (7-fold cross-validation) of all samples showing that chocolate consumption has little effect on the model. (e) Coefficients plot derived from all the plasma samples. The O-PLS-DA coefficients plots are presented using a back-scaling transformation, as described previously,<sup>19</sup> which allows each variable to be plotted with a color code which relates to the significance of class discrimination as calculated from the correlation matrix. Positive peaks are from metabolites that are higher in the “chocolate indifferent” class.

Bruker, Germany). Plasma samples (360  $\mu$ L) were introduced into a 4 mm NMR tube with 40  $\mu$ L of <sup>2</sup>H<sub>2</sub>O as locking substance and measured on a Bruker Avance 600 MHz spectrometer equipped with an inverse probe. Three <sup>1</sup>H NMR spectra were acquired for each sample, (a) a standard spectrum using the first increment of the noesy sequence (D1-90°-t1-90°-tm-90°-free induction decay [FID]) with water suppression,<sup>15</sup> (b) Carr-Purcell–Meiboom–Gill (CPMG, D1-90°-( $\tau$ -180°- $\tau$ )-n-FID) spin-echo sequence<sup>16</sup> with water suppression (the analysis reported in this manuscript did not use these data and results will be reported later), and (c) diffusion-edited (D1-90°-G1-180°-G1-

90°-G2- $\Delta$ -90°-G1-180°-G1-90°-G2- $\tau$ -90°-FID) sequence.<sup>17</sup> The standard spectra were acquired with a relaxation delay D1 of 2.5 s during which the water resonance is selectively irradiated, and a fixed interval t1 of 3  $\mu$ s. The water resonance is irradiated for a second time during the mixing time tm of 100 ms. CPMG spin-echo spectra were measured using a spin-echo loop time ( $2n\tau$ ) of 19.2 ms and a relaxation delay of 2.5 s. Diffusion-edited spectra were obtained using a relaxation delay of 1 s, pulsed field gradients G1 and G2 set at 46.8 G·cm<sup>-1</sup>, and a diffusion delay  $\Delta$  of 120 ms during which the molecules are allowed to diffuse. Urines samples (500  $\mu$ L) were adjusted to pH 6.8 using

100  $\mu$ L of phosphate buffer solution ( $\text{KH}_2\text{PO}_4$ , final concentration of 0.2 M) containing 1 mM of sodium 3-(trimethylsilyl)-[2,2,3,3- $^2\text{H}_4$ ]-1-propionate (TSP). Urine spectra were then registered on a Bruker Avance 600 MHz spectrometer equipped with an inverse cryogenic probe. Spectra were acquired using the standard sequence given above, with a relaxation delay of 2.5 s and a mixing time  $t_m$  of 100 ms. For each plasma and urine sample, 32 FIDs were collected into 65 536 data points using a spectral width of 12 019.2 Hz and an acquisition time of 2.7 s.

**NMR Data Processing.**  $^1\text{H}$  NMR spectra were processed using the software package TOPSPIN (version 1.3, Bruker, Germany). The FIDs were multiplied by an exponential weighting function corresponding to a line broadening of 0.3 Hz for standard plasma spectra. A line broadening of 1 Hz was applied to plasma CPMG and diffusion-edited data as well as to urine spectra. The acquired NMR spectra were manually phase- and baseline-corrected and referenced to the chemical shift of the methyl resonance of lactate at  $\delta$  1.33 for plasma spectra and of TSP at  $\delta$  0.00 for urine samples.

**Clinical Quantitative Measurements.** Free fatty acids (data not shown) were measured using a Wako enzymatic method on a XPAND system (Dade Behring, Switzerland). High-density lipoprotein (HDL) and low-density lipoprotein (LDL) cholesterol were determined using the AHD and ALDL Cholesterol assay systems (Dade Behring, Switzerland).

**Chemometrics.** Full resolution spectra incorporating data points within the  $\delta$  0.4–9.5 region were used for statistical multivariate analysis excluding the water residue signal between  $\delta$  4.7–5 (plasma data set) and  $\delta$  4.5–6.5 (urine data set). Statistical analysis was performed using the software package SIMCA-P+ (version 11.0.0, Umetrics AB, Umeå, Sweden) and in-house developed MATLAB (The MathWorks, Inc., Natick, MA) routines. Initial data analyses were conducted using principal components analysis (PCA), and detailed classification studies used projection to latent structure-discriminant analysis with filtering of data variables orthogonal to sample class (O-PLS-DA)<sup>18</sup> with unit-variance scaling (each parameter has a mean of zero and a variance of one). This approach provides a way to filter metabolic information which is not correlated to the predefined classes, while the loadings yield information on which NMR signals are associated with the observed clustering, thus, giving a means for metabolic interpretation. In the O-PLS-DA method, the variation in the **X** matrix (the NMR spectra) and the **Y** matrix (the descriptive variables or the classes) is separated into three parts. The first part contains the variation common to **X** and **Y**, the second part contains the specific variation for **X**, so-called structured noise, and the last part contains the residual variance. The O-PLS-DA method allows improved interpretation of the models because the structured noise is modeled separately from the variation common to the **X** and **Y** matrices. To test the validity of the model against over-fitting, the cross-validation parameter  $Q^2$  was computed. The O-PLS-DA loadings plot was processed according the method described by Cloarec et al.<sup>19</sup> This consists of combining the back-scaled O-PLS-DA loadings from a model where the data had been autoscaled to unit variance with the variable weights of the same model in the same plot. For this purpose, each O-PLS-DA loading is first multiplied by the standard deviation (back scaling) of its corresponding variable and then plotted as a function of its related chemical shift but with a color code linked to the weights of the selected latent variable. In this way, the common resonances from metabolites

**Table 1.** O-PLS-DA Model Summary for Discriminating NMR Spectra of Plasma and Urine<sup>a</sup>

biofluids	NMR spectra	time points	$R^2\text{X}$	$R^2\text{Y}$	$Q^2$	$P$ -value, Tcv
Plasma	Standard	All time points	0.73	0.40	0.36	$3.2 \times 10^{-25}$
		P2-1 and P4-1	0.76	0.41	0.15	$5.2 \times 10^{-8}$
	CPMG spin-echo	All time points	0.40	0.74	0.33	$7.5 \times 10^{-64}$
		P2-1 and P4-1	0.33	0.57	-0.09	$7.5 \times 10^{-10}$
Urine	Diffusion-edited	All time points	0.61	0.58	0.28	$2.4 \times 10^{-42}$
		P2-1 and P4-1	0.63	0.84	0.30	$7.4 \times 10^{-9}$
	Standard	All time points	0.20	0.78	0.39	$6.8 \times 10^{-30}$
		U1	0.23	0.85	0.06	$7.6 \times 10^{-10}$

<sup>a</sup> O-PLS-DA models were generated with 1 predictive component and 2 orthogonal components to discriminate between 2 groups for plasma and urine samples. Models were calculated using all time points for test days 1 and 2. Also, models were calculated for the specific time points P2-1 and P4-1 before chocolate or placebo consumption. The  $R^2\text{X}$  and  $R^2\text{Y}$  values show how much of the variation in the data sets **X** and **Y**, respectively, is explained by the model. The  $Q^2$  value represents the predictability of the models, and relates to its statistical validity. A negative value indicates that differences between groups are statistically non-significant. Since the cross-validated scores (Tcv) are assumed to follow a normal distribution, a paired Student's  $t$  test was applied between the two groups, and  $P$ -values were calculated at a confidence level of 95%.

involved in the discrimination, that is, “chocolate desiring” and “chocolate indifferent” subjects, are highlighted. The interpretation of the loadings is therefore more straightforward because the resulting plot provides roughly the same shape as that of a real spectrum. The standard 7-fold cross-validation method (repeatedly leaving out a seventh of the samples and predicting them back into the model) was applied to establish the robustness of the model. The assignment of metabolites was achieved using a range of two-dimensional NMR spectroscopic techniques on selected samples.

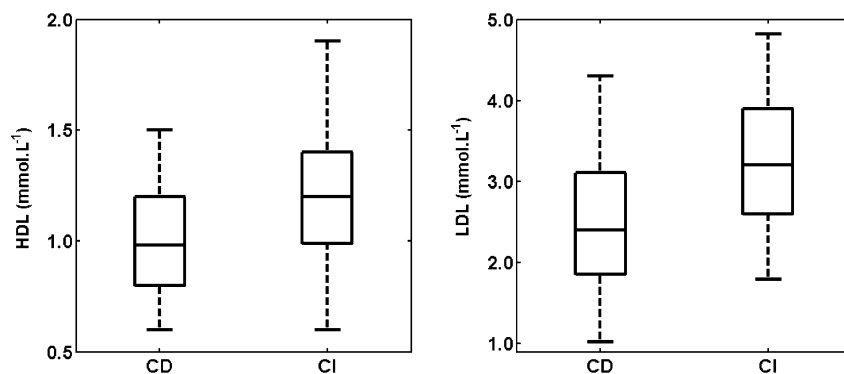
## Results

**Analysis of  $^1\text{H}$  NMR Spectroscopic Data on Plasma.** A typical standard  $^1\text{H}$  NMR spectrum of human blood plasma is shown in Figure 2. Such spectra exhibit a broad set of resonances arising from lipoprotein-bound fatty acyl groups found in triglycerides, phospholipids, and cholesteryl esters, together with peaks from the glyceryl moiety of triglycerides and the choline head group of phosphatidylcholine. Also, many sharper peaks arising from the major low molecular weight molecules present in plasma are observed.<sup>20</sup> PCA was performed on the whole set of standard NMR spectra of plasma taken when the subjects were under identical dietary regime. The scores plot of the first, second, and fourth principal components (Figure 2) explained 73% of the total spectral metabolic variance and showed partial clustering of the samples, when coded according to “chocolate desiring” or “chocolate indifferent”, indicating an underlying class-determining metabolic fingerprint. This clustering indicates that there are underlying metabolic differences between classes that can be revealed by a simple unsupervised model. To determine more precisely the metabolic biomarkers that discriminated the classes, we used a cross-validated orthogonal-filtered partial least-squares discriminant analysis<sup>18</sup> (O-PLS-DA) approach which showed a global structuring of the sample points grouping “chocolate desiring” and “chocolate indifferent” participants in distinct parts of the scores plot analogous to the PC analysis (Figure 2). NMR spectra edited on the basis of molecular diffusion coefficients, which have small molecule peaks eliminated (data not shown), gave identical results in terms of class separation (Table 1), showing that the discriminating substances are plasma macromolecular species. NMR

**Table 2.** Assignment of Statistically Significant Increases in NMR Signal Intensities in Plasma of “Chocolate Desiring” and “Chocolate Indifferent” Subjects<sup>a</sup>

increased in	signal assignments	moieties	$\delta$ <sup>1</sup> H (ppm) and multiplicity	<i>P</i> -values
“Chocolate indifferent”	lipids (mainly LDL)	CH <sub>3</sub>	0.86(m)	$5.2 \times 10^{-5}$
	lipids (mainly LDL)	(CH <sub>2</sub> ) <sub><i>n</i></sub>	1.25(m)	$1.8 \times 10^{-6}$
	lipids (mainly LDL)	CH <sub>2</sub> -C=C	1.99(m)	$1.2 \times 10^{-5}$
	lipids	=C-CH <sub>2</sub> -C=	2.75(m)	$3.7 \times 10^{-5}$
	choline in phospholipids	N-(CH <sub>3</sub> ) <sub>3</sub> , OCH <sub>2</sub> , NCH <sub>2</sub>	3.21(s), 4.05(t), 3.51(t)	$2.6 \times 10^{-5}$
Chocolate desiring”	unsaturated lipids	-CH=CH-	5.28(m)	$2.3 \times 10^{-10}$
	albumin	$\epsilon$ -CH <sub>2</sub>	2.90(t), 2.96(t), 3.02(t)	$4.2 \times 10^{-3}$

<sup>a</sup>  $\delta$  <sup>1</sup>H, chemical shifts calibrated against the lactate signal at  $\delta$  1.33; s, singlet; t, triplet; m, multiplet. Since metabolite concentrations do not follow a normal distribution, the nonparametric Wilcoxon matched-pairs signed rank test was applied to preselected metabolite peak intensities scaled to the noise level between the two groups, and *P*-values were calculated at a confidence level of 95%.



**Figure 3.** Comparison of blood plasma concentrations of high-density lipoprotein (HDL) and low-density lipoprotein (LDL) between the “chocolate desiring” (CD) and “chocolate indifferent” (CI) subjects displayed using box and whisker plots. Statistical analysis was performed using Wilcoxon matched-pairs signed rank test at a confidence level of 95% on the clinical quantitative measurements of HDL and LDL done on a representative subset of plasma samples. The blood plasma levels of HDL ( $n = 135$ ,  $P = 1.2 \times 10^{-4}$ ) and LDL ( $n = 135$ ,  $P = 9.9 \times 10^{-7}$ ) were significantly different between the two groups.

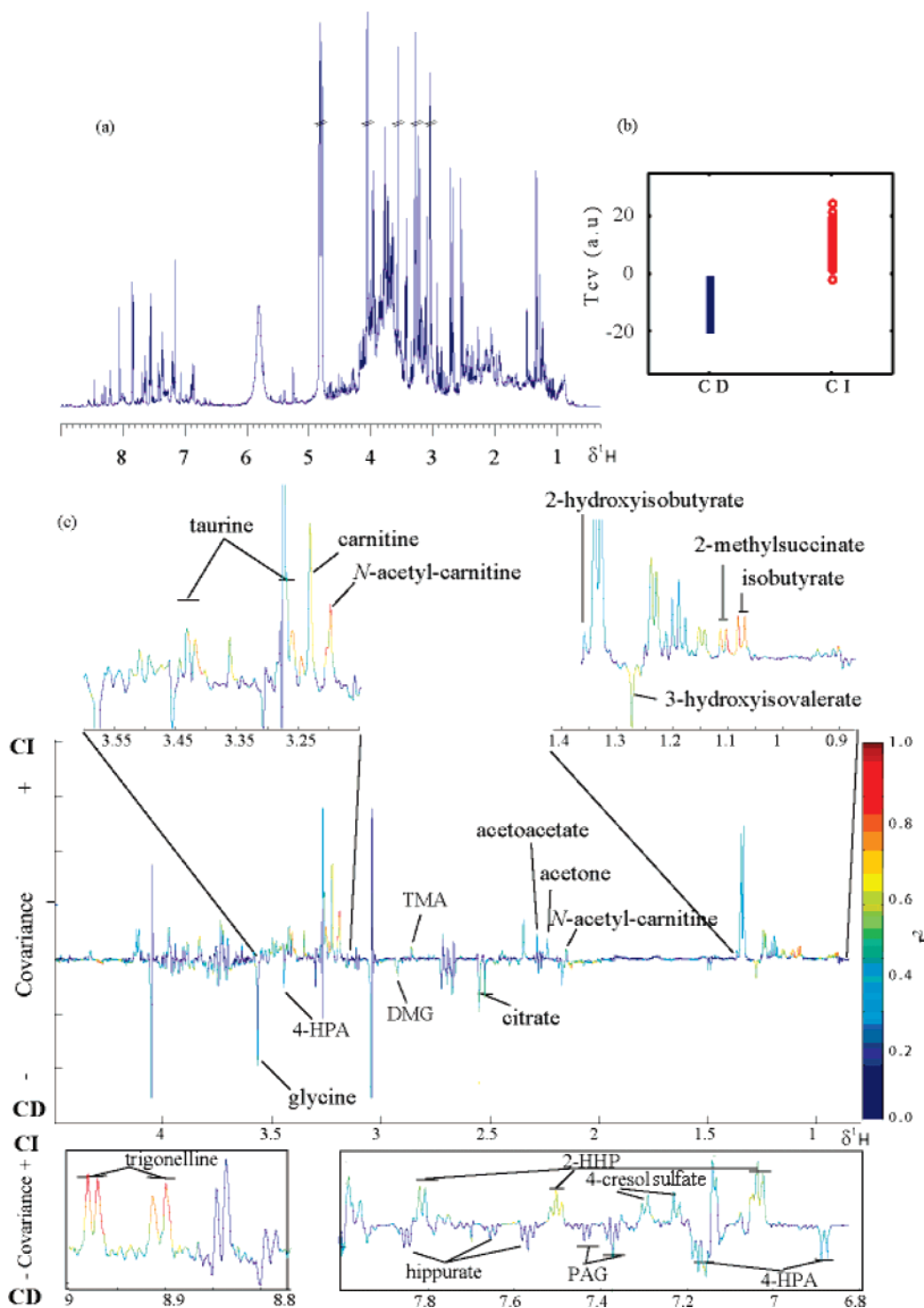
data generated using a spin-echo pulse sequence,<sup>20</sup> which attenuates the signals of macromolecules, support this observation as noted with the nonvalidity of the models generated on samples collected before chocolate/placebo intake (Table 1).

Interpretation of the back-scaled O-PLS-DA loadings<sup>19</sup> for the first latent component indicated that “chocolate indifferent” class subjects are distinguished by significantly higher levels of plasma lipids (Figure 2, Table 2) incorporated into lipoproteins. Plasma lipoproteins, namely, chylomicrons, chylomicron remnants, very low-density lipoprotein, intermediate-density lipoprotein, LDL, and HDL, differ in their composition (free and esterified cholesterol, triacylglycerols, phospholipids, and apolipoproteins), their size, and in their biological function.<sup>21</sup> The <sup>1</sup>H NMR spectrum of plasma contains information correlated to both chemical composition of lipoproteins and their particle size, which allows changes in distribution of lipoprotein classes and subclasses to be monitored.<sup>22,23</sup> The spectral regions from  $\delta$  0.8 to 1.4 comprises the complex methyl and methylene signals from fatty acyl groups, and by correlation analysis, it is the central parts of these peaks, mainly dominated by LDL particles, that exhibit the strongest correlation with the “chocolate indifferent” class.<sup>23</sup> To validate these observations from the data-driven metabolomic experiments, we performed independent quantitative lipoprotein analyses for HDL and LDL confirming the NMR observations (Figure 3). There was no correlation between BMI and dietary preference class, and indeed, nearly all subjects were essentially within the normal range (Supplementary Table 1 in Supporting Information). In addition, increased levels of unsaturated and polyunsaturated

fatty acyl species in lipoproteins and a decreased level of albumin were also correlated to the “chocolate indifferent” class (Table 2).

**Analysis of <sup>1</sup>H NMR Spectroscopic Data on Urine.** Urinary metabolic profiles give time-averaged representations of the recent homeostatic metabolic changes of the individual and carry indirect information on the gut microbial metabolic activities, as observed through the excretion of many polar microbial-mammalian cometabolites.<sup>6,24</sup> A typical <sup>1</sup>H NMR spectrum of human urine is shown in Figure 4. Application of O-PLS-DA resulted in clear clustering of samples (Figure 4, Table 3) for all time points according to chocolate preferences. Following the same procedure as previously described, the O-PLS-DA loadings plot revealed resonances, which exhibit a strong correlation with the classes. Some signals strongly correlated with the “chocolate indifferent” class appear within the spectral region from  $\delta$  1.0 to 1.2 (Figure 4). From the correlation analysis, increased levels of 3-hydroxyisovalerate, dimethylglycine (DMG), glycine, and citrate were observed for the “chocolate desiring” class. On the other hand, isobutyrate, methylsuccinate, acetone, acetoacetate, trimethylamine (TMA), taurine, trigonelline, carnitine, and *N*-acetyl-carnitine were seen to increase in the “chocolate indifferent” class. Within the aromatic region from  $\delta$  6.8 to 8.0 (Figure 4), the two phenylacetyl derivatives, phenylacetylglutamine (PAG) and 4-hydroxyphenylacetate (4-HPA), as well as 2-hydroxyhippurate (2-HHP), are strongly correlated with the “chocolate desiring” and “chocolate indifferent” classes, respectively.

Moreover, the analysis of correlations between aromatic metabolites ascribed to gut microflora metabolism in each class



**Figure 4.** Multivariate data analysis of <sup>1</sup>H NMR urine metabolic profiles. (a) A 600 MHz <sup>1</sup>H NMR spectrum of human urine. (b) O-PLS-DA cross-validated scores plot ( $O^2 = 0.39$  (7-fold cross-validation) of all samples). (c) O-PLS-DA loadings plot from the model back-scaled and plotted as a function of chemical shifts. The back-scaled coefficients plot is color coded to the weights of the selected latent variable. Positive peaks are from metabolites that are higher in the “chocolate indifferent” class. Individuals belonging to the “chocolate desiring” class showed higher levels of 3-hydroxyisovalerate, citrate, dimethylglycine (DMG), glycine, phenylacetylglutamine (PAG), and 4-hydroxyphenylacetate (4-HPA). “Chocolate indifferent” subjects had higher urinary excretion of isobutyrate, 2-methylsuccinate, trimethylamine, *N*-acetyl-carnitine, carnitine, taurine, 2-hydroxyhippurate, and trigonelline.

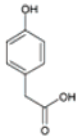
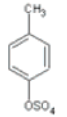
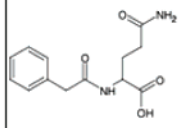
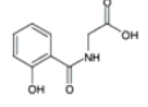
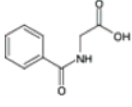
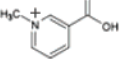
was performed to provide deeper insights into the class-specific microbial activities (Figure 5). The correlation matrix reports the Pearson correlation coefficients between <sup>1</sup>H NMR intensities scaled to the noise level obtained from the different aromatic metabolites and their respective *P*-values calculated at the

confidence levels of 95% and 90% for the “chocolate desiring” (CD, blue) and the “chocolate indifferent” (CI, red) subjects. Positive and negative correlations show the multicollinearity between metabolites whose concentrations are interdependent such as in the case of a substrate–product biochemical reaction

**Table 3.** Assignment of Statistically Significant Increases in NMR Signal Intensities in Urine of “Chocolate Desiring” and “Chocolate Indifferent” Subjects<sup>a</sup>

high in	signal assignments	moieties	$\delta$ <sup>1</sup> H (ppm) and multiplicity	<i>P</i> -values
“Chocolate indifferent”	isobutyrate	CH <sub>3</sub> , CH	1.07(d), 2.39(m)	$2.5 \times 10^{-2}$
	methylsuccinate	CH <sub>3</sub> , CH, CH <sub>2</sub>	1.11(d), 2.13(dd), 2.52(dd), 2.62(m)	$1.2 \times 10^{-5}$
	acetone	CH <sub>3</sub>	2.24(s)	0.02
	acetoacetate	CH <sub>3</sub>	2.29(s)	0.87
	trimethylamine	(CH <sub>3</sub> ) <sub>3</sub> -N	2.87(s)	$1.0 \times 10^{-4}$
	<i>N</i> -acetyl-carnitine	N(CH <sub>3</sub> ) <sub>2</sub> , COCH <sub>3</sub> , CH, CH <sub>2</sub> COOH, CH <sub>2</sub> COOH, CH <sub>2</sub> N	3.19(s), 2.14(s), 3.85(dd), 2.65(dd), 2.51(dd), 3.61(d)	$2.1 \times 10^{-3}$
	carnitine	CH <sub>3</sub> , CH <sub>2</sub> COOH, NCH <sub>2</sub>	3.23(s), 2.44(dd), 3.43(m)	$5.1 \times 10^{-3}$
	taurine	S-CH <sub>2</sub> , N-CH <sub>2</sub>	3.26(t), 3.40(t)	$5.6 \times 10^{-3}$
	2-hydroxyhippurate	6-CH, 4-CH, 3,6-CH, CH <sub>2</sub>	7.80(dd), 7.49(m), 7.02(m), 3.99(s)	$1.1 \times 10^{-2}$
	trigonelline	2-CH, 4-CH, 5-CH, 6-CH, CH <sub>3</sub>	9.28(s), 8.90(d), 8.21(dd), 8.99(d), 4.48(s)	$3.5 \times 10^{-2}$
	2-hydroxyisobutyrate	(CH <sub>3</sub> ) <sub>2</sub> -C	1.36(s)	$1.1 \times 10^{-2}$
	4-cresol	CH <sub>3</sub> , 2,6-CH, 3,5-CH	2.34(s), 7.21(d), 7.28(d)	$2.6 \times 10^{-2}$
	3-hydroxyisovalerate	(CH <sub>3</sub> ) <sub>2</sub> -C	1.27(s)	$6.1 \times 10^{-2}$
	phenylacetylglutamine	3,5-CH, 2,4,6-CH, 1-CH, 3-CH <sub>2</sub> , 2-CH <sub>2</sub>	7.43(m), 7.36(m), 4.18(m), 2.27(t), 2.11(m)	0.60
“Chocolate desiring”	citrate	half CH <sub>2</sub> , half CH <sub>2</sub>	2.54(d), 2.65(d)	0.60
dimethylglycine	CH <sub>3</sub>	2.93(s)	$7.8 \times 10^{-2}$	
4-hydroxyphenylacetate	CH <sub>2</sub> , 2,6-CH, 3,5-CH	3.45(s), 7.17(d), 6.87(d)	0.90	
glycine	CH <sub>2</sub>	3.57(s)	0.23	

<sup>a</sup>  $\delta$  <sup>1</sup>H, chemical shifts calibrated against the TSP signal at  $\delta$  0.0; s, singlet; d, doublet; t, triplet; m, multiplet; dd, doublet of doublets. Wilcoxon matched-pairs signed rank test was applied to preselected metabolite peak intensities scaled to the noise level between the two groups, and *P*-values were calculated at a confidence level of 95%.

						
	4-HPA	4-cresol sulfate	PAG	2-HHP	hippurate	trigonelline
4-HPA		<b>-0.3 (0.0860)*</b>	<b>0.6 (0.01x10<sup>-3</sup>)**</b>	<b>0.1 (0.5620)</b>	<b>-0.3 (0.0620)*</b>	<b>0.3 (0.0560)*</b>
4-cresol sulfate	<b>0.1 (0.3990)</b>		<b>0.3 (0.0220)**</b>	<b>0.1 (0.5450)</b>	<b>0.0 (0.8150)</b>	<b>0.1 (0.4570)</b>
PAG	<b>0.3 (0.0540)**</b>	<b>0.5 (0.0010)**</b>		<b>0.1 (0.3490)</b>	<b>-0.2 (0.1120)</b>	<b>0.3 (0.0340)**</b>
2-HHP	<b>-0.1 (0.6850)</b>	<b>0.2 (0.3240)</b>	<b>0.1 (0.5310)</b>		<b>-0.1 (0.4740)</b>	<b>-0.1 (0.3880)</b>
hippurate	<b>-0.1 (0.5800)</b>	<b>0.3 (0.0850)*</b>	<b>0.2 (0.2160)</b>	<b>0.3 (0.0300)**</b>		<b>0.0 (0.9950)</b>
trigonelline	<b>0.1 (0.4940)</b>	<b>0.1 (0.5560)</b>	<b>0.2 (0.1850)</b>	<b>0.3 (0.0470)**</b>	<b>-0.1 (0.7360)</b>	

**Figure 5.** Correlation matrix between urinary host and microbial metabolites in “chocolate indifferent” and “chocolate desiring” individuals. This figure reports the Pearson correlation coefficients between <sup>1</sup>H NMR intensities scaled to the noise level obtained from the different aromatic metabolites and their respective *P*-values calculated at the confidence levels of 95% (\*\*\*) and 90% (\*) for the “chocolate desiring” (CD, blue) and the “chocolate indifferent” (CI, red) subjects. Positive and negative correlations show the multicollinearity between metabolites whose concentrations are interdependent such as in the case of a substrate–product biochemical reaction or under some common regulatory mechanisms across different pathways. Phenylacetylglutamine (PAG) shows significant and non-class specific correlations with 4-hydroxyphenylacetate (4-HPA) and 4-cresol sulfate, two microbial metabolites, which may indicate a relationship between microflora activity and the host liver and kidney metabolism. However, the correlations of 2-hydroxyhippurate (2-HHP) with hippurate and trigonelline are specifically observed in the CI group, while the CD group is characterized by the correlation of trigonelline with PAG. These observations highlight a class-specific microbial modulation of dietary flavonoids and niacin metabolism. Moreover, the conversion of 4-HPA to 4-cresol sulfate exhibits a negative correlation in the urine profiles of CD subjects confirming an already established substrate–product biochemical relationship associated with the gut activity of *C. difficile*. Interestingly, this relationship shows a non-significant correlation in the CI individuals. These findings suggest differential mammalian–microbial metabolism for the considered metabolites between the “chocolate desiring” and the “chocolate indifferent” individuals.

or under some common regulatory mechanisms across different pathways. Statistically significant correlations were observed for the following pairs of metabolites: 4-HPA/PAG, 4-cresol sulfate/PAG, 4-Cresol sulfate/Hippurate, 2-HHP/hippurate, and 2-HHP/trigonelline in the “chocolate indifferent” group, and 4-cresol sulfate/4-HPA, PAG/4-HPA, PAG/4-Cresol, hippurate/4-HPA, trigonelline/4-HPA, and trigonelline/PAG in the “chocolate desiring” group (Figure 5).

## Discussion

Consumption of chocolate in the test volunteers had no effect on the metabolic phenotype separation since class clustering was already obtained on samples collected before chocolate/placebo intake. This observation precludes the segregation of classes due to residual metabolites or direct metabolic effects of chocolate consumption and supports the notion of stable metabolic imprinting linked to dietary preferences. The metabolic profiles collected on both test days, thus, indicated that “chocolate preference” is an objectively definable and measurable individual preference class, although we note that dietary chocolate itself is likely to be only one indicator of a more complex dietary preference background. The main metabolic determinants in these metabolotypes are the decreased concentration of LDL and the elevated level of albumin in the “chocolate desiring” group. Albumin, the most abundant plasma protein, is involved in many physiological processes including transport of free fatty acids generated from the lipoprotein lipase action in the lipoprotein metabolism. A relationship between albumin concentration and apoprotein B-100 secretion and LDL catabolism has also been reported on a human hepatocyte model.<sup>25</sup> Thus, intriguingly, the chocolate-preferring class has intrinsically different lipoprotein and lipid status even in the absence of chocolate stimulation.

Furthermore, the class-specific metabolic phenotype is also encoded in the urinary profiles, suggesting both distinct energy and microbiota metabolism. The increased urinary excretion of PAG and citrate in the “chocolate desiring” subjects suggests a different modulation of the tricarboxylic acid cycle,<sup>26</sup> as well as variations in renal tubular pH and its subsequent changes in aconitase activity.<sup>27</sup> Moreover, the relative higher urinary levels of carnitine and *N*-acetyl-carnitine observed in the “chocolate indifferent” group could be related to a different basal energy metabolism, for example, lipid oxidation, in relation with the trend observed with the excretion of the ketone bodies acetone and acetoacetate.<sup>28,29</sup> Aromatic compounds, such as phenylacetate and hippurate derivatives, methylamines, many short chain fatty acids, and their hydroxylation products are strongly associated with gut microbial metabolism<sup>24,30–33</sup> some of them being previously associated with chocolate consumption.<sup>34</sup> The changes in urinary excretion of trigonelline indicate a class-specific metabolism of niacin, an essential vitamin involved in major physiological functions as a coenzyme in tissue respiration, carbohydrate, and lipid metabolism. Niacin requirements are satisfied by both dietary sources and biosynthesis through a tryptophan-mediated metabolism ensured by the liver and the gut microflora.<sup>35</sup> In addition to this, class-specific changes between the levels of the aromatic metabolites 4-HPA and 2-HHP on the one hand, and between TMA, DMG, and glycine on the other hand, provide further evidence of a gut microbial oriented metabolism depending on chocolate preferences.

In addition, the correlation matrix between aromatic metabolites illustrates class-specific mammalian–microbial co-

metabolism. For instance, the significant correlations of PAG with the microbial metabolites 4-HPA and 4-cresol sulfate in the whole population may highlight a relationship between microflora activity and the host metabolism of aromatic compounds. Furthermore, the conversion of 4-HPA to 4-cresol sulfate exhibits a negative correlation in the urine profiles of “chocolate desiring” subjects, which corresponds to a molecular process that has been associated with the presence of *Clostridium difficile*,<sup>30</sup> but possibly other Clostridial strains as well. This observation suggests a differential management of the metabolic pool of the precursor 4-HPA by gut bacteria. Interestingly, the correlation of 2-HHP with hippurate, two metabolites deriving from the metabolism of dietary flavonoids and hydroxycinnamates by colon microorganisms,<sup>36</sup> and the particular correlation of trigonelline with 2-HHP, might indicate a specific microbial modulation of dietary flavonoids and niacin metabolism in “chocolate indifferent” subjects. Altogether, these findings suggest differential mammalian–microbial metabolism for the considered metabolites between the “chocolate desiring” and the “chocolate indifferent” individuals.

Our observations demonstrate imprinted differences in the gut microbial metabolic activities of the individuals that appear to depend on their previous dietary consumption habits. These results are supported by previous reports on the impact of dietary habits on the re-colonization of the gut microflora<sup>11,34</sup> and imply that dietary influences on microbial activities may be much more subtle than previously suspected. Indeed, it is widely thought that the gut microbial populations in individuals are stable,<sup>37,38</sup> but these data suggest that metabolic activity of the gut microbes and, consequently, the metabolites that are fed to the host may be more finely modulated by diet than previously thought. We note that a specific dietary preference appears to influence the functional ecology and biochemistry of the gut in healthy individuals in the sense that excreted metabolites closely reflect the total metabolic activities of the microbiome.

In conclusion, these data suggest the occurrence of a metabolic imprinting of the basal metabolic phenotype in relation to a behavioral/psychological dietary preference that is characterized by “chocolate desiring” or “chocolate indifference”. This imprinting is independent of the ingested food, as chocolate consumption versus placebo has no direct effect. The plasma metabolotype variation with preference class is mainly characterized by differences in the lipoprotein profiles. In addition, the metabolic differences observed in urine suggest considerable differences in gut microbial metabolic activities that may be of long-term health significance to the host. The observed metabolic imprinting provides evidence for a link between specific dietary preference patterns and metabolic phenotype. Obtaining broader and more profound insights into the cause and effects of the metabolic differences underlying dietary preferences would have considerable implications. The nutrimetabonomics approach presented here, which allows human nutritional preference to be linked to metabolic phenotypes of individuals, can then be extended to a range of functional food products and eating habits. In particular, the comparison of metabolic biomarkers associated with different eating habits would assess if similar molecular mechanisms are involved across different “desiring” phenotypes. Such applications of nutrimetabonomics will provide a future basis for classification of dietary responses and the possibility of optimized or personalized nutritional management.

**Abbreviations:** CPMG, Carr–Purcell–Meiboom–Gill; NMR, nuclear magnetic resonance; O-PLS-DA, orthogonal projection to latent structure discriminant analysis.

**Acknowledgment.** We thank the Metabolic unit of the Nestlé Research Center for the study design and all sample provision. The research was funded by Nestlé.

**Supporting Information Available:** A Supplementary Table describing the background information on the subjects participating in the study and the chocolate eating habits questionnaire. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- Rezzi, S.; Ramadan, Z.; Fay, L. B.; Kochhar, S. *J. Proteome Res.* **2007**, *6*, 513–525.
- Drewnowski, A. *Annu. Rev. Nutr.* **1997**, *17*, 237–253.
- Drewnowski, A.; Henderson, S. A.; Levine, A.; Hann, C. *Public Health Nutr.* **1999**, *2*, 513–519.
- Breen, F. M.; Plomin, R.; Wardle, J. *Physiol. Behav.* **2006**, *88*, 443–447.
- 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. *Nature* **2006**, *440*, 1073–1077.
- 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. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12511–12516.
- Stella, C.; Beckwith-Hall, B.; Cloarec, O.; Holmes, E.; Lindon, J. C.; Powell, J.; vanderOuderaa, F.; Bingham, S.; Cross, A. J.; Nicholson, J. K. *J. Proteome Res.* **2006**, *5*, 2780–2788.
- Lichtenstein, A. H.; Appel, L. J.; Brands, M.; Carnethon, M.; Daniels, S.; Franch, H. A.; Franklin, B.; Kris-Etherton, P.; Harris, W. S.; Howard, B.; Karanja, N.; Lefevre, M.; Rudel, L.; Sacks, F.; Van Horn, L.; Winston, M.; Wylie-Rosett, J. *Circulation* **2006**, *114*, 82–96.
- Kuller, L. H. *Nutr. Rev.* **2006**, *64*, S15–S26.
- Martin, F. P.; Dumas, M. E.; Wang, Y.; Legido-Quigley, C.; Yap, I. K.; Tang, H.; 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. *Mol. Syst. Biol.* **2007**, *3*, 112.
- Turnbaugh, P.; Ley, R.; Mahowald, M.; Magrini, V.; Mardis, E.; Gordon, J. *Nature* **2006**, *444*, 1027–1031.
- Ley, R.; Turnbaugh, P.; Klein, S.; Gordon, J. *Nature* **2006**, *444*, 1023–1024.
- Gavaghan, C. L.; Wilson, I. D.; Nicholson, J. K. *J. Proteome Res.* **2006**, *5*, 378–384.
- Nicholson, J. K.; Connelly, J.; Lindon, J. C.; Holmes, E. *Nat. Rev. Drug Discovery* **2002**, *1*, 153–161.
- Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546–4553.
- Meiboom, S.; Gill, D. *Rev. Sci. Instrum.* **1958**, *29*, 688–691.
- Wu, D. H.; Chen, A. D.; Johnson, C. S. *J. Magn. Reson., Ser. A* **1995**, *115*, 260–264.
- Trygg, J.; Wold, S. *J. Chemom.* **2002**, *16*, 119–128.
- Cloarec, O.; Dumas, M. E.; Trygg, J.; Craig, A.; Barton, R. H.; Lindon, J. C.; Nicholson, J. K.; Holmes, E. *Anal. Chem.* **2005**, *77*, 517–526.
- Nicholson, J. K.; Foxall, P. J.; Spraul, M.; Farrant, R. D.; Lindon, J. C. *Anal. Chem.* **1995**, *67*, 793–811.
- Jackson, R. L.; Morrisett, J. D.; Gotto, A. M., Jr. *Physiol. Rev.* **1976**, *56*, 259–316.
- Hiltunen, Y.; Ala-Korpela, M.; Jokisaari, J.; Eskelinen, S.; Kiviniitty, K.; Savolainen, M.; Kesaniemi, Y. A. *Magn. Reson. Med.* **1991**, *21*, 222–232.
- Otvos, J. D.; Jeyarajah, E. J.; Bennett, D. W. *Clin. Chem.* **1991**, *37*, 377–386.
- Nicholson, J. K.; Holmes, E.; Wilson, I. D. *Nat. Rev. Microbiol.* **2005**, *3*, 431–438.
- Cianflone, K.; Vu, H.; Zhang, Z.; Sniderman, A. D. *Atherosclerosis* **1994**, *107*, 125–135.
- Yang, D.; Brunengraber, H. *J. Nutr.* **2000**, *130*, 991S–994S.
- Nicholson, J. K.; Higham, D. P.; Timbrell, J. A.; Sadler, P. J. *Mol. Pharmacol.* **1989**, *36*, 398–404.
- Bales, J. R.; Bell, J. D.; Nicholson, J. K.; Sadler, P. J. *Magn. Reson. Med.* **1986**, *3*, 849–856.
- Brass, E. P. *Am. J. Clin. Nutr.* **2000**, *72*, 618S–623S.
- Selmer, T.; Andrei, P. I. *Eur. J. Biochem.* **2001**, *268*, 1363–1372.
- Gostner, A.; Blaut, M.; Schaffer, V.; Kozianowski, G.; Theis, S.; Klingeberg, M.; Dombrowski, Y.; Martin, D.; Ehrhardt, S.; Taras, D.; Schwiertz, A.; Kleessen, B.; Luhrs, H.; Schaubert, J.; Dorbath, D.; Menzel, T.; Scheppach, W. *Br. J. Nutr.* **2006**, *95*, 40–50.
- Wang, Y.; Holmes, E.; Nicholson, J. K.; Cloarec, O.; Chollet, J.; Tanner, M.; Singer, B. H.; Utzinger, J. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12676–12681.
- Zeisel, S. H.; Blusztajn, J. K. *Annu. Rev. Nutr.* **1994**, *14*, 269–296.
- Rios, L. Y.; Gonthier, M. P.; Remesy, C.; Mila, I.; Lapierre, C.; Lazarus, S. A.; Williamson, G.; Scalbert, A. *Am. J. Clin. Nutr.* **2003**, *77*, 912–918.
- Nagayama, K.; Fujita, A. *J. Vitaminol.* **1957**, *3*, 24–29.
- Rechner, A. R.; Kuhnle, G.; Bremner, P.; Hubbard, G. P.; Moore, K. P.; Rice-Evans, C. A. *Free Radical Biol. Med.* **2002**, *33*, 220–235.
- Moore, W. E.; Holdeman, L. V. *Appl. Microbiol.* **1974**, *27*, 961–979.
- Savage, D. C. *Annu. Rev. Microbiol.* **1977**, *31*, 107–133.

PR070431H