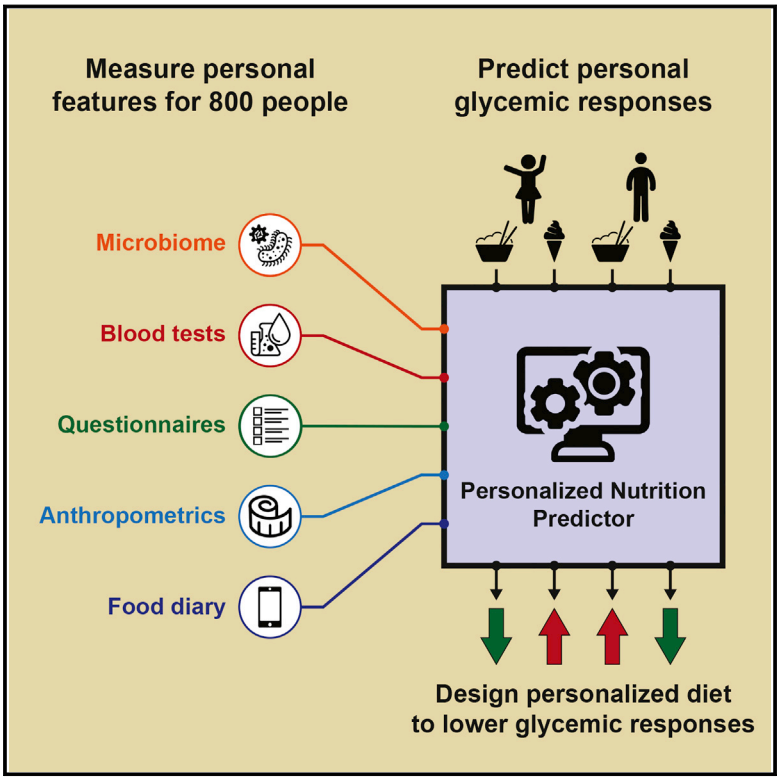
Article

Cell

Personalized Nutrition by Prediction of Glycemic Responses

Graphical Abstract



Highlights

- High interpersonal variability in post-meal glucose observed in an 800-person cohort
- Using personal and microbiome features enables accurate glucose response prediction
- Prediction is accurate and superior to common practice in an independent cohort
- Short-term personalized dietary interventions successfully lower post-meal glucose

Authors

David Zeevi, Tal Korem, Niv Zmora, ..., Zamir Halpern, Eran Elinav, Eran Segal

Correspondence

eran.elinav@weizmann.ac.il (E.E.), eran.segal@weizmann.ac.il (E.S.)

In Brief

People eating identical meals present high variability in post-meal blood glucose response. Personalized diets created with the help of an accurate predictor of blood glucose response that integrates parameters such as dietary habits, physical activity, and gut microbiota may successfully lower post-meal blood glucose and its long-term metabolic consequences.

Zeevi et al., 2015, Cell *163*, 1079–1094
November 19, 2015
http://dx.doi.org/10.1016/j.cell.2015.11.001

CellPress

Cell

Personalized Nutrition by Prediction of Glycemic Responses

David Zeevi,[1,2,8] Tal Korem,[1,2,8] Niv Zmora,[3,4,5,8] David Israeli,[6,8] Daphna Rothschild,[1,2] Adina Weinberger,[1,2] Orly Ben-Yacov,[1,2] Dar Lador,[1,2] Tali Avnit-Sagi,[1,2] Maya Lotan-Pompan,[1,2] Jotham Suez,[3] Jemal Ali Mahdi,[3] Elad Matot,[1,2] Gal Malka,[1,2] Noa Kosower,[1,2] Michal Rein,[1,2] Gili Zilberman-Schapira,[3] Lenka Dohnalová,[3] Meirav Pevsner-Fischer,[3] Rony Bikovsky,[1,2] Zamir Halpern,[5,7] Eran Elinav,[3,9,*] and Eran Segal[1,2,9,*]

[1]Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel
[2]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 7610001, Israel
[3]Immunology Department, Weizmann Institute of Science, Rehovot 7610001, Israel
[4]Internal Medicine Department, Tel Aviv Sourasky Medical Center, Tel Aviv 6423906, Israel
[5]Research Center for Digestive Tract and Liver Diseases, Tel Aviv Sourasky Medical Center, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 6423906, Israel
[6]Day Care Unit and the Laboratory of Imaging and Brain Stimulation, Kfar Shaul Hospital, Jerusalem Center for Mental Health, Jerusalem 9106000, Israel
[7]Digestive Center, Tel Aviv Sourasky Medical Center, Tel Aviv 6423906, Israel
[8]Co-first author
[9]Co-senior author
*Correspondence: eran.elinav@weizmann.ac.il (E.E.), eran.segal@weizmann.ac.il (E.S.)
http://dx.doi.org/10.1016/j.cell.2015.11.001

SUMMARY

Elevated postprandial blood glucose levels constitute a global epidemic and a major risk factor for prediabetes and type II diabetes, but existing dietary methods for controlling them have limited efficacy. Here, we continuously monitored week-long glucose levels in an 800-person cohort, measured responses to 46,898 meals, and found high variability in the response to identical meals, suggesting that universal dietary recommendations may have limited utility. We devised a machine-learning algorithm that integrates blood parameters, dietary habits, anthropometrics, physical activity, and gut microbiota measured in this cohort and showed that it accurately predicts personalized postprandial glycemic response to real-life meals. We validated these predictions in an independent 100-person cohort. Finally, a blinded randomized controlled dietary intervention based on this algorithm resulted in significantly lower postprandial responses and consistent alterations to gut microbiota configuration. Together, our results suggest that personalized diets may successfully modify elevated postprandial blood glucose and its metabolic consequences.

INTRODUCTION

Blood glucose levels are rapidly increasing in the population, as evident by the sharp incline in the prevalence of prediabetes and impaired glucose tolerance estimated to affect, in the U.S. alone, 37% of the adult population (Bansal, 2015). Prediabetes, characterized by chronically impaired blood glucose responses, is a significant risk factor for type II diabetes mellitus (TIIDM), with up to 70% of prediabetics eventually developing the disease (Nathan et al., 2007). It is also linked to other manifestations, collectively termed the metabolic syndrome, including obesity, hypertension, non-alcoholic fatty liver disease, hypertriglyceridemia, and cardiovascular disease (Grundy, 2012). Thus, maintaining normal blood glucose levels is considered critical for preventing and controlling the metabolic syndrome (Riccardi and Rivellese, 2000).

Dietary intake is a central determinant of blood glucose levels, and thus, in order to achieve normal glucose levels it is imperative to make food choices that induce normal postprandial (postmeal) glycemic responses (PPGR; Gallwitz, 2009). Postprandial hyperglycemia is an independent risk factor for the development of TIIDM (American Diabetes Association., 2015a), cardiovascular disease (Gallwitz, 2009), and liver cirrhosis (Nishida et al., 2006) and is associated with obesity (Blaak et al., 2012), and enhanced all-cause mortality in both TIIDM (Cavalot et al., 2011) and cancer (Lamkin et al., 2009).

Despite their importance, no method exists for predicting PPGRs to food. The current practice is to use the meal carbohydrate content (American Diabetes Association., 2015b; Bao et al., 2011), even though it is a poor predictor of the PPGR (Conn and Newburgh, 1936). Other methods aimed at estimating PPGRs are the glycemic index, which quantifies PPGR to consumption of a single tested food type, and the derived glycemic load (Jenkins et al., 1981). It thus has limited applicability in assessing the PPGR to real-life meals consisting of arbitrary food combinations and varying quantities (Dodd et al., 2011), consumed at different times of the day and at different proximity to physical activity and other meals. Indeed, studies examining the effect of diets with a low glycemic index on TIIDM risk, weight loss, and cardiovascular risk factors yielded mixed results (Greenwood et al., 2013; Kristo et al., 2013; Schwingshackl and Hoffmann, 2013).

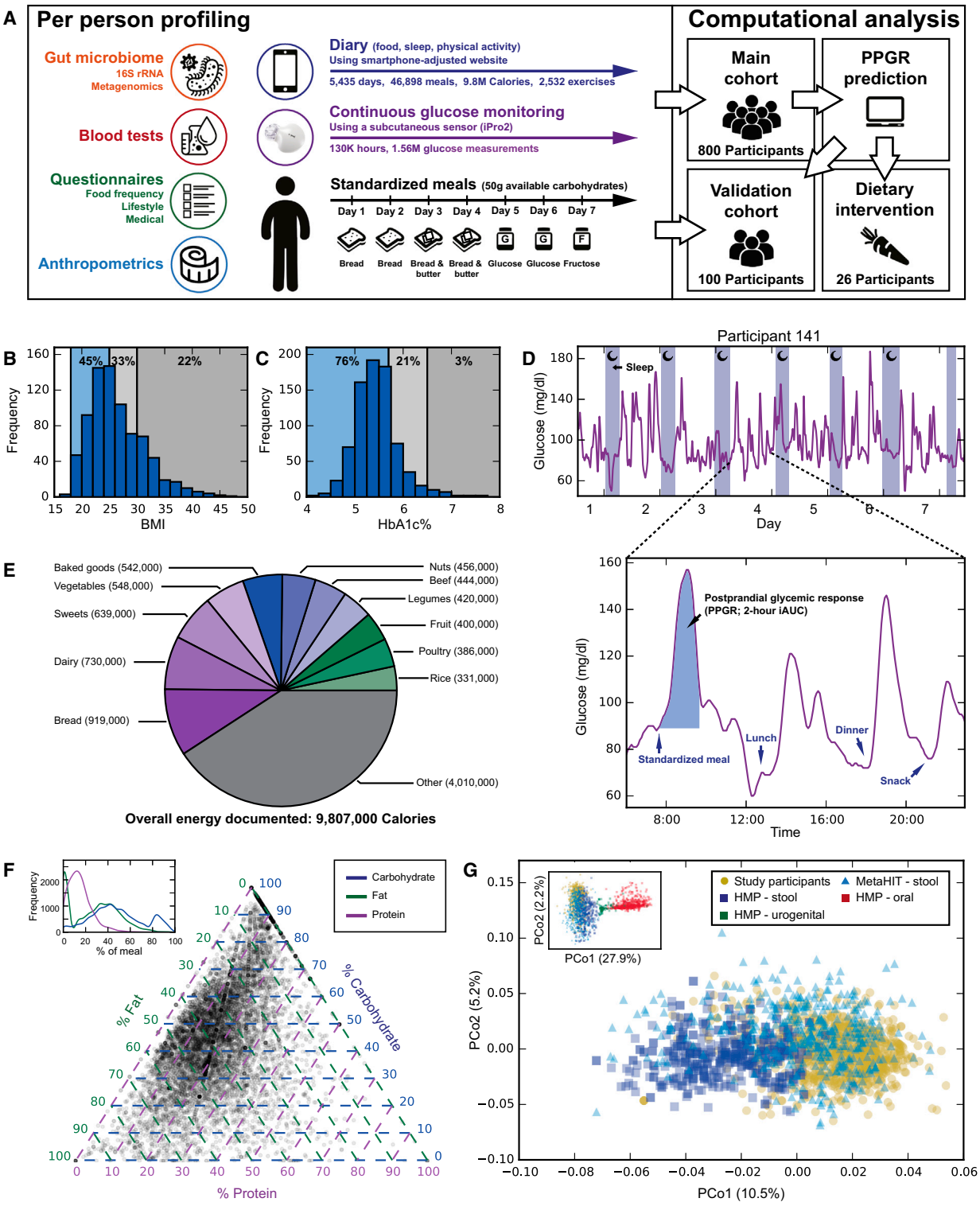


Figure 1. Profiling of Postprandial Glycemic Responses, Clinical Data, and Gut Microbiome

(A) Illustration of our experimental design.

(B and C) Distribution of BMI and glycated hemoglobin (HbA1c%) in our cohort. Thresholds for overweight (BMI $\geq$ 25 kg/m^2), obese (BMI $\geq$ 30 kg/m^2), prediabetes (HbA1c% $\geq$ 5.7%) and TIIDM ($\geq$6.5%) are shown.

(legend continued on next page)

Table 1. Cohorts Description

	Main Cohort	Validation Cohort	KS p Value
Number of participants (n)	800	100	
Sex (% female)	60%	60%	1
Age (y) Mean ± SD	43.3 ± 13.1	42.4 ± 12.6	0.972
BMI (kg/m^2) Mean ± SD	26.4 ± 5.1	26.5 ± 4.8	0.867
BMI ≥ 25	428 (54%)	50 (50%)	
BMI ≥ 30	173 (22%)	18 (18%)	
HbA1c% Mean ± SD	5.43 ± 0.45	5.50 ± 0.55	0.492
HbA1c% ≥ 5.7	189 (24%)	31 (31%)	
HbA1c% ≥ 6.5	23 (3%)	3 (3%)	
Total cholesterol (non-fasting, mg/dl) Mean ± SD	186.8 ± 37.5	182.7 ± 35.7	0.231
HDL cholesterol (non-fasting, mg/dl) Mean ± SD	59.0 ± 17.8	55.0 ± 16.1	0.371
Waist-to-hip circumference ratio Mean ± SD	0.83 ± 0.12	0.84 ± 0.07	0.818

KS - Kolmogorov-Smirnov test. See also Figure S1.

More broadly, ascribing a single PPGR to each food assumes that the response is solely an intrinsic property of the consumed food. However, the few small-scale (n = 23–40) studies that examined interpersonal differences in PPGRs found high variability in the response of different people to the same food (Vega-López et al., 2007; Vrolix and Mensink, 2010), but the factors underlying this variability have not been systematically studied.

Factors that may affect interpersonal differences in PPGRs include genetics (Carpenter et al., 2015), lifestyle (Dunstan et al., 2012), insulin sensitivity (Himsworth, 1934), and exocrine pancreatic and glucose transporters activity levels (Gibbs et al., 1995). Another factor that may be involved is the gut microbiota. Pioneering work by Jeffrey Gordon and colleagues previously showed that it associates with the propensity for obesity and its complications, and later works also demonstrated associations with glucose intolerance, TIIDM, hyperlipidemia, and insulin resistance (Le Chatelier et al., 2013; Karlsson et al., 2013; Qin et al., 2012; Suez et al., 2014; Turnbaugh et al., 2006; Zhang et al., 2013). However, little is known about the association of gut microbiota with PPGRs.

Here, we set out to quantitatively measure individualized PPGRs, characterize their variability across people, and identify factors associated with this variability. To this end, we continuously monitored glucose levels during an entire week in a cohort of 800 healthy and prediabetic individuals and also measured blood parameters, anthropometrics, physical activity, and self-reported lifestyle behaviors, as well as gut microbiota composition and function. Our results demonstrate high interpersonal variability in PPGRs to the same food. We devised a machine learning algorithm that integrates these multi-dimensional data and accurately predicts personalized PPGRs, which we further validated in an independently collected 100-person cohort. Moreover, we show that personally tailored dietary interventions based on these predictions result in significantly improved PPGRs accompanied by consistent alterations to the gut microbiota.

RESULTS

Measurements of Postprandial Responses, Clinical Data, and Gut Microbiome

To comprehensively characterize PPGRs, we recruited 800 individuals aged 18–70 not previously diagnosed with TIIDM (Figure 1A, Table 1). The cohort is representative of the adult non-diabetic Israeli population (Israeli Center for Disease Control, 2014), with 54% overweight (BMI ≥ 25 kg/m^2) and 22% obese (BMI ≥ 30 kg/m^2, Figures 1B, 1C, and S1). These properties are also characteristic of the Western adult non-diabetic population (World Health Organization, 2008).

Each participant was connected to a continuous glucose monitor (CGM), which measures interstitial fluid glucose every 5 min for 7 full days (the "connection week"), using subcutaneous sensors (Figure 1D). CGMs estimate blood glucose levels with high accuracy (Bailey et al., 2014) and previous studies found no significant differences between PPGRs extracted from CGMs and those obtained from either venous or capillary blood (Vrolix and Mensink, 2010). We used blinded CGMs and thus participants were unaware of their CGM levels during the connection week. Together, we recorded over 1.5 million glucose measurements from 5,435 days.

While connected to the CGM, participants were instructed to log their activities in real-time, including food intake, exercise and sleep, using a smartphone-adjusted website (www.personalnutrition.org) that we developed (Figure S2A). Each food item within every meal was logged along with its weight by selecting it from a database of 6,401 foods with full nutritional values based on the Israeli Ministry of Health database that we further improved and expanded with additional items from certified sources. To increase compliance, participants were informed that accurate logging is crucial for them to receive an accurate analysis of their PPGRs to food (ultimately provided to each of them). During the connection week, participants were asked to follow their normal daily routine and dietary habits, except for the first meal of every day, which we provided as one of four different types of standardized meals, each consisting of 50 g of available carbohydrates. This resulted in a total of 46,898 real-life meals with close-to or full nutritional values (median of 54

(D) Example of continuous glucose monitoring (CGM) for one participant during an entire week. Colored area within zoom-in shows the incremental area under the glucose curve (iAUC) which we use to quantify the meal's PPGR.
(E) Major food components consumed by energy intake.
(F) Distribution of meals (dots) by macronutrient content. Inset shows histogram of meals per macronutrient.
(G) Bray-Curtis based PCoA of metagenome-based bacterial abundances of stool samples in our cohort and in the U.S. HMP and European MetaHIT cohorts. Inset shows PCoA when samples from other HMP body sites are added. See also Figure S2.

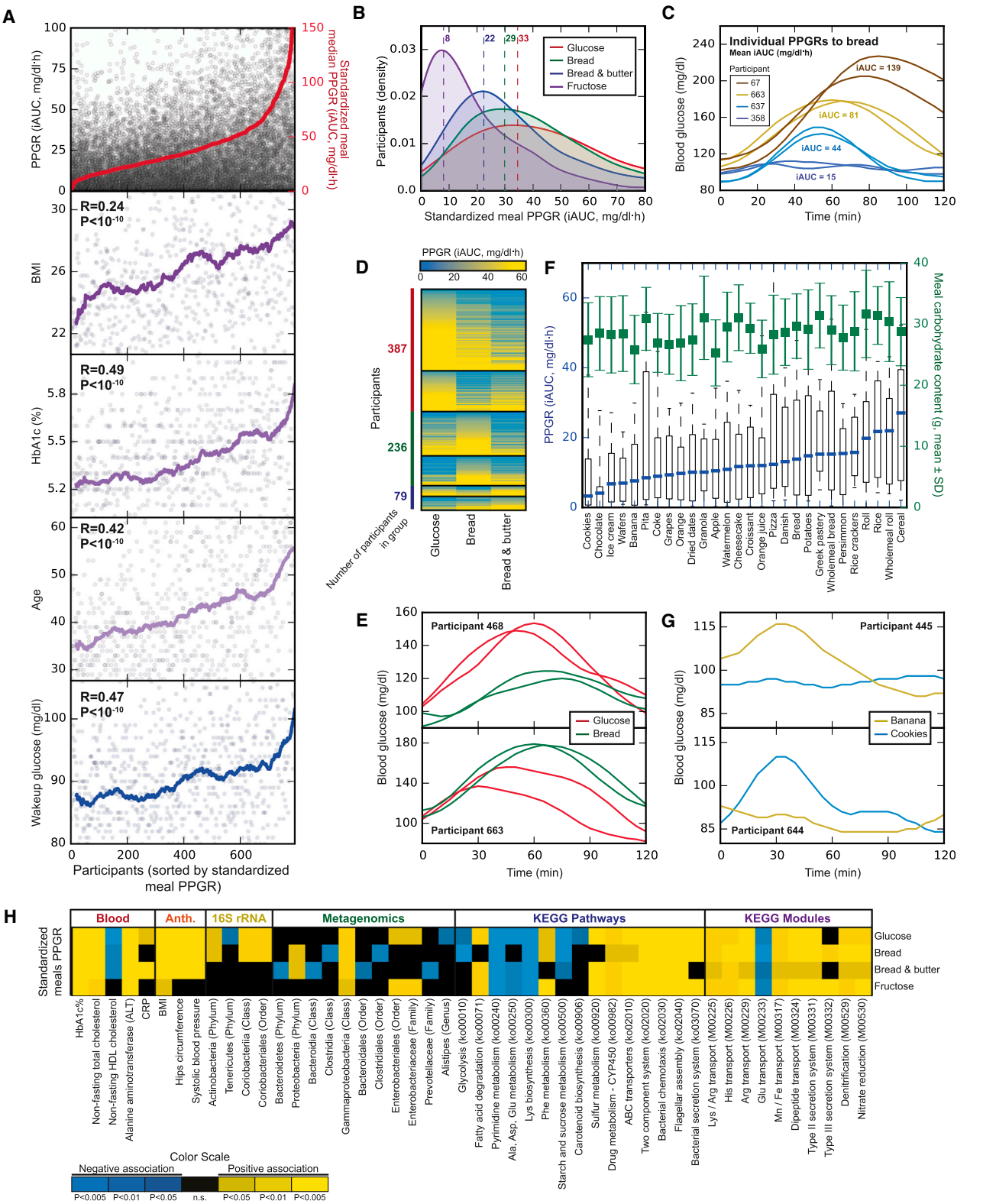


Figure 2. High Interpersonal Variability in the Postprandial Glycemic Response to the Same Meal

(A) PPGRs associate with risk factors. Shown are PPGRs, BMI, HbA1c%, age, and wakeup glucose of all participants, sorted by median standardized meal PPGR (top, red dots). Correlation of factors with the median PPGRs to standardized meals is shown along with a moving average line.

(legend continued on next page)

meals per participant) and 5,107 standardized meals. The PPGR of each meal was calculated by combining reported meal time with CGM data and computing the incremental area under the glucose curve in the 2 hr after the meal (iAUC; Wolever and Jenkins, 1986; Figure 1D).

Prior to CGM connection, a comprehensive profile was collected from each participant, including: food frequency, lifestyle, and medical background questionnaires; anthropometric measures (e.g., height, hip circumference); a panel of blood tests; and a single stool sample, used for microbiota profiling by both 16S rRNA and metagenomic sequencing.

With a total of ~10,000,000 Calories logged, our data provide a global view into the cohort's dietary habits, showing the fraction that each food source contributes to the cohort's overall energy intake (e.g., dairy, 7%; sweets, 6%; Figure 1E), and macronutrient intake (Figures S2B–S2D). Analysis of the caloric breakdown of every meal by macronutrients revealed that protein intake varies relatively little across meals (80% of meals have 5%–35% protein), while fat and carbohydrates have a wide and bimodal distribution, where one of the modes corresponds to fat-free meals and constitutes 18% of all meals (Figure 1F).

Principal coordinates analysis (PCoA) on the Bray-Curtis dissimilarity between metagenome-based relative abundances (RA) revealed a similar degree of variability in the microbiomes of our cohort and stool samples of the US HMP (Human Microbiome Project Consortium, 2012) and European MetaHIT (Nielsen et al., 2014) cohorts (Figure 1G). The first two principal coordinates show some distinction between our cohort and the other cohorts, but when HMP samples from other body sites are added to the PCoA, stool samples from all three cohorts cluster together and separate from the rest, indicative of overall similarity in the gut microbiota composition of individuals from these three distinct geographical regions (Figure 1G).

Postprandial Glycemic Responses Associate with Multiple Risk Factors

Our data replicate known associations of PPGRs with risk factors, as the median standardized meal PPGR was significantly correlated with several known risk factors including BMI ($R = 0.24$, $p < 10^{-10}$), glycated hemoglobin (HbA1c%, $R = 0.49$, $p < 10^{-10}$), wakeup glucose ($R = 0.47$, $p < 10^{-10}$), and age ($R = 0.42$, $p < 10^{-10}$, Figure 2A). These associations are not confined to extreme values but persist along the entire range of PPGR values, suggesting that the reduction in levels of risk factors is continuous across all postprandial values, with lower values associated with lower levels of risk factors even within the normal value ranges (Figure 2A).

Utilizing the continuous nature of the CGMs, we also examined the association between risk factors and the glucose level of each participant at different percentiles (0–100) with respect to all glucose measurements from the connection week. These levels are affected by the PPGRs while also reflecting the general glycemic control state of the participant. All percentiles significantly associated with risk factors (wakeup glucose, BMI, HbA1c%, and age; Figures S3A–S3D). The percentile at which the glucose level correlation was highest varied across risk factors. For example, BMI had the highest correlation with the 40th glucose value percentile, whereas for HbA1c% percentile 95 had the highest correlation (Figures S3A and S3C). These results suggest that the entire range of glucose levels of an individual may have clinical relevance, with different percentiles being more relevant for particular risk factors.

High Interpersonal Variability in the Postprandial Response to Identical Meals

Next, we examined intra- and interpersonal variability in the PPGR to the same food. First, we assessed the extent to which PPGRs to three types of standardized meals that were given twice to every participant (Figure 1A), are reproducible within the same person. Indeed, the two replicates showed high agreement ($R = 0.77$ for glucose, $R = 0.77$ for bread with butter, $R = 0.71$ for bread, $p < 10^{-10}$ in all cases), demonstrating that the PPGR to identical meals is reproducible within the same person and that our experimental system reliably measures this reproducibility. However, when comparing the PPGRs of different people to the same meal, we found high interpersonal variability, with the PPGRs of every meal type (except fructose) spanning the entire range of PPGRs measured in our cohort (Figures 2B, 2C, and S3E–S3H). For example, the average PPGR to bread across 795 people was 44 ± 31 mg/dl*h (mean ± SD), with the bottom 10% of participants exhibiting an average PPGR below 15 mg/dl*h and the top 10% of participants exhibiting an average PPGR above 79 mg/dl*h. The large interpersonal differences in PPGRs are also evident in that the type of meal that induced the highest PPGR differs across participants and that different participants might have opposite PPGRs to pairs of different standardized meals (Figures 2D and 2E).

Interpersonal variability was not merely a result of participants having high PPGRs to all meals, since high variability was also observed when the PPGR of each participant was normalized to his/her own PPGR to glucose (Figures S3I–S3K). For white bread and fructose, for which such normalized PPGRs were previously measured, the mode of the PPGR distribution in our cohort had excellent agreement with published values (Foster-Powell et al.,

(B) Kernel density estimation (KDE) smoothed histogram of the PPGR to four types of standardized meals provided to participants (each with 50 g of available carbohydrates). Dashed lines represent histogram modes (See also Figure S3).

(C) Example of high interpersonal variability and low intra-personal variability in the PPGR to bread across four participants (two replicates per participant consumed on two different mornings).

(D) Heatmap of PPGR (average of two replicates) of participants (rows) to three types of standardized meals (columns) consumed in replicates. Clustering is by each participant's relative rankings of the three meal types.

(E) Example of two replicates of the PPGR to two standardized meals for two participants exhibiting reproducible yet opposite PPGRs.

(F) Box plot (box, IQR; whiskers, 10–90 percentiles) of the PPGR to different real-life meals along with amount of carbohydrates consumed (green; mean ± std).

(G) Same as (E), for a pair of real-life meals, each containing 20 g of carbohydrates.

(H) Heatmap (subset) of statistically significant associations ($p < 0.05$, FDR corrected) between participants' standardized meals PPGRs and participants' clinical and microbiome data (See also Figure S4 for the full heatmap).

2002), further validating the accuracy of our data (bread: 65 versus 71; fructose: 15 versus 19, Figures S3I and S3K).

Next, we examined variability in the PPGRs to the multiple real-life meals reported by our participants. Since real-life meals vary in amounts and may each contain several different food components, we only examined meals that contained 20–40 g of carbohydrates and had a single dominant food component whose carbohydrate content exceeded 50% of the meal's carbohydrate content. We then ranked the resulting dominant foods that had at least 20 meal instances by their population-average PPGR (Figure 2F). For foods with a published glycemic index, our population-average PPGRs agreed with published values ($R = 0.69$, $p < 0.0005$), further supporting our data (Table S1). For example, the average PPGR to rice and potatoes was relatively high, whereas that for ice cream, beer, and dark chocolate was relatively low, in agreement with published data (Atkinson et al., 2008; Foster-Powell et al., 2002). Similar to standardized meals, PPGRs to self-reported meals highly varied across individuals, with both low and high responders noted for each type of meal (Figures 2F and 2G).

Postprandial Variability Is Associated with Clinical and Microbiome Profiles

We found multiple significant associations between the standardized meal PPGRs of participants and both their clinical and gut microbiome data (Figures 2H and S4). Notably, the TIIDM and metabolic syndrome risk factors HbA1c%, BMI, systolic blood pressure, and alanine aminotransferase (ALT) activity are all positively associated with PPGRs to all types of standardized meals, reinforcing the medical relevance of PPGRs. In most standardized meals, PPGRs also exhibit a positive correlation with CRP, whose levels rise in response to inflammation (Figure 2H).

With respect to microbiome features, the phylogenetically related Proteobacteria and Enterobacteriaceae both exhibit positive associations with a few of the standardized meals PPGR (Figure 2H). These taxa have reported associations with poor glycemic control, and with components of the metabolic syndrome including obesity, insulin resistance, and impaired lipid profile (Xiao et al., 2014). RAs of Actinobacteria are positively associated with the PPGR to both glucose and bread, which is intriguing since high levels of this phylum were reported to associate with a high-fat, low-fiber diet (Wu et al., 2011).

At the functional level, the KEGG pathways of bacterial chemotaxis and of flagellar assembly, reported to increase in mice fed high-fat diets and decrease upon prebiotics administration (Everard et al., 2014), exhibit positive associations with several standardized meal PPGRs (Figure 2H). The KEGG pathway of ABC transporters, reported to be positively associated with TIIDM (Karlsson et al., 2013) and with a Western high-fat/high-sugar diet (Turnbaugh et al., 2009), also exhibits positive association with several standardized meal PPGRs (Figure 2H). Several bacterial secretion systems, including both type II and type III secretion systems that are instrumental in bacterial infection and quorum sensing (Sandkvist, 2001) are positively associated with most standardized meal PPGRs (Figure 2H). Finally, KEGG modules for transport of the positively charged amino acids lysine and arginine are associated with high PPGR to standardized foods, while transport of the negatively charged amino acid glutamate is associated with low PPGRs to these foods.

Taken together, these results show that PPGRs vary greatly across different people and associate with multiple person-specific clinical and microbiome factors.

Prediction of Personalized Postprandial Glycemic Responses

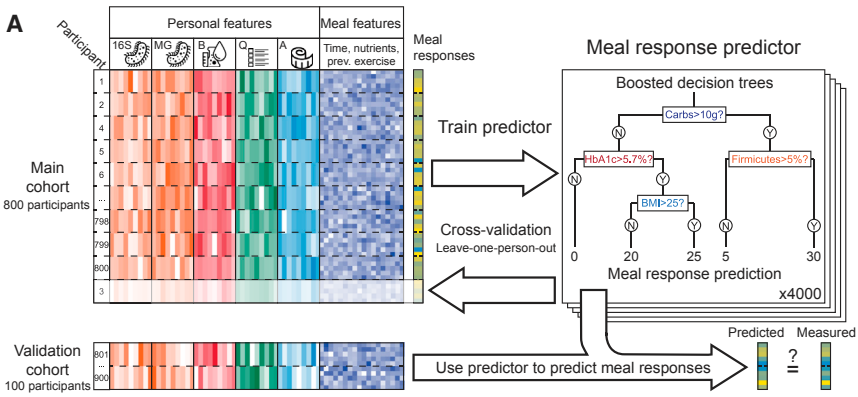
We next asked whether clinical and microbiome factors could be integrated into an algorithm that predicts individualized PPGRs. To this end, we employed a two-phase approach. In the first, discovery phase, the algorithm was developed on the main cohort of 800 participants, and performance was evaluated using a standard leave-one-out cross validation scheme, whereby PPGRs of each participant were predicted using a model trained on the data of all other participants. In the second, validation phase, an independent cohort of 100 participants was recruited and profiled, and their PPGRs were predicted using the model trained only on the main cohort (Figure 3A).

Given non-linear relationships between PPGRs and the different factors, we devised a model based on gradient boosting regression (Friedman, 2001). This model predicts PPGRs using the sum of thousands of different decision trees. Trees are inferred sequentially, with each tree trained on the residual of all previous trees and making a small contribution to the overall prediction (Figure 3A). The features within each tree are selected by an inference procedure from a pool of 137 features representing meal content (e.g., energy, macronutrients, micronutrients); daily activity (e.g., meals, exercises, sleep times); blood parameters (e.g., HbA1c%, HDL cholesterol); CGM-derived features; questionnaires; and microbiome features (16S rRNA and metagenomic RAs, KEGG pathway and module RAs and bacterial growth dynamics - PTRs; Korem et al., 2015).

As a baseline reference, we used the "carbohydrate counting" model, as it is the current gold standard for predicting PPGRs (American Diabetes Association., 2015b; Bao et al., 2011). On our data, this model that consists of a single explanatory variable representing the meal's carbohydrate amount achieves a modest yet statistically significant correlation with PPGRs ($R = 0.38$, $p < 10^{-10}$, Figure 3B). A model using only meal Caloric content performs worse ($R = 0.33$, $p < 10^{-10}$, Figure 3C). Our predictor that integrates the above person-specific factors predicts the held-out PPGRs of individuals with a significantly higher correlation ($R = 0.68$, $p < 10^{-10}$, Figure 3D). This correlation approaches the presumed upper bound limit set by the 0.71–0.77 correlation that we observed between the PPGR of the same person to two replicates of the same standardized meal.

Validation of Personalized Postprandial Glycemic Response Predictions on an Independent Cohort

We further validated our model on an independent cohort of 100 individuals that we recruited separately. Data from this additional cohort were not available to us while developing the algorithm. Participants in this cohort underwent the same profiling as in the main 800-person cohort. No significant differences were found between the main and validation cohorts in key parameters, including age, BMI, non-fasting total and HDL cholesterol, and HbA1c% (Table 1, Figure S1).



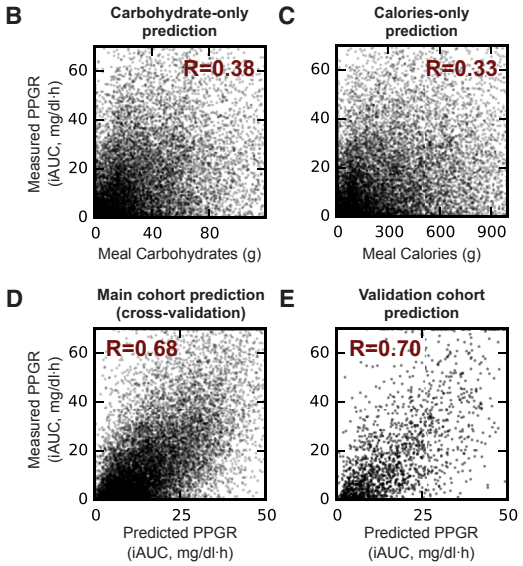


Figure 3. Accurate Predictions of Personalized Postprandial Glycemic Responses

(A) Illustration of our machine-learning scheme for predicting PPGRs.

(B–E) PPGR predictions. Dots represent predicted (x axis) and CGM-measured PPGR (y axis) for meals, for a model based: only on the meal's carbohydrate content (B); only on the meal's Caloric content (C); our predictor evaluated in leave-one-person-out cross validation on the main 800-person cohort (D); and our predictor evaluated on the independent 100-person validation cohort (E). Pearson correlation of predicted and measured PPGRs is indicated.

Notably, our algorithm, derived solely using the main 800 participants cohort, achieved similar performance on the 100 participants of the validation cohort (R = 0.68 and R = 0.70 on the main and validation cohorts, respectively, Figures 3D and 3E). The reference carbohydrate counting model achieved the same performance as in the main cohort (R = 0.38). This result further supports the ability of our algorithm to provide personalized PPGR predictions.

Factors Underlying Personalized Predictions

To gain insight into the contribution of the different features in the algorithm's predictions, we examined partial dependence plots (PDP), commonly used to study functional relations between features used in predictors such as our gradient boosting regressor and an outcome (PPGRs in our case; Hastie et al., 2008). PDPs graphically visualize the marginal effect of a given feature on prediction outcome after accounting for the average effect of all other features. While this effect may be indicative of feature importance, it may also be misleading due to higher-order interactions (Hastie et al., 2008). Nonetheless, PDPs are commonly used for knowledge discovery in large datasets such as ours.

As expected, the PDP of carbohydrates (Figure 4A) shows that as the meal carbohydrate content increases, our algorithm predicts, on average, a higher PPGR. We term this relation, of higher predicted PPGR with increasing feature value, as non-beneficial (with respect to prediction), and the opposite relation, of lower predicted PPGR with increasing feature value, as beneficial (also with respect to prediction; see PDP legend in Figure 4). However, since PDPs display the overall contribution of each feature across the entire cohort, we asked whether the relationship between carbohydrate amount and PPGRs varies across people. To this end, for each participant we computed the slope of the linear regression between the PPGR and carbohydrate amount of all his/her meals. As expected, this slope was positive for nearly all (95.1%) participants, reflective of higher PPGRs in meals richer in carbohydrates. However, the magnitude of this slope varies greatly across the cohort, with the PPGR of some people correlating well with the carbohydrate content (i.e., carbohydrates "sensitive") and that of others exhibiting equally high PPGRs but little relationship to the amount of carbohydrates (carbohydrate "insensitive"; Figure 4B). This result suggests that carbohydrate sensitivity is also person specific.

The PDP of fat exhibits a beneficial effect for fat since our algorithm predicts, on average, lower PPGR as the meal's ratio of fat to carbohydrates (Figure 4C) or total fat content (Figure S5A) increases, consistent with studies showing that adding fat to meals may reduce the PPGR (Cunningham and Read, 1989). However, here too, we found that the effect of fat varies across people. We compared the explanatory power of a linear regression between each participant's PPGR and meal carbohydrates, with that of regression using both fat and carbohydrates. We then used the difference in Pearson R between the two models as a quantitative measure of the added contribution of fat (Figure 4D). For some participants we observed a reduction in PPGR with the addition of fat, while for others meal fat content

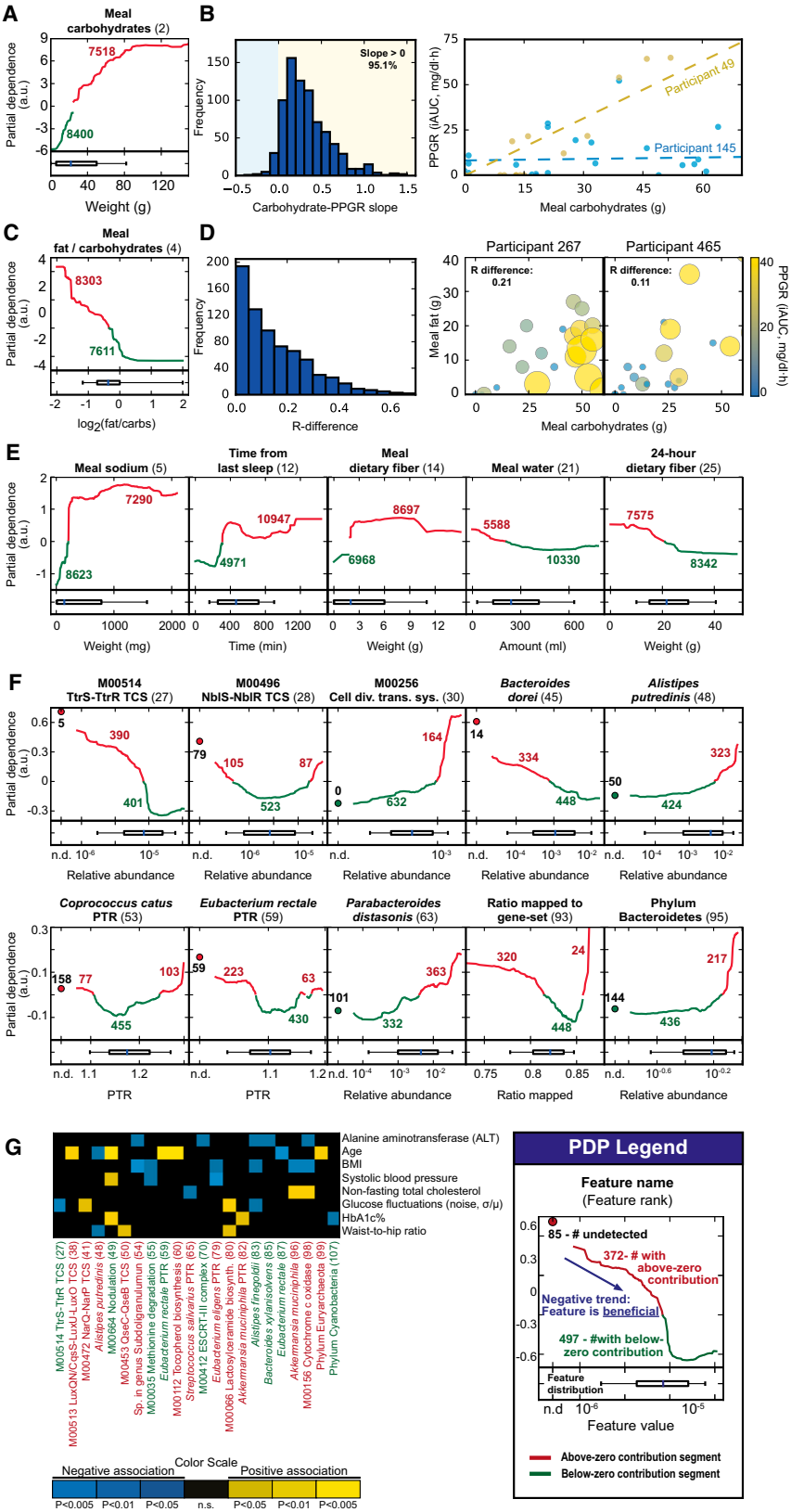


Figure 4. Factors Underlying the Prediction of Postprandial Glycemic Responses

(A) Partial dependence plot (PDP) showing the marginal contribution of the meal's carbohydrate content to the predicted PPGR (y axis, arbitrary units) at each amount of meal carbohydrates (x axis). Red and green indicate above and below zero contributions, respectively (number indicate meals). Boxplots (bottom) indicate the carbohydrates content at which different percentiles (10, 25, 50, 75, and 90) of the distribution of all meals across the cohort are located. See PDP legend.

(B) Histogram of the slope (computed per participant) of a linear regression between the carbohydrate content and the PPGR of all meals. Also shown is an example of one participant with a low slope and another with a high slope.

(C) Meal fat/carbohydrate ratio PDP.

(D) Histogram of the difference (computed per participant) between the Pearson R correlation of two linear regression models, one between the PPGR and the meal carbohydrate content and another when adding fat and carbohydrate*fat content. Also shown is an example of the carbohydrate and fat content of all meals of one participant with a relatively low R difference (carb alone correlates well with PPGR) and another with a relatively high difference (meals with high fat content have lower PPGRs). Dot color and size correspond to the meal's PPGR.

(E) Additional PDPs.

(F) Microbiome PDPs. The number of participants in which the microbiome feature was not detected is indicated (left, n.d.). Boxplots (box, IQR; whiskers 10–90 percentiles) based only on detected values.

(G) Heatmap of statistically significant correlations (Pearson) between microbiome features termed beneficial (green) or non-beneficial (red) and several risk factors and glucose parameters.

See also Figure S5.

did not add much to the explanatory power of the regressor based only on the meal's carbohydrates content (Figure 4D).

Interestingly, while dietary fibers in the meal increase the predicted PPGR, their long-term effect is beneficial as higher amount of fibers consumed in the 24 hr prior to the meal reduces the predicted PPGR (Figure 4E). The meal's sodium content, the time that passed since last sleeping, and a person's cholesterol levels or age all exhibit non-beneficial PDPs, while the PDPs of the meal's alcohol and water content display beneficial effects (Figures 4E and S5A). As expected, the PDP of HbA1c% shows a non-beneficial effect with increased PPGR at higher HbA1c% values; intriguingly, higher PPGRs are predicted, on average, for individuals with HbA1c% above ~5.5%, which is very close to the prediabetes threshold of 5.7% (Figure S5A).

The 72 PDPs of the microbiome-based features used in our predictor were either beneficial (21 factors), non-beneficial (28), or non-decisive (23) in that they mostly decreased, increased, or neither, as a function of the microbiome feature. The resulting PDPs had several intriguing trends. For example, growth of *Eubacterium rectale* was mostly beneficial, as in 430 participants with high inferred growth for *E. rectale* it associates with a lower PPGR (Figure 4F). Notably, *E. rectale* can ferment dietary carbohydrates and fibers to produce metabolites useful to the host (Duncan et al., 2007), and was associated with improved postprandial glycemic and insulinemic responses (Martínez et al., 2013), as well as negatively associated with TIIDM (Qin et al., 2012). RAs of *Parabacteroides distasonis* were found non-beneficial by our predictor (Figure 4F) and this species was also suggested to have a positive association with obesity (Ridaura et al., 2013). As another example, the KEGG module of cell-division transport system (M00256) was non-beneficial, and in the 164 participants with the highest levels for it, it associates with a higher PPGR (Figure 4F). *Bacteroides thetaiotaomicron* was non-beneficial (Figure S5B), and it was associated with obesity and was suggested to have increased capacity for energy harvest (Turnbaugh et al., 2006). In the case of *Alistipes putredinis* and the Bacteroidetes phylum, the non-beneficial classification that our predictor assigns to both of them is inconsistent with previous studies that found them to be negatively associated with obesity (Ridaura et al., 2013; Turnbaugh et al., 2006). This may reflect limitations of the PDP analysis or result from a more complex relationship between these features, obesity, and PPGRs.

To assess the clinical relevance of the microbiome-based PDPs, we computed the correlation between several risk factors and overall glucose parameters, and the factors with beneficial and non-beneficial PDPs across the entire 800-person cohort. We found 20 statistically significant correlations ($p < 0.05$, FDR corrected) where microbiome factors termed non-beneficial correlated with risk factors, and those termed beneficial exhibited an anti-correlation (Figure 4G). For example, higher levels of the beneficial methionine degradation KEGG module (M00035) resulted in lower PPGRs in our algorithm, and across the cohort, this module anti-correlates with systolic blood pressure and with BMI (Figure 4G). Similarly, fluctuations in glucose levels across the connection week correlates with nitrate respiration two-component regulatory system (M00472) and with lactosylceramide biosynthesis (M00066), which were both termed non-beneficial. Glucose fluctuations also anti-correlate with levels of the tetrathionate respiration two-component regulatory system (M00514) and with RAs of *Alistipes finegoldii*, both termed beneficial (Figure 4G). In 14 other cases, factors with beneficial or non-beneficial PDPs were correlated and anti-correlated with risk factors, respectively.

These results suggest that PPGRs are associated with multiple and diverse factors, including factors unrelated to meal content.

Personally Tailored Dietary Interventions Improve Postprandial Responses

Next, we asked whether personally tailored dietary interventions based on our algorithm could improve PPGRs. We designed a two-arm blinded randomized controlled trial and recruited 26 new participants. A clinical dietitian met each participant and compiled 4–6 distinct isocaloric options for each type of meal (breakfast, lunch, dinner, and up to two intermediate meals), accommodating the participant's regular diet, eating preferences, and dietary constraints. Participants then underwent the same 1-week profiling of our main 800-person cohort (except that they consumed the meals compiled by the dietitian), thus providing the inputs (microbiome, blood parameters, CGM, etc.) that our algorithm needs for predicting their PPGRs.

Participants were then blindly assigned to one of two arms (Figure 5A). In the first, "prediction arm," we applied our algorithm in a leave-one-out scheme to rank every meal of each participant in the profiling week (i.e., the PPGR to each predicted meal was hidden from the predictor). We then used these rankings to design two 1-week diets: (1) a diet composed of the meals predicted by the algorithm to have low PPGRs (the "good" diet); and (2) a diet composed of the meals with high predicted PPGRs (the "bad" diet). Every participant then followed each of the two diets for a full week, during which they were connected to a CGM and a daily stool sample was collected (if available). The order of the 2 diet weeks was randomized for each participant and the identity of the intervention weeks (i.e., whether they are "good" or "bad") was kept blinded from CRAs, dietitians and participants.

The second, "expert arm," was used as a gold standard for comparison. Participants in this arm underwent the same process as the prediction arm except that instead of using our predictor for selecting their "good" and "bad" diets a clinical dietitian and a researcher experienced in analyzing CGM data (collectively termed "expert") selected them based on their measured PPGRs to all meals during the profiling week. Specifically, meals that according to the expert's analysis of their CGM had low and high PPGRs in the profiling week were selected for the "good" and "bad" diets, respectively. Thus, to the extent that PPGRs are reproducible within the same person, this expert-based arm should result in the largest differences between the "good" and "bad" diets because the selection of meals in the intervention weeks is based on their CGM data.

Notably, for 10 of the 12 participants of the predictor-based arm, PPGRs in the "bad" diet were significantly higher than in the "good" diet ($p < 0.05$, Figure 5C). Differences between the two diets are also evident in fewer glucose spikes and fewer fluctuations in the raw week-long CGM data (Figure 5B). The

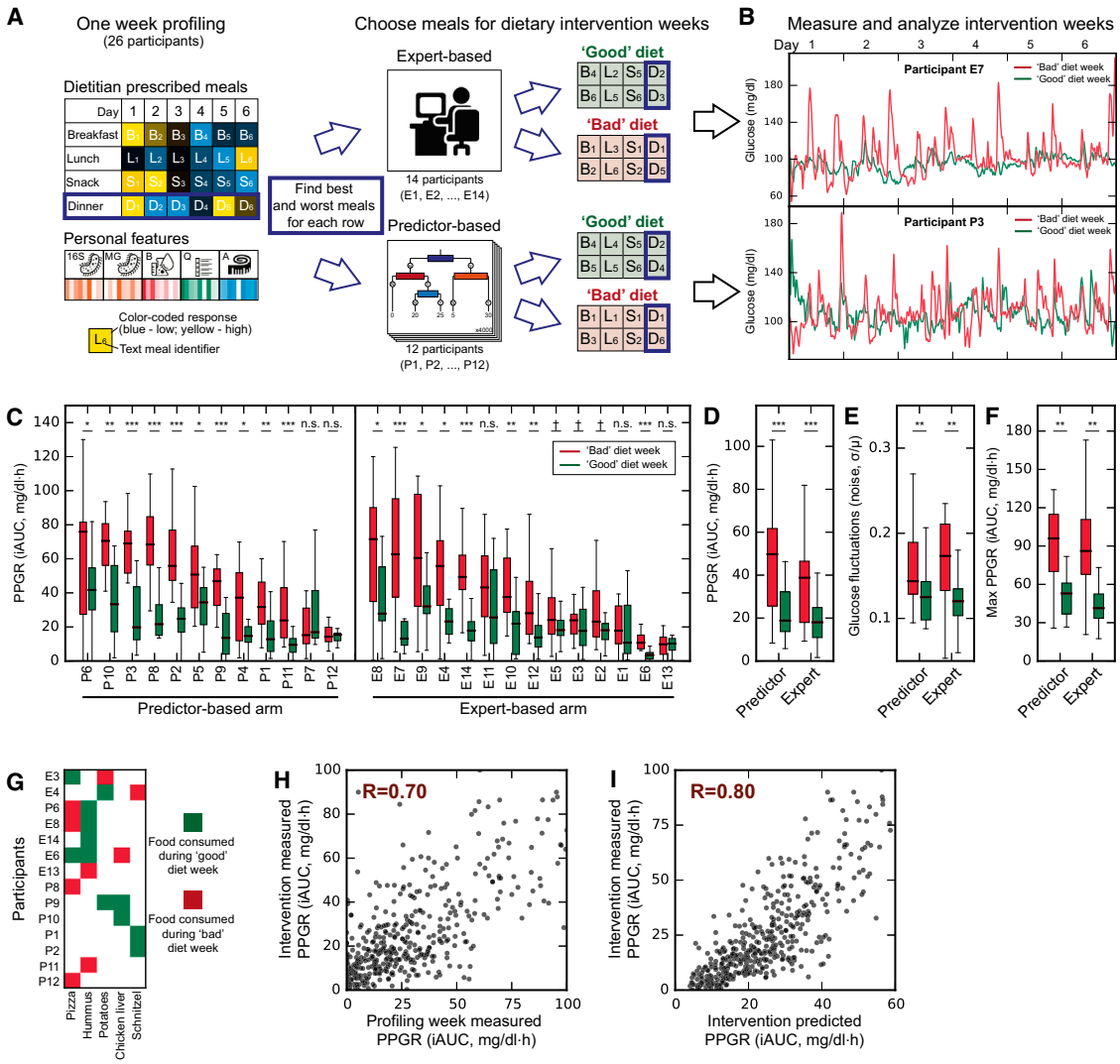


Figure 5. Personally Tailored Dietary Interventions Improve Postprandial Glycemic Responses

(A) Illustration of the experimental design of our two-arm blinded randomized controlled trial.

(B) Continuous glucose measurements of one participant from the expert arm (top) and another from the predictor arm (bottom) across their "good" (green) and "bad" (red) diet weeks.

(C) Boxplot of meal PPGRs during the "bad" (red) and "good" (green) diet weeks for participants in both the predictor (left) and expert (right) arms. Statistical significance is marked (Mann-Whitney *U*-test, ***p < 0.001, **p < 0.01, *p < 0.05, † p < 0.1, n.s. not significant).

(D) As in (C), but for a grouping of all meals of all participants in each study arm (p, Wilcoxon signed-rank test).

(E) Boxplot of the blood glucose fluctuations (noise) of participants in both the "bad" (red) and "good" (green) diet weeks for both study arms. Blood glucose fluctuations per participant are defined as the ratio between the standard deviation and mean of his/her weeklong blood glucose levels (p, Wilcoxon signed-rank test).

(F) As in (E), but for the maximum PPGR of each participant.

(G) Subset of dominant food components prescribed in the "good" (green) diet of some participants and in the "bad" (red) diet of other participants. See also Figure S6 for the full matrix.

(H) Dot plot between the CGM-measured PPGR of meals during the profiling week (x axis) and the average CGM-measured PPGR of the same meals during the dietary intervention weeks (y axis). Meals of all participants in both study arms are shown.

(I) As in (H), but when PPGRs in the dietary intervention weeks are predicted by our predictor using only the first profiling week data of each participant. Boxplots - box, IQR; whiskers 1.5*IQR.

success of the predictor was comparable to that of the expert-based arm, in which significantly lower PPGRs in the "good" versus the "bad" diet were observed for 8 of its 14 participants ($p < 0.05$, 11 of 14 participants with $p < 0.1$, Figure 5C).

When combining the data across all participants, the "good" diet exhibited significantly lower PPGRs than the "bad" diet ($p < 0.05$, Figure 5D) as well as improvement in other measures of blood glucose metabolism in both study arms, specifically,

lower fluctuations in glucose levels across the CGM connection week ($p < 0.05$, Figure 5E), and a lower maximal PPGR ($p < 0.05$, Figure 5F) in the "good" diet.

Both study arms constitute personalized nutritional interventions and thus demonstrate the efficacy of this approach in lowering PPGRs. However, the predictor-based approach has broader applicability since it can predict PPGRs to arbitrary unseen meals, whereas the "expert"-based approach will always require CGM measurements of the meals it prescribes.

Post hoc examination of the prescribed diets revealed the personalized aspect of the diets in both arms in that multiple dominant food components (as in Figure 2F) prescribed in the "good" diet of some participants were prescribed in the "bad" diet of others (Figures 5G and S6). This occurs when components induced opposite CGM-measured PPGRs across participants (expert arm) or were predicted to have opposite PPGRs (predictor arm).

The correlation between the measured PPGR of meals during the profiling week and the average CGM-measured PPGR of the same meals during the dietary intervention was 0.70 (Figure 5H), which is similar to the reproducibility observed for standardized meals ($R = 0.71–0.77$). Thus, as in the case of standardized meals, a meal's PPGR during the profiling week was not identical to its PPGR in the dietary intervention week. Notably, using only the first profiling week data of each participant, our algorithm predicted the average PPGRs of meals in the dietary intervention weeks with an even higher correlation ($R = 0.80$, Figure 5I). Since our predictor also incorporates context-specific factors (e.g., previous meal content, time since sleep), this result also suggests that such factors may be important determinants of PPGRs.

Taken together, these results show the utility of personally tailored dietary interventions for improving PPGRs in a short-term intervention period, and the ability of our algorithm to devise such interventions.

Alterations in Gut Microbiota Following Personally Tailored Dietary Interventions

Finally, we used the daily microbiome samples collected during the intervention weeks to ask whether the interventions induced significant changes in the gut microbiota. Previous studies showed that even short-term dietary interventions of several days may significantly alter the gut microbiota (David et al., 2014; Korem et al., 2015).

We detected changes following the dietary interventions that were significant relative to a null hypothesis of no change derived from the first week, in which there was no intervention, across all participants (Figures 6A and 6B). While many of these significant changes were person-specific, several taxa changed consistently in most participants ($p < 0.05$, FDR corrected, Figure 6C and S7). Moreover, in most cases in which the consistently changing taxa had reported associations in the literature, the direction of change in RA following the "good" diet was in agreement with reported beneficial associations. For example, *Bifidobacterium adolescentis*, for which low levels were reported to be associated with greater weight loss (Santacruz et al., 2009), generally decrease in RA following the "good" diet and increase following the "bad" diet (Figure 6C,D). Similarly, TIIDM has been associated with low levels of *Roseburia inulinivorans* (Qin et al., 2012; Figure 6E), *Eubacterium eligens* (Karlsson et al., 2013), and *Bacteroides vulgatus* (Ridaura et al., 2013), and all these bacteria increase following the "good" diet and decrease following the "bad" diet (Figure 6C). The Bacteroidetes phylum, for which low levels associate with obesity and high fasting glucose (Turnbaugh et al., 2009), increases following the "good" diet and decreases following the "bad" diet (Figure 6C). Low levels of Anaerostipes associate with improved glucose tolerance and reduced plasma triglyceride levels in mice (Everard et al., 2011) and indeed these bacteria decrease following the "good" diet and increase following the "bad" diet (Figure 6C). Finally, low levels of *Alistipes putredinis* associate with obesity (Ridaura et al., 2013) and this bacteria increased following the "good" diet (Figure 6C).

These findings demonstrate that while both baseline microbiota composition and personalized dietary intervention vary between individuals, several consistent microbial changes may be induced by dietary intervention with a consistent effect on PPGR.

DISCUSSION

In this work we measured 46,898 PPGRs to meals in a population-based cohort of 800 participants. We demonstrate that PPGRs are highly variable across individuals even when they consume the same standardized meals. We further show that an algorithm that integrates clinical and microbiome features can accurately predict personalized PPGRs to complex, real-life meals even in a second independently collected validation cohort of 100 participants. Finally, personalized dietary interventions based on this algorithm induced lower PPGRs and were accompanied by consistent gut microbiota alterations.

Our study focused on PPGRs, as they were shown to be important in achieving proper glycemic control, and when disturbed are considered an independent disease risk factor (American Diabetes Association., 2015a; Gallwitz, 2009). PPGRs in our study also associated with several risk factors, including BMI, HbA1c%, and wakeup glucose. In addition to its centrality in glucose homeostasis, PPGRs serves as a convenient and accurate endpoint, enabling continuous "point-of-care" collection of dozens of quantitative measurements per person during a relatively short follow up period. Such continuous assessment of PPGRs is complementary to other equally important clinical parameters such as BMI and HbA1c%, for which changes typically occur over longer timescales and are thus difficult to correlate to nutritional responses in real time.

In line with few small-scale studies that previously examined individual PPGRs (Vega-López et al., 2007; Vrolix and Mensink, 2010), we demonstrate on 800 individuals that the PPGR of different people to the same food can greatly vary. The most compelling evidence for this observation is the controlled setting of standardized meals, provided to all participants in replicates. This high interpersonal variability suggests that at least with regard to PPGRs, approaches that grade dietary ingredients as universally "good" or "bad" based on their average PPGR in the population may have limited utility for an individual.

We report several associations between microbiome features and variability in PPGRs across people. In some cases, such as for Actinobacteria, Proteobacteria, and Enterobacteriaceae, the

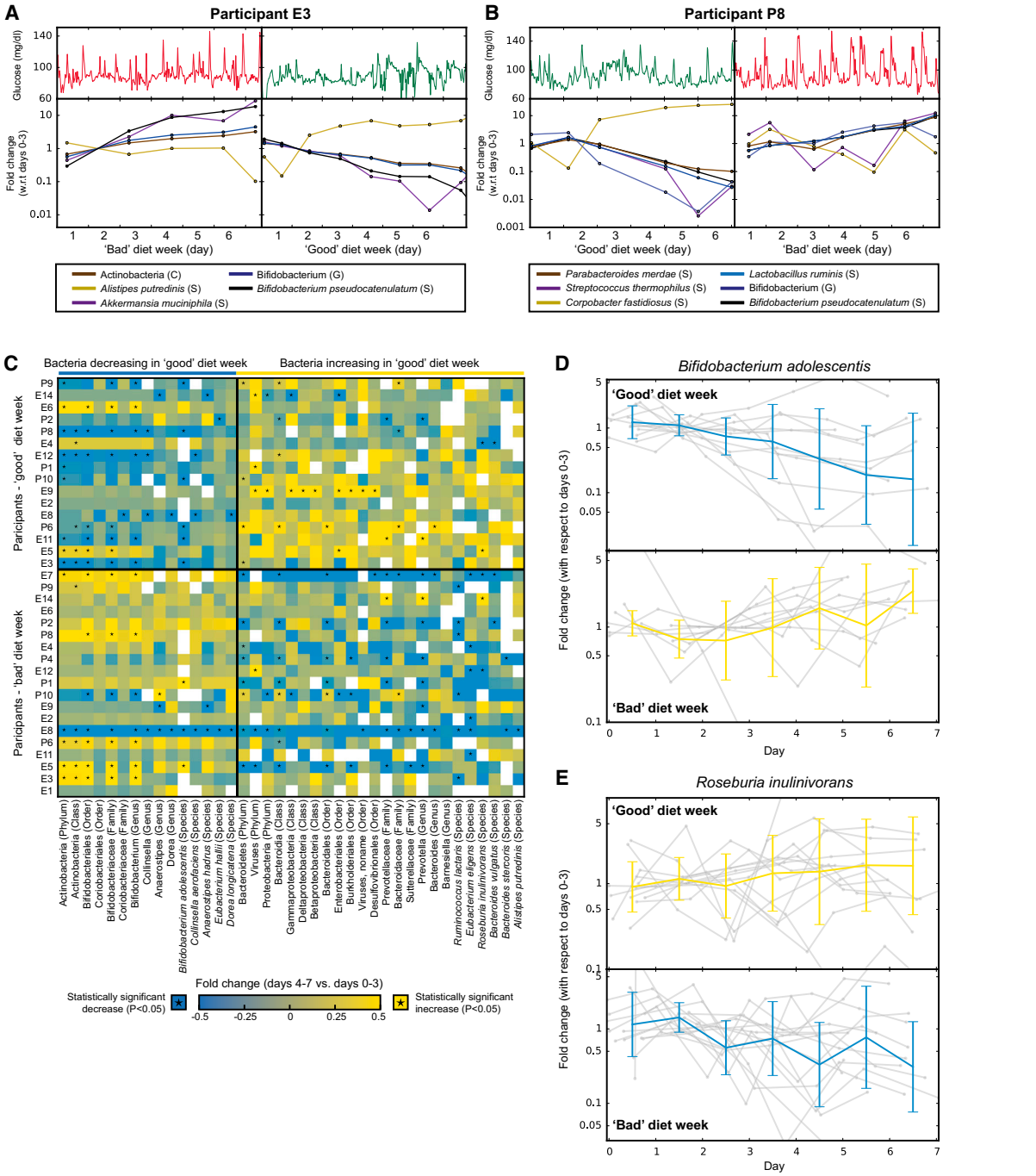


Figure 6. Dietary Interventions Induce Consistent Alterations to the Gut Microbiota Composition

(A) Top: Continuous glucose measurements of a participant from the expert arm for both the "bad" diet (left) and "good" diet (right) week. Bottom: Fold change between the relative abundance (RA) of taxa in each day of the "bad" (left) or "good" (right) weeks and days 0–3 of the same week. Shown are only taxa that exhibit statistically significant changes with respect to a null hypothesis of no change derived from changes in the first profiling week (no intervention) of all participants.

(B) As in (A) for a participant from the predictor arm. See also Figure S7 for changes in all participants.

(C) Heatmap of taxa with opposite trends of change in RA between "good" and "bad" intervention weeks that was consistent across participant and statistically significant (Mann-Whitney *U*-test between changes in the "good" and "bad" weeks, $p < 0.05$, FDR corrected). Left and right column blocks shows bacteria increasing and decreasing in their RA following the "good" diet, respectively, and conversely for the "bad" diet. Colored entries represent the (log) fold change between the RA of a taxon (x axis) between days 4–7 and 0–3 within each participant (y axis). Asterisks indicate a statistically significant fold change. See also Figure S7 for all changes.

(legend continued on next page)

direction of our associations are consistent with previous associations reported between these taxa and higher-level phenotypes such as dietary habits, obesity and overall glycemic control (Wu et al., 2011; Xiao et al., 2014), raising testable hypotheses about how these taxa may mediate these host metabolic effects. However, in most other cases we identify yet unknown associations with particular biosynthesis pathways or transport and secretion systems, which may be contributed by different taxa in different individuals. These correlations thus provide concrete new pointers for further mechanistic research, aimed at establishing causal roles for these bacterial taxa and functional pathways in determining PPGRs.

Our study further attempts to analyze real-life meals that are consumed in complex food combinations, at different times of the day, and in varying proximity to previous meals, physical activity, and sleep. While clearly of higher translational relevance, the use of "real-life" nutritional input also introduces noise into the meal composition data. Despite that, our results show that predictions for such meals can be made informative by integrating data from a large cohort into a carefully structured predictor. Even better predictions can likely be achieved with further research.

Our algorithm takes as input a comprehensive clinical and microbiome profile and employs a data-driven unbiased approach to infer the major factors that are predictive of PPGRs. Introspecting the resulting algorithm shows that its predictions integrate multiple diverse features that are unrelated to the content of the meal itself. These include contents of previous meals, time since sleep, proximity to exercise, and several microbiome-based factors. With respect to microbiome factors, our algorithm identifies multiple functional pathways and bacterial taxa as either beneficial or non-beneficial, such that in participants with increasing levels for these factors the algorithm predicts a lower or higher PPGR, respectively. In many such cases, microbiome factors found to be beneficial with respect to PPGRs are also negatively associated with risk factors such as HbA1c% and cholesterol levels.

Dietary interventions based on our predictor showed significant improvements in multiple aspects of glucose metabolism, including lower PPGRs and lower fluctuations in blood glucose levels within a short 1-week intervention period. It will be interesting to evaluate the utility of such personalized intervention over prolonged periods of several months and even years. If successful, prolonged individualized dietary control of the PPGR may be useful in controlling, ameliorating, or preventing a set of disorders associated with chronically impaired glucose control, including obesity, prediabetes, TIIDM, and non-alcoholic fatty liver disease (Grundy, 2012). These intriguing possibilities, and the microbiome changes that accompany them, merit further studies. Of equal interest and importance, our individualized nutritional study protocols may be applicable to address other clinically relevant issues involving nutritional modifications, such as TIIDM and TIDM patient-specific determination of medication (e.g., insulin and oral hypoglycemics) dosing and timing.

Employing similar individualized prediction of nutritional effects on disease development and progression may also be valuable in rationally designing nutritional interventions in a variety of inflammatory, metabolic, and neoplastic multi-factorial disorders. More broadly, accurate personalized predictions of nutritional effects in these scenarios may be of great practical value, as they will integrate nutritional modifications more extensively into the clinical decision-making scheme.

EXPERIMENTAL PROCEDURES

Human Cohorts

Approved by Tel Aviv Sourasky Medical Center Institutional Review Board (IRB), approval numbers TLV-0658-12, TLV-0050-13 and TLV-0522-10; Kfar Shaul Hospital IRB, approval number 0-73; and Weizmann Institute of Science Bioethics and Embryonic Stem Cell Research oversight committee. Reported to http://clinicaltrials.gov/, NCT: NCT01892956.

Study Design

Study participants were healthy individuals aged 18–70 able to provide informed consent and operate a glucometer. Prior to the study, participants filled medical, lifestyle, and nutritional questionnaires. At connection week start, anthropometric, blood pressure and heart-rate measurements were taken by a CRA or a certified nurse, as well as a blood test. Glucose was measured for 7 days using the iPro2 CGM with Enlite sensors (Medtronic, MN, USA), independently calibrated with the Contour BGM (Bayer AG, Leverkusen, Germany) as required. During that week participants were instructed to record all daily activities, including standardized and real-life meals, in real-time using their smartphones; meals were recorded with exact components and weights. Full inclusion and exclusion criteria are detailed in Supplemental Experimental Procedures. Questionnaires used can be found in Data S1.

Standardized Meals

Participants were given standardized meals (glucose, bread, bread and butter, bread and chocolate, and fructose), calculated to have 50 g of available carbohydrates. Participants were instructed to consume these meals immediately after their night fast, not to modify the meal, and to refrain from eating or performing strenuous physical activity before, and for 2 hr following consumption.

Stool Sample Collection

Participants sampled their stool following detailed printed instructions. Sampling was done using a swab (n = 776) or both a swab and an OMNIgene-GUT (OMR-200; DNA Genotek) stool collection kit (n = 413, relative abundances (RA) for the same person are highly correlated ($R = 0.99$ $p < 10^{-10}$) between swabs and OMNIgene-GUT collection methods). Collected samples were immediately stored in a home freezer (−20°C), and transferred in a provided cooler to our facilities where it was stored at −80°C (−20°C for OMNIgene-GUT kits) until DNA extraction. All samples were taken within 3 days of connection week start.

Genomic DNA Extraction and Filtering

Genomic DNA was purified using PowerMag Soil DNA isolation kit (MoBio) optimized for Tecan automated platform. For shotgun sequencing, 100 ng of purified DNA was sheared with a Covaris E220X sonicator. Illumina compatible libraries were prepared as described (Suez et al., 2014). For 16S rRNA sequencing, PCR amplification of the V3/4 region using the 515F/806R 16S

(D) For *Bifidobacterium adolescentis*, which decreased significantly following the "good" diet interventions (see panel C), shown is the average and standard deviation of the (log) fold change of all participants in each day of the "good" (top) diet week relative to days 0–3 of the "good" week. Same for the "bad" diet week (bottom) in which *B. adolescentis* increases significantly (see panel C). Grey lines show fold changes (log) in individual participants.
(E) As in (D), for *Roseburia inulinivorans*.

rRNA gene primers was performed followed by 500 bp paired-end sequencing (Illumina MiSeq).

Microbial Analysis

We used USearch8.0 (Edgar, 2013) to obtain RA from 16S rRNA reads. We filtered metagenomic reads containing Illumina adapters, filtered low quality reads and trimmed low quality read edges. We detected host DNA by mapping with GEM (Marco-Sola et al., 2012) to the Human genome with inclusive parameters, and removed those reads. We obtained RA from metagenomic sequencing via MetaPhlAn2 (Truong et al., 2015) with default parameters. We assigned length-normalized RA of genes, obtained by similar mapping with GEM to the reference catalog of (Li et al., 2014), to KEGG Orthology (KO) entries (Kanehisa and Goto, 2000), and these were then normalized to a sum of 1. We calculated RA of KEGG modules and pathways by summation. We considered only samples with >10K reads of 16S rRNA, and >10 M metagenomic reads (>1.5 M for daily samples in diet intervention cohort).

Associating PPGRs with Risk Factors and Microbiome Profile

We calculated the median PPGR to standardized meals for each participant who consumed at least four of the standardized meals and correlated it with clinical parameters (Pearson). We also calculated the mean PPGR of replicates of each standardized meal (if performed) and correlated (Pearson) these values with (a) blood tests; (b) anthropometric measurements; (c) 16S rRNA RA at the species to phylum levels; (d) MetaPhlAn tag-level RA; and (e) RA of KEGG genes. We capped RA at a minimum of 1e-4 (16S rRNA), 1e-5 (MetaPhlAn) and 2e-7 (KEGG gene). For 16S rRNA analysis we removed taxa present in less than 20% of participants. Correlations on RAs were performed in logspace.

Enrichment analysis of higher phylogenetic levels (d) and KEGG pathways and modules (e) was performed by Mann-Whitney *U*-test between −log(p value)*sign(R) of above correlations (d, e) of tags or genes contained in the higher order groups and −log(p value)*sign(R) of the correlations of the rest of the tags or genes.

FDR Correction

FDR was employed at the rate of 0.15, per tested variable (e.g., glucose standardized PPGR) per association test (e.g., with blood tests) for analyses in Figure 2G and Figure S4; per phylogenetic level in Figure 6 and Figure S7; and on the entire association matrix in Figure 4G.

Meal Preprocessing

We merged meals logged less than 30 min apart and removed meals logged within 90 min of other meals. We also removed very small (<15 g and <70 Calories) meals and meals with very large (>1 kg) components, meals with incomplete logging and meals consumed at the first and last 12 hr of the connection week.

PPGR Predictor

Microbiome derived features were selected according to number of estimators using them in an additional predictor run on training data. For detailed feature list see Supplemental Experimental Procedures. We predicted PPGRs using stochastic gradient boosting regression, such that 80% of the samples and 40% of the features were randomly sampled for each estimator. The depth of the tree at each estimator was not limited, but leaves were restricted to have at least 60 instances (meals). We used 4000 estimators with a learning rate of 0.002.

Microbiome Changes during Dietary Intervention

We determined the significantly changing taxa of each participant by a *Z* test of fold-change in RA between the beginning and end of each intervention week against a null hypothesis of no change and standard deviation calculated from at least 25-fold changes across the first profiling week (no intervention) of corresponding taxa from all participants with similar initial RA. We checked whether a change was consistent across the cohort for each taxa by performing Mann-Whitney *U*-test between the *Z* statistics of the "good" intervention weeks and those of the "bad" intervention weeks across all participants.

A detailed description of methods used in this paper can be found in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the data reported in this paper is ENA: PRJEB11532.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, one table and one dataset and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.11.001.

AUTHOR CONTRIBUTIONS

T.K. and D.Z. conceived the project, designed and conducted all analyses, interpreted the results, wrote the manuscript and are listed in random order. D.R. conceived and directed the dietary intervention (DI) study and designed and conducted analyses. T.K., D.Z., and A.W designed protocols and supervised data collection. N.Z., D.I., Z.H., and E.E. coordinated and supervised clinical aspects of data collection. T.K., D.Z., N.Z. and D.I. equally contributed to this work. A.W. conceived the project, developed protocols, directed and performed sample sequencing. M.R. and O.B.-Y. supervised the DI study. O.B.-Y. conducted analyses and wrote the manuscript. D.L. conducted analyses, interpreted results and advised nutritional decisions. T.A.-S. and M.L.-P. developed protocols and together with E.M. performed metagenomic extraction and sequencing. N.Z., J.S., J.A.M., G.Z.-S., L.D., and M.P.-F. developed protocols and performed 16S sequencing. G.M., N.K. and R.B. coordinated and designed data collection. Z.H. conceived the project and provided infrastructure. E.E. and E.S. conceived and directed the project and analyses, designed data collection protocols, designed and conducted the analyses, interpreted the results, and wrote the manuscript.

ACKNOWLEDGMENTS

We thank the Segal and Elinav group members for fruitful discussions; Keren Segal, Yuval Dor, Tali Raveh-Sadka, Michal Levo, and Leeat Keren for fruitful discussions and critical insights to the manuscript; Guy Raz and Ran Chen for website development; Shira Zelber-Sagi for discussions; and Noya Horowitz for writing and submitting documents for review by IRBs. This research was supported by the Weizmann Institute of Science. T.K., D.Z., and D.R. are supported by the Ministry of Science, Technology, and Space, Israel. T.K. is supported by the Foulkes Foundation. E.E. is supported by Yael and Rami Ungar, Israel; Leona M. and Harry B. Helmsley Charitable Trust; the Gurwin Family Fund for Scientific Research; Crown Endowment Fund for Immunological Research; estate of Jack Gitlitz; estate of Lydia Hershkovich; the Benoziyo Endowment Fund for the Advancement of Science; John L. and Vera Schwartz, Pacific Palisades; Alan Markovitz, Canada; Cynthia Adelson, Canada; estate of Samuel and Alwyn J. Weber; Mr. and Mrs. Donald L. Schwarz, Sherman Oaks; grants funded by the European Research Council (ERC); the Israel Science Foundation (ISF); E.E. is the incumbent of the Rina Gudinski Career Development Chair. E.S. is supported by a research grant from Jack N. Halpern, and Mr. and Mrs. Donald L. Schwarz and by grants from the ERC and the ISF.

Received: October 5, 2015
Revised: October 29, 2015
Accepted: October 30, 2015
Published: November 19, 2015

REFERENCES

American Diabetes Association (2015a). (5) Prevention or delay of type 2 diabetes. Diabetes Care *38* (*Suppl*), S31–S32.

American Diabetes Association (2015b). (4) Foundations of care: education, nutrition, physical activity, smoking cessation, psychosocial care, and immunization. Diabetes Care *38* (*Suppl*), S20–S30.

Atkinson, F.S., Foster-Powell, K., and Brand-Miller, J.C. (2008). International tables of glycemic index and glycemic load values: 2008. Diabetes Care *31*, 2281–2283.

Bailey, T.S., Ahmann, A., Brazg, R., Christiansen, M., Garg, S., Watkins, E., Welsh, J.B., and Lee, S.W. (2014). Accuracy and acceptability of the 6-day Enlite continuous subcutaneous glucose sensor. Diabetes Technol. Ther. *16*, 277–283.

Bansal, N. (2015). Prediabetes diagnosis and treatment: A review. World J. Diabetes *6*, 296–303.

Bao, J., Gilbertson, H.R., Gray, R., Munns, D., Howard, G., Petocz, P., Colagiuri, S., and Brand-Miller, J.C. (2011). Improving the estimation of mealtime insulin dose in adults with type 1 diabetes: the Normal Insulin Demand for Dose Adjustment (NIDDA) study. Diabetes Care *34*, 2146–2151.

Blaak, E.E., Antoine, J.-M., Benton, D., Björck, I., Bozzetto, L., Brouns, F., Diamant, M., Dye, L., Hulshof, T., Holst, J.J., et al. (2012). Impact of postprandial glycaemia on health and prevention of disease. Obes. Rev. *13*, 923–984.

Carpenter, D., Dhar, S., Mitchell, L.M., Fu, B., Tyson, J., Shwan, N.A.A., Yang, F., Thomas, M.G., and Armour, J.A.L. (2015). Obesity, starch digestion and amylase: association between copy number variants at human salivary (AMY1) and pancreatic (AMY2) amylase genes. Hum. Mol. Genet. *24*, 3472–3480.

Cavalot, F., Pagliarino, A., Valle, M., Di Martino, L., Bonomo, K., Massucco, P., Anfossi, G., and Trovati, M. (2011). Postprandial blood glucose predicts cardiovascular events and all-cause mortality in type 2 diabetes in a 14-year follow-up: lessons from the San Luigi Gonzaga Diabetes Study. Diabetes Care *34*, 2237–2243.

Conn, J.W., and Newburgh, L.H. (1936). The glycemic response to isogluco-genic quantities of protein and carbohydrate. J. Clin. Invest. *15*, 665–671.

Cunningham, K.M., and Read, N.W. (1989). The effect of incorporating fat into different components of a meal on gastric emptying and postprandial blood glucose and insulin responses. Br. J. Nutr. *61*, 285–290.

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. Nature *505*, 559–563.

Dodd, H., Williams, S., Brown, R., and Venn, B. (2011). Calculating meal glycemic index by using measured and published food values compared with directly measured meal glycemic index. Am. J. Clin. Nutr. *94*, 992–996.

Duncan, S.H., Belenguer, A., Holtrop, G., Johnstone, A.M., Flint, H.J., and Lobley, G.E. (2007). Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. Appl. Environ. Microbiol. *73*, 1073–1078.

Dunstan, D.W., Kingwell, B.A., Larsen, R., Healy, G.N., Cerin, E., Hamilton, M.T., Shaw, J.E., Bertovic, D.A., Zimmet, P.Z., Salmon, J., and Owen, N. (2012). Breaking up prolonged sitting reduces postprandial glucose and insulin responses. Diabetes Care *35*, 976–983.

Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods *10*, 996–998.

Everard, A., Lazarevic, V., Derrien, M., Girard, M., Muccioli, G.G., Neyrinck, A.M., Possemiers, S., Van Holle, A., François, P., de Vos, W.M., et al. (2011). Responses of gut microbiota and glucose and lipid metabolism to prebiotics in genetic obese and diet-induced leptin-resistant mice. Diabetes *60*, 2775–2786.

Everard, A., Lazarevic, V., Gaïa, N., Johansson, M., Ståhlman, M., Backhed, F., Delzenne, N.M., Schrenzel, J., François, P., and Cani, P.D. (2014). Microbiome of prebiotic-treated mice reveals novel targets involved in host response during obesity. ISME J. *8*, 2116–2130.

Foster-Powell, K., Holt, S.H.A., and Brand-Miller, J.C. (2002). International table of glycemic index and glycemic load values: 2002. Am. J. Clin. Nutr. *76*, 5–56.

Friedman, J.H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. Ann. Stat. *29*, 1189–1232.

Gallwitz, B. (2009). Implications of postprandial glucose and weight control in people with type 2 diabetes: understanding and implementing the International Diabetes Federation guidelines. Diabetes Care *32* (*Suppl 2*), S322–S325.

Gibbs, E.M., Stock, J.L., McCoid, S.C., Stukenbrok, H.A., Pessin, J.E., Stevenson, R.W., Milici, A.J., and McNeish, J.D. (1995). Glycemic improvement in diabetic db/db mice by overexpression of the human insulin-regulatable glucose transporter (GLUT4). J. Clin. Invest. *95*, 1512–1518.

Greenwood, D.C., Threapleton, D.E., Evans, C.E.L., Cleghorn, C.L., Nykjaer, C., Woodhead, C., and Burley, V.J. (2013). Glycemic index, glycemic load, carbohydrates, and type 2 diabetes: systematic review and dose-response meta-analysis of prospective studies. Diabetes Care *36*, 4166–4171.

Grundy, S.M. (2012). Pre-diabetes, metabolic syndrome, and cardiovascular risk. J. Am. Coll. Cardiol. *59*, 635–643.

Hastie, T., Tibshirani, R., and Friedman, J. (2008). The Elements of Statistical Learning: Data Mining, Inference and Prediction (Stanford: Springer).

Himsworth, H.P. (1934). Dietetic factors influencing the glucose tolerance and the activity of insulin. J. Physiol. *81*, 29–48.

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207–214.

Israeli Center for Disease Control (2014). Health 2013 (Jerusalem: Israeli Ministry of Health).

Jenkins, D.J., Wolever, T.M., Taylor, R.H., Barker, H., Fielden, H., Baldwin, J.M., Bowling, A.C., Newman, H.C., Jenkins, A.L., and Goff, D.V. (1981). Glycemic index of foods: a physiological basis for carbohydrate exchange. Am. J. Clin. Nutr. *34*, 362–366.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.

Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C.J., Fagerberg, B., Nielsen, J., and Bäckhed, F. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature *498*, 99–103.

Korem, T., Zeevi, D., Suez, J., Weinberger, A., Avnit-Sagi, T., Pompan-Lotan, M., Matot, E., Jona, G., Harmelin, A., Cohen, N., et al. (2015). Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. Science *349*, 1101–1106.

Kristo, A.S., Matthan, N.R., and Lichtenstein, A.H. (2013). Effect of diets differing in glycemic index and glycemic load on cardiovascular risk factors: review of randomized controlled-feeding trials. Nutrients *5*, 1071–1080.

Lamkin, D.M., Spitz, D.R., Shahzad, M.M.K., Zimmerman, B., Lenihan, D.J., Degeest, K., Lubaroff, D.M., Shinn, E.H., Sood, A.K., and Lutgendorf, S.K. (2009). Glucose as a prognostic factor in ovarian carcinoma. Cancer *115*, 1021–1027.

Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J.-M., Kennedy, S., et al.; MetaHIT consortium (2013). Richness of human gut microbiome correlates with metabolic markers. Nature *500*, 541–546.

Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J.R., Prifti, E., Nielsen, T., et al.; MetaHIT Consortium; MetaHIT Consortium (2014). An integrated catalog of reference genes in the human gut microbiome. Nat. Biotechnol. *32*, 834–841.

Marco-Sola, S., Sammeth, M., Guigó, R., and Ribeca, P. (2012). The GEM mapper: fast, accurate and versatile alignment by filtration. Nat. Methods *9*, 1185–1188.

Martínez, I., Lattimer, J.M., Hubach, K.L., Case, J.A., Yang, J., Weber, C.G., Louk, J.A., Rose, D.J., Kyureghian, G., Peterson, D.A., et al. (2013). Gut microbiome composition is linked to whole grain-induced immunological improvements. ISME J. *7*, 269–280.

Nathan, D.M., Davidson, M.B., DeFronzo, R.A., Heine, R.J., Henry, R.R., Pratley, R., and Zinman, B.; American Diabetes Association (2007). Impaired fasting glucose and impaired glucose tolerance: implications for care. Diabetes Care *30*, 753–759.

Nielsen, H.B., Almeida, M., Juncker, A.S., Rasmussen, S., Li, J., Sunagawa, S., Plichta, D.R., Gautier, L., Pedersen, A.G., Le Chatelier, E., et al.; MetaHIT

Consortium; MetaHIT Consortium (2014). Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat. Biotechnol. *32*, 822–828.

Nishida, T., Tsuji, S., Tsujii, M., Arimitsu, S., Haruna, Y., Imano, E., Suzuki, M., Kanda, T., Kawano, S., Hiramatsu, N., et al. (2006). Oral glucose tolerance test predicts prognosis of patients with liver cirrhosis. Am. J. Gastroenterol. *101*, 70–75.

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature *490*, 55–60.

Riccardi, G., and Rivellese, A.A. (2000). Dietary treatment of the metabolic syndrome–the optimal diet. Br. J. Nutr. *83* (*Suppl 1*), S143–S148.

Ridaura, V.K., Faith, J.J., Rey, F.E., Cheng, J., Duncan, A.E., Kau, A.L., Griffin, N.W., Lombard, V., Henrissat, B., Bain, J.R., et al. (2013). Gut microbiota from twins discordant for obesity modulate metabolism in mice. Science *341*, 1241214.

Sandkvist, M. (2001). Type II secretion and pathogenesis. Infect. Immun. *69*, 3523–3535.

Santacruz, A., Marcos, A., Wärnberg, J., Martí, A., Martin-Matillas, M., Campoy, C., Moreno, L.A., Veiga, O., Redondo-Figuero, C., Garagorri, J.M., et al.; EVASYON Study Group (2009). Interplay between weight loss and gut microbiota composition in overweight adolescents. Obesity (Silver Spring) *17*, 1906–1915.

Schwingshackl, L., and Hoffmann, G. (2013). Long-term effects of low glycemic index/load vs. high glycemic index/load diets on parameters of obesity and obesity-associated risks: a systematic review and meta-analysis. Nutr. Metab. Cardiovasc. Dis. *23*, 699–706.

Suez, J., Korem, T., Zeevi, D., Zilberman-Schapira, G., Thaiss, C.A., Maza, O., Israeli, D., Zmora, N., Gilad, S., Weinberger, A., et al. (2014). Artificial sweeteners induce glucose intolerance by altering the gut microbiota. Nature *514*, 181–186.

Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., and Segata, N. (2015). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat. Methods *12*, 902–903.

Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. Nature *444*, 1027–1031.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., et al. (2009). A core gut microbiome in obese and lean twins. Nature *457*, 480–484.

Vega-López, S., Ausman, L.M., Griffith, J.L., and Lichtenstein, A.H. (2007). Interindividual variability and intra-individual reproducibility of glycemic index values for commercial white bread. Diabetes Care *30*, 1412–1417.

Vrolix, R., and Mensink, R.P. (2010). Variability of the glycemic response to single food products in healthy subjects. Contemp. Clin. Trials *31*, 5–11.

Wolever, T.M., and Jenkins, D.J. (1986). The use of the glycemic index in predicting the blood glucose response to mixed meals. Am. J. Clin. Nutr. *43*, 167–172.

World Health Organization (2008). Global Health Observatory Data Repository (Disease and Injury Country Estimates).

Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S.A., Bewtra, M., Knights, D., Walters, W.A., Knight, R., et al. (2011). Linking long-term dietary patterns with gut microbial enterotypes. Science *334*, 105–108.

Xiao, S., Fei, N., Pang, X., Shen, J., Wang, L., Zhang, B., Zhang, M., Zhang, X., Zhang, C., Li, M., et al. (2014). A gut microbiota-targeted dietary intervention for amelioration of chronic inflammation underlying metabolic syndrome. FEMS Microbiol. Ecol. *87*, 357–367.

Zhang, X., Shen, D., Fang, Z., Jie, Z., Qiu, X., Zhang, C., Chen, Y., and Ji, L. (2013). Human gut microbiota changes reveal the progression of glucose intolerance. PLoS ONE *8*, e71108.

Cell

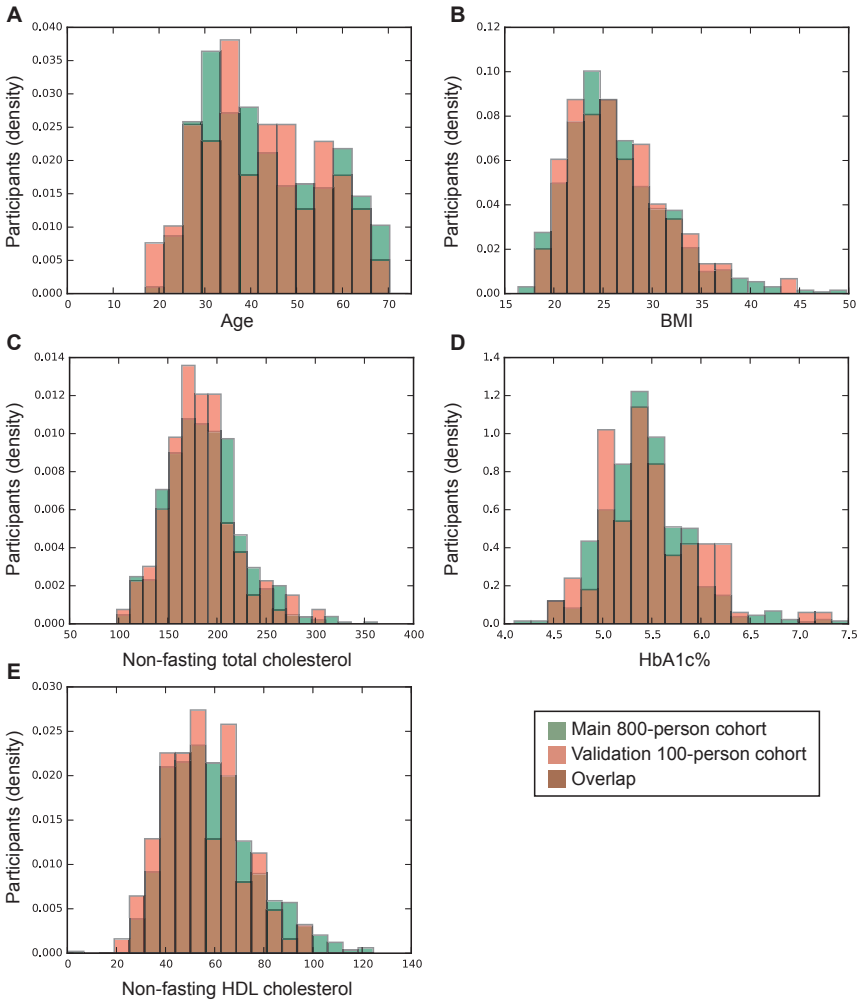


Figure S1. Histograms Showing Distribution of Major Risk Factors in the Main 800-Person Cohort and the Validation 100-Person Cohort, Related to Table 1 and Figure 3

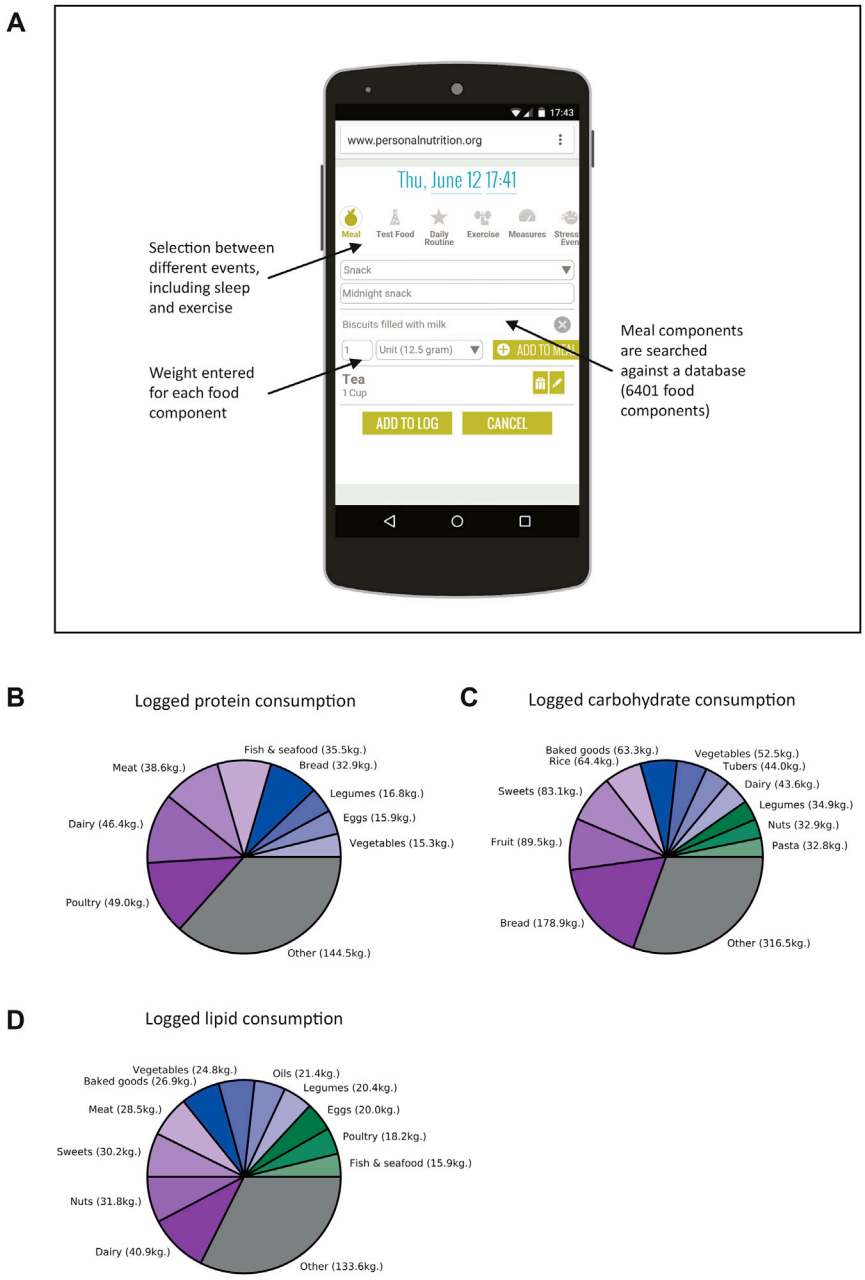


Figure S2. Logging of Lifestyle and Nutritional Information, Related to Figure 1

(A) Key features of the smartphone-adjusted website that we developed for meal logging and daily activity collection. (B–D) Major food components consumed by the cohort by protein (B), carbohydrates (C), and fat (D).

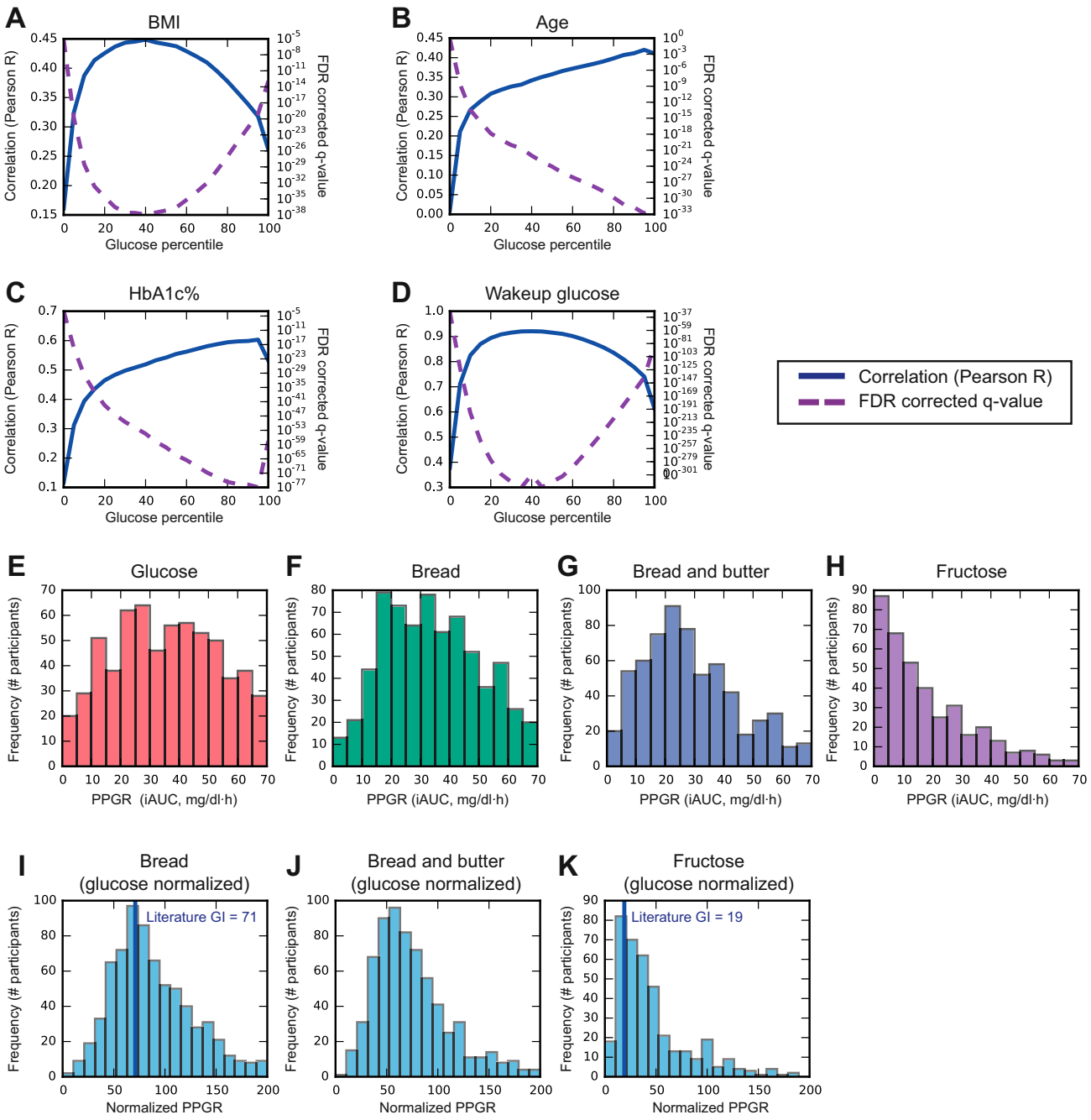


Figure S3. Glycemic Data Are Interpersonally Variable and Correlated with Risk Factors, Related to Figure 2

(A–D) For each participant, the CGM measured glucose level at all percentiles (0th, 1st, 2nd, ..., 100th percentiles) across the entire connection week (x axis) were correlated with BMI (A), age (B), HbA1c% (C) and wakeup glucose levels (D). For each such percentile, shown is the Pearson R (left y axis, full line) and FDR-corrected q-value (right y axis, dashed line). (E–H) Histograms as in Figure 2B, without smoothing, of the postprandial glycemic response (PPGR) to glucose (E), bread (F), bread and butter (G) and fructose (H), the standardized meals provided to participants (each with 50 g of available carbohydrates). (I–K) Histograms as in Figure 2B, demonstrating similarly high variability in the PPGR of each participant to bread (I), bread & butter (J) and fructose (K), except here PPGRs were normalized to the PPGR of the participant to 50 g. glucose. Purple vertical line denotes published glycemic index (GI) of bread and fructose (Foster-Powell et al., 2002).

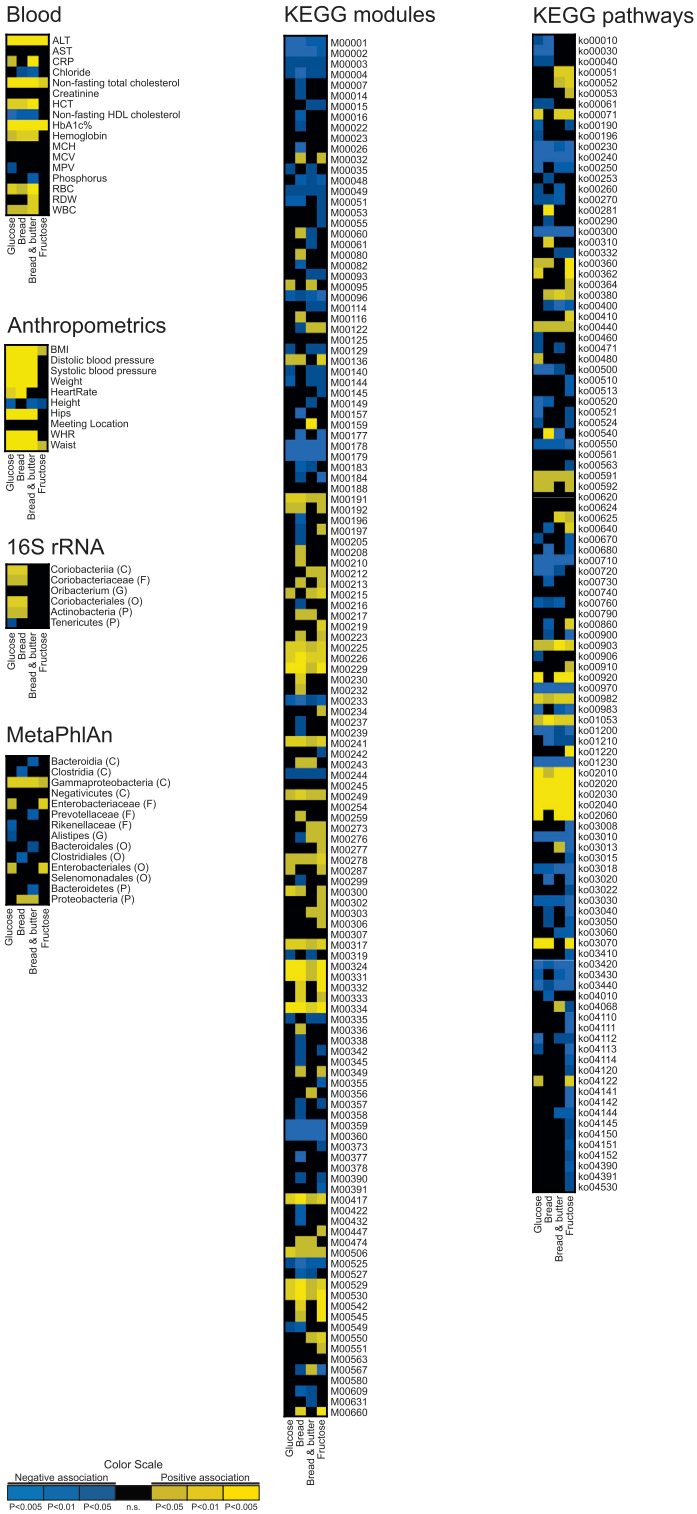


Figure S4. Interpersonal Differences in Postprandial Glycemic Responses (PPGRs) to Standardized Meals Associate with Clinical and Microbiome Features, Related to Figure 2

Complete set of statistically significant associations from Figure 2H. Shown are all statistically significant associations ($p < 0.05$, FDR corrected) found between participants' standardized meals PPGRs and participants' blood parameters, anthropometrics, 16S rRNA and metagenomic derived abundances, and abundances of KEGG pathways and modules. Yellow and blue entries indicate significant positive and negative associations, respectively.

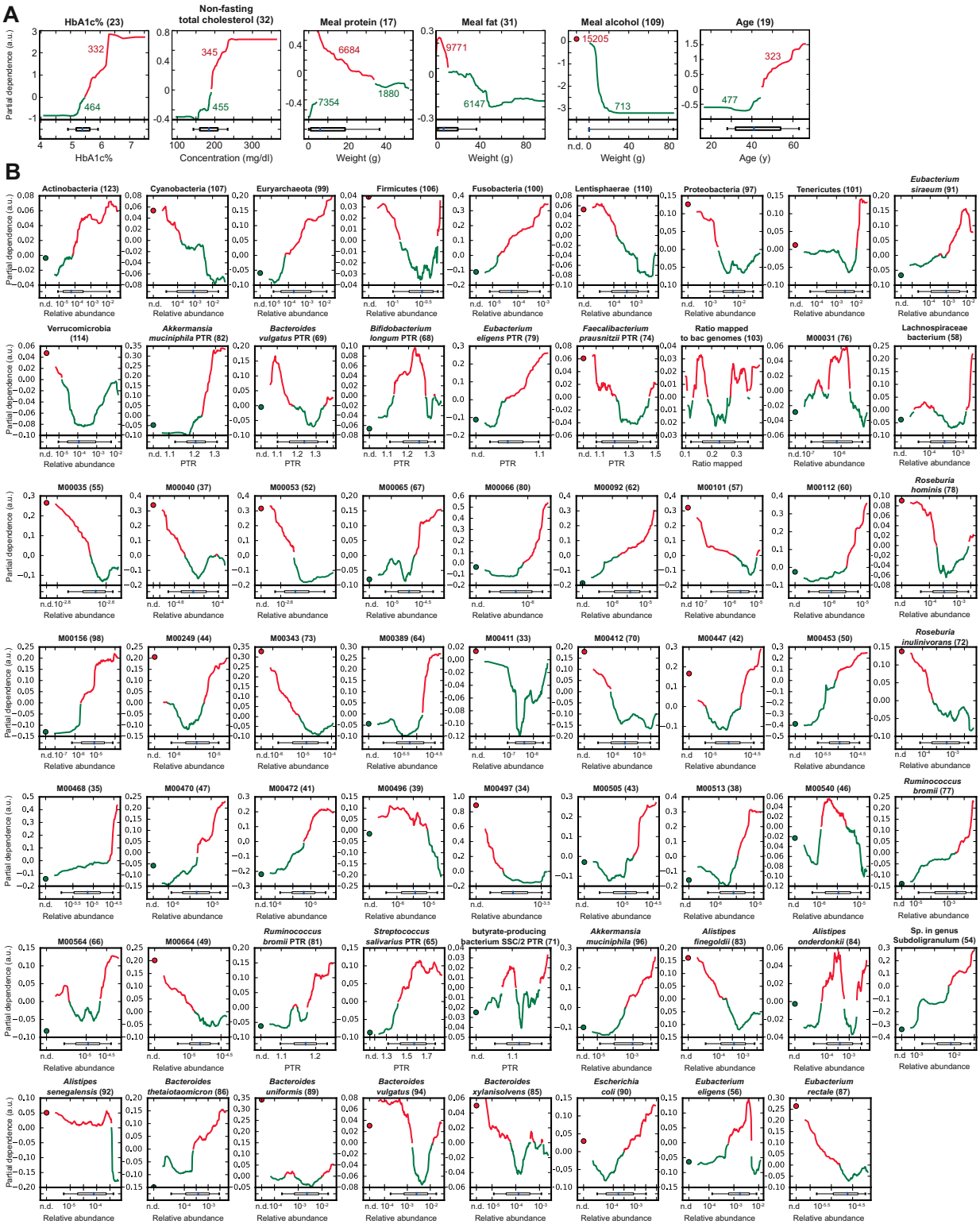


Figure S5. Additional Features Underlying the Prediction of Postprandial Glycemic Responses, Related to Figure 4

(A) Shown are partial dependency plots (PDPs, as in Figure 4), for additional features. (B) Shown are PDPs for all microbiome-derived features.

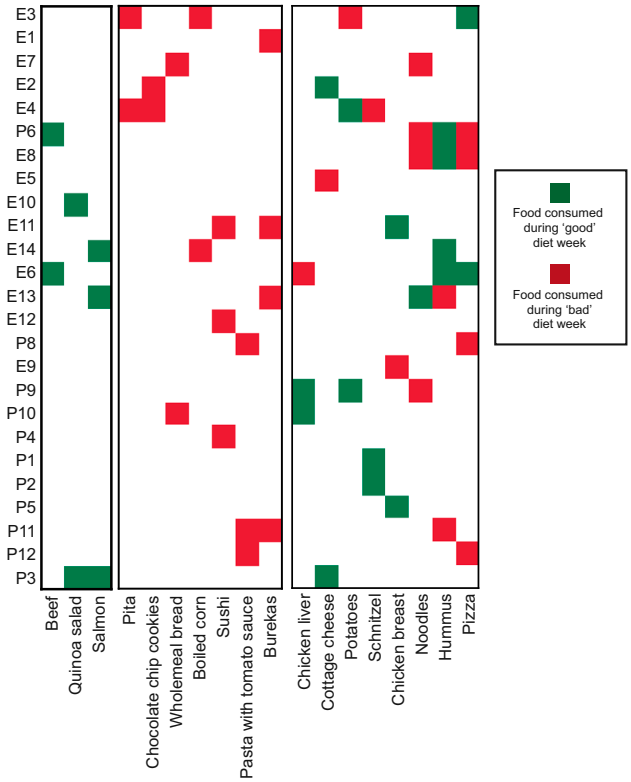
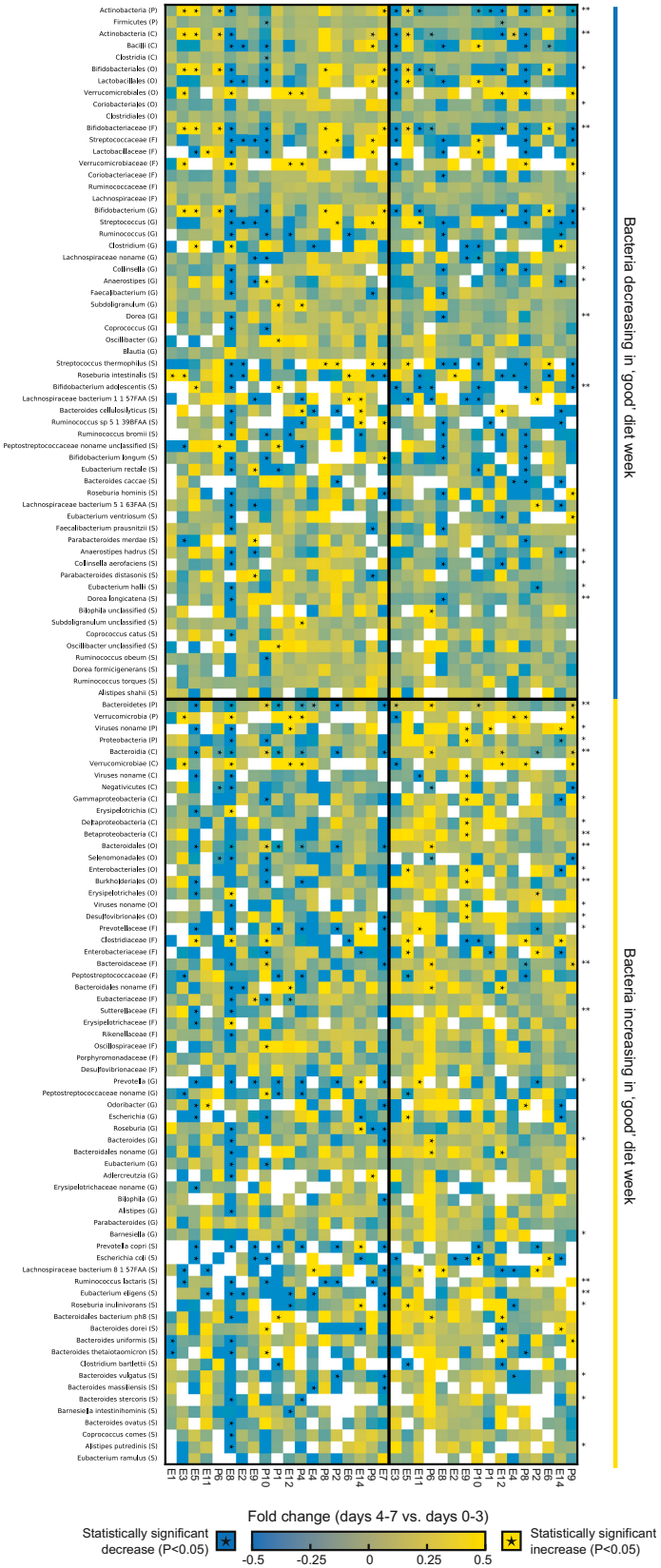


Figure S6. Partition of Dominant Food Components into "Good" and "Bad" Diets of Participants across the Diet Intervention Weeks, Related to Figure 5

Colored entries indicate a food component (x axis) that was prescribed in the "good" (green) or "bad" (red) diet week of a participant (y axis). Food components are shown in submatrices of foods that were only prescribed in the "good" diet week (left column block), food components prescribed only in the "bad" diet week (middle column block), and foods that were prescribed in the "good" diet of some participants and in the "bad" diet of other participants (right column block). Only dominant food component are shown, defined as foods whose carbohydrate content was more than 50% of the entire Caloric content of the meal in which they were prescribed.



(legend on next page)

Figure S7. Dietary Interventions Induce Significant Alterations to the Gut Microbiota Configuration, Related to Figure 6

Shown are all bacteria present in at least one stool sample in at least 10 different intervention weeks. Entries are colored according to the (log) fold change between the relative abundance of a bacteria (y axis) in days 4–7 and the relative abundance of that bacteria at days 0–3 within each participant (x axis). Asterisks indicate a statistically significant fold change. Left and right columns blocks correspond to changes in "bad" and "good" intervention weeks, respectively. Top and bottom row blocks correspond to a trend of decreasing and increasing abundance in the "good" intervention week, respectively, and conversely in the "bad" week. Taxa with statistically significant (Mann-Whitney *U*-test) patterns of consistent change following the "good" diet that is opposite to a consistent change following the "bad" diet are marked to the right of the heatmap (** $p < 0.01$, * $p < 0.05$, FDR corrected). P, phylum; C, class; O, order; F, family; G, genus; S, species.

Cell

Supplemental Information

Personalized Nutrition by Prediction of Glycemic Responses

David Zeevi, Tal Korem, Daphna Rothschild, Adina Weinberger, Niv Zmora, David Israeli, Orly Ben-Yacov, Dar Lador, Tali Avnit-Sagi, Maya Lotan-Pompan, Jotham Suez, Jemal Ali Mahdi, Elad Matot, Gal Malka, Noa Kosower, Michal Rein, Gili Zilberman-Schapira, Lenka Dohnalová, Meirav Pevsner-Fischer, Rony Bikovsky, Zamir Halpern, Eran Elinav, and Eran Segal

Supplemental Experimental Procedures

Study enrollment. Participants of all three cohorts registered via an online website (www.personalnutrition.org). Prior to participation, subjects attended a two-hour meeting ("connection meeting") in groups of up to 12 people, in which they received detailed instructions regarding participation in the study, including personal instructions on usage of a blood glucose monitor (BGM) and usage of our logging website (see below). The meeting also included a personal meeting with a physician which authorized participation and acquired informed consent.

Inclusion and exclusion criteria. Participants included in the study were males and females aged 18-70 able to provide an informed consent and technically operate a glucometer (for calibration of the CGM) throughout the connection week. Exclusion criteria included: (i) pregnancy; (ii) usage of antibiotics within three months prior to participation; (iii) chronically active inflammatory or neoplastic disease in the three years prior to enrollment; (iv) chronic gastrointestinal disorder, including Inflammatory Bowel Disease and Celiac disease; (v) skin disease, including contact dermatitis, precluding proper attachment of the continuous glucose monitor; (vi) active neuropsychiatric disorder; (vii) myocardial infarction or cerebrovascular accident in the 6 months prior to participation; (viii) chronic immunosuppressive medication usage; (ix) pre-diagnosed type I or type II diabetes mellitus (the latter not applicable for the dietary intervention cohort).

Lifestyle and medical background, and food frequency questionnaires. All questionnaires were distributed to participants online prior to the connection week and were manually validated by trained clinical research associates (CRAs). Ambiguities and unclarities in the questionnaires were personally discussed with each participant. Food frequency questionnaires were aimed at characterizing long-term dietary habits, and were created by certified clinical dietitians according to commonly used practices. All questionnaires are available as Data S1.

Blood tests. Blood was drawn by a certified nurse at the beginning of the connection week. Fasting was not required prior to the test. The following tests were performed: complete blood count (CBC), glycosylated hemoglobin (HbA1c%), liver aminotransferases activity (ALT, AST), total and HDL cholesterol levels, thyroid stimulating hormone (TSH), blood electrolytes (sodium, potassium, calcium, phosphorus and chloride), creatinine levels, and C-reactive protein (CRP).

Anthropometrics. Measurements were performed by trained CRAs and included weight, height, and waist and hip circumference. In addition, blood pressure and heart-rate were measured by a certified nurse.

Continuous glucose monitoring. Glucose was measured using the iPro2™ continuous glucose monitor (CGM; Medtronic, MN, USA), which measures interstitial glucose levels using the subcutaneous Enlite™ sensors. To calibrate CGM measurements to blood glucose measurements, participants also measured blood glucose by fingerpricks, using the Contour™ BGM (Bayer AG, Leverkusen, Germany). Participants were required to perform four measurements per day, and were given reminders sent to their cellular phones using an automated messaging system. Participants were asked to measure blood glucose prior to meals for added accuracy, as recommended by the iPro2 manual. CareLink online software (Medtronic) was used to perform calibration for CGM measurements, as directed by the iPro2 manual.

Activity logging. Participants were instructed to record their daily activities during the connection week using a proprietary smartphone-adjusted website that we developed for this purpose. Participants were asked to record, for each meal, exact components and component weights. In addition, participants were asked to record the consumption of standardized meals, sleep and wakeup times, physical activity, including duration and intensity, stressful events upon occurrence, hunger levels, dietary supplements and prescription drug consumption. A special section of the connection meeting was dedicated to correct logging, with special emphasis on logging meals at the correct time and correctly recording food components. Logging of each participant was reviewed by trained CRAs in real-time during the connection week and once more when the week ended. Unclarities in the log were personally resolved with each participant.

Standardized meals. We supplied our participants with standardized meals at the start of each connection week. Meals were prepared the same way for all participants. All standardized meals were calculated to have 50g of available carbohydrates. Standardized meals were: (a) 50g glucose (two replicates; 1564 meals, 794 participants); (b) 110g of white bread (two replicates; 1569 meals, 795 participants); (c) 110g of white bread and 30g of butter (two replicates, dairy-consuming participants only; 1399 meals, 709 participants); (d) 55g of white bread and 50g of dark chocolate (two replicates, non dairy-consuming participants; 137 meals, 69 participants); (e) 50g fructose (one replicate, only participants connected to CGM in the evening; 438 meals, 438 participants).

Participants were instructed to (a) consume the standardized meal as their first meal in the morning, immediately after their night fast; (b) to eat the entire meal provided without any modifications, and without eating anything else for two hours; (c) to refrain from performing strenuous physical activity before, and for two hours following standardized meal consumption. Adherence with these instructions was validated with each participant by the study CRAs.

Computing postprandial glycemic responses. Logged meal times and continuous glucose measurements (CGMs) were used to calculate the incremental area under the curve (iAUC) following a meal as previously described (Wolever and Jenkins, 1986). To reduce noise, the median of all glucose values from the 30-minute period prior to the meal was taken as the initial glucose level, above which the incremental area was calculated. We filtered out meals that had incomplete glucose measurements in the time window of 30 minutes before and 2 hours after the logged meal time.

Sequencing depth. Metagenomic samples were sequenced to a mean depth of 17,129,579±258,720, 19,840,830±895,788, 4,341,958±298,411 (mean±sem) for the main, validation and dietary intervention cohorts, respectively. 16S samples were sequenced to a mean depth of 135,324±5,452 and 101,931±17,005 (mean±sem) for the main and validation cohorts respectively.

Quality control of metagenomic reads and removal of host DNA. We applied Trimmomatic (Bolger et al., 2014) with the following parameters:

ILLUMINACLIP:<Trueseq3 adapters fasta file>:2:30:10 LEADING:25 TRAILING:25 MINLEN:50

For files originating from HiSeq and MiSeq, we added SLIDINGWINDOW:3:25 to the command. We removed host DNA by mapping to the human genome (hg19, downloaded from https://genome.ucsc.edu) and removing any mapped reads (see section below).

Mapping of metagenomic sequencing reads. Mapping was performed using the GEM mapper (Marco-Sola et al., 2012) with the following parameters:

-q offset-33 --gem-quality-threshold 26 -m 0.1 -e 0.1 --min-matched-bases 0.8 --max-big-indel-length 15 -s 3 -d 'all' -D 1 -v -T 2

with the addition of modifier:

-p -E 0.3 --max-extendable-matches 'all' --max-extensions-per-match 5

for paired end reads, and -m set to 0.05 for mapping to the human genome. Resulted mappings were retained as long as they had at least 50 matched bases with minimal quality of 26.

Bacterial relative abundance calculation. We calculated relative abundances from metagenomic reads by using MetaPhlAn2 (Truong et al., 2015) with default parameters, and from 16S rRNA sequencing reads by using USearch8.0 (Edgar, 2013) as described in the manual at http://drive5.com/usearch/manual/uparse_pipeline.html, using the commands fastq_mergepairs, derep_fulllength, sortbysize, otu_radius_pct, uchime_ref, usearch_global, followed by the script assign_taxonomy.py from QIIME (Caporaso et al., 2010).

PCoA in Fig. 1G. Oral samples are samples from buccal mucosa and tongue dorsum. Urogenital samples are samples from mid-vagina, posterior fornix and vaginal introitus.

Genetic content relative abundance calculation. We mapped reads to the integrated reference catalog of the human gut microbiome (Li et al., 2014). For each gene in the catalog, we counted the fraction of reads mapped to it from each sample, normalized by gene length in kilobases. Reads mapping to more than one location were split so that each location received an equal fraction of the mapped read. We subsequently assigned mapped reads to KEGG Orthology (KO) entries (Kanehisa and Goto, 2000) using the gene annotation table available at http://meta.genomics.cn/. We then calculated gene relative abundances by normalizing the KEGG genes of each sample to sum to 1. To calculate the abundances of KEGG pathways and modules, we summed the relative abundance of genes in each pathway and module.

Enrichment analysis of higher-order microbiome taxa and function. To test for associations at the level of KEGG pathways and modules and for associations with MetaPhlAn higher taxonomical levels (i.e., entities that each contain multiple individual features - multiple genes per KEGG pathway and module and multiple species per higher taxonomical level), we ordered gene and tag level associations by the sign of the correlation (i.e., positive or negative) multiplied by the negative of the logged p-values, such that the most significant positive correlations would have a higher rank, and the most significant negative correlations would have a lower rank. For each group (pathway, module and MetaPhlAn higher taxonomic level), we computed the P-value and direction of association by performing the Mann-Whitney *U*-test between genes or MetaPhlAn tags belonging to that group and those not belonging to that group.

Meal preprocessing. We employed the following filters consecutively to the meals that served as input to the predictor: (i) We merged meals logged less than 30 minutes apart, as long as the earliest meal contained

more than 50 Calories. The merged meal was assigned the summed values of all of its components and the time of the earliest component. (ii) We removed meals logged within 90 minutes from other meals containing more than 50 Calories. (iii) We removed meals with components weighing more than 1kg unless they contained less than 20 Calories. (iv) We removed meals with incomplete logging. (v) We removed meals with less than 70 Calories and less than 15 grams of total weight. (vi) We removed meals consumed in the first and last 12 hours of the CGM connection period. This entire process resulted in 15,918 and 2,102 meals in the main and validation cohorts, respectively (median of 20 and 21 meals per participant, respectively).

Predictor implementation. Our predictor is based on code adapted from the sklearn 0.15.2 (Pedregosa et al., 2011) GradientBoostingRegressor class

Predictor features. The following features were incorporated in our predictor: (i) Meal features - amount of alcohol (g), caffeine (mg), carbohydrate (g), dietary fibers (g), energy (Cal.), fat (g), protein (g), sodium (mg), sugars (g), water (g), as well as the carbohydrates-to-fat ratio. (ii) Log-based features - time to next and last exercise and sleep; amount of water consumed one hour before and in the two hours following the meal; total amount of carbohydrates consumed in the 3, 6 and 12 hours prior to the meal; total amount of Calories consumed in the 2, 3, 6, and 12 hours prior to the meal; total amount of fibers consumed 12 and 24 hours prior to the meal; and the hour of the day in which the meal was consumed. (iv) CGM-derived features - iAUC and glucose trend of 1, 2, and 4 hours prior to the meal. (v) Clinical features - blood test results. (vi) Personal features - age, sex, smoking habits, and self reported hunger, physical activity, stress levels and defecation routine. (vii) Microbiome features - relative abundances of 16S rRNA based phyla existing in more than 20% of the cross-validation training cohort; relative abundance of the 30 KEGG modules used by most estimators when running the same predictor with all KEGG modules using a 10-fold cross-validation scheme on the main cohort; 20 metagenome-based species relative abundances selected similarly to the KEGG modules; 10 PTRs calculated as in (Korem et al., 2015), selected similarly to the KEGG modules; Percentage of reads mapped to host genome, gene-set database, and database of full genomes described in (Korem et al., 2015). In all these features relative abundances were corrected to mapping percentage and missing values were capped to a minimum of 1e-9 (0.5 for PTRs).

Glycemic changes during dietary intervention. To compare the glycemic consequences of the two dietary intervention weeks (‘good’ and ‘bad’ weeks): (1) For each participant, we compared the PPGRs of all major meals (breakfast, lunch, dinner) in the ‘good’ week to all PPGRs in the ‘bad’ week using the Mann-Whitney *U*-test (**Fig. 5C**). (2) For each arm separately, we compared the average PPGR to a major meal type in the ‘good’ week and the corresponding type in the ‘bad’ week (e.g., average breakfast PPGR in the ‘good’ diet vs. the ‘bad’ diet). We used Wilcoxon signed-rank test, pairing between each meal type of each participant (**Fig. 5D**). (3) We compared the glucose level fluctuations (**Fig. 5E**) and max PPGR (**Fig. 5F**) across the entire ‘good’ diet week vs. the entire ‘bad’ diet week, again, pairing between different participants and using the Wilcoxon signed-rank test. Major meals were defined as >300 Calories.

Microbiome changes during dietary intervention. To identify significant changes in relative bacterial abundances during the dietary intervention weeks (‘good’ and ‘bad’ diet weeks) we compared changes that occurred in these weeks to changes that occurred in the first profiling week in which no intervention was performed. To this end, we defined the change in a specific taxon in a given intervention week as the fold change between the mean of its log relative abundances in days 4.5-8.5 and days -1-3.5 relative to the time of connection to the CGM. Next, for each phylogenetic level, we examined the relative abundances of all taxa belonging to that level from the first profiling week of all participants. As the standard deviation of the fold-changes may depend on the relative abundance itself, we further divided those taxa into equally sized bins based on the mean log relative abundance at the start of the profiling weeks, and calculated the standard deviation of the fold change in each bin. The number of bins was selected such that at least 25 fold-changes would be present in each bin (10 bins were used for the species level, 5 bins for the genus level, 3 bins for the phylum level, and 4 bins for all other levels). We determined whether a specific change in the relative abundance of each taxa in a certain intervention week was significant by testing it against a null hypothesis of no change with standard deviation taken from those calculated in the profiling week as described for corresponding phylogenetic level and starting relative abundance, using a *Z*-test. We removed from further analysis taxa for which a change could not be detected in at least 10 different intervention weeks. Next, we checked whether a change was consistent across the cohort for each taxa by performing Mann-Whitney *U*-test between the *Z* statistics of the ‘good’ intervention weeks and those of the ‘bad’ intervention weeks across all participants. The test was only performed when there were at least 6 statistics available for ‘good’ weeks and for ‘bad’ weeks.

MEDIAN RESPONSE RANK	FOOD	MEAN GI	GI SEM	GI DERIVED FOOD DESCRIPTION	GI SOURCE
LOW - 1	Cookies	42	5	LU P'tit Déjeuner Chocolat (LU, France)	1
2	Chocolate	49	6	Cadbury's Confectionery, Ringwood, Australia	2
3	Ice cream	36	8	Ice cream, NS (Canada)	1
4	Wafers	74		Graham Wafers (Christie Brown and Co, Toronto, Canada)	1
5	Banana	46		Banana (Canada)	1
6	Pita	57		Pita bread, white (Canada)	1
7	Coke	63		Coca Cola®, soft drink (Coca Cola Bottling Co., Atlanta, GA, USA)	2
8	Grapes	43		Grapes, NS (Canada)	1
9	Orange	31		Oranges, raw	1
10	Dried dates	42	4	Dates, dried (Australia)	2
11	Granola	57	2	Museli	1
12	Apple	34		Apple, NS (Canada)	1
13	Watermelon	72	13	Watermelon, raw (Australia)	2
14	Cheesecake				
15	Croissant	67		Croissant (Food City, Toronto, Canada)	1
16	Orange juice	50	2	Orange Juice, mean of 4 studies	2
17	Pizza	60		Pizza, cheese (Pillsbury Canada Ltd, Toronto, Canada)	1
18	Danish				
19	Bread	70		White wheat bread*	2
20	Potatoes	69	5	Potato, white with skin, baked (UK)	2
21	Greek pastery				
22	Wholemeal bread				
23	Persimmon				
24	Rice crackers	91		Rice cracker, plain (Sakada, Japan)	1
25	Roll				
26	Rice	86	10	Broken rice, white, cooked in rice cooker (Lion Foods, Thailand)	1
27	Wholemeal roll				
HIGH - 28	Cereal	81	6	Corn flakes	2

Table S1. Glycemic indices of food items. Related to Figure 2. Shown are food items presented in Fig. 2E. 1 - (Foster-Powell et al., 2002); 2 - (Atkinson et al., 2008).

Supplemental references

Atkinson, F.S., Foster-Powell, K., and Brand-Miller, J.C. (2008). International tables of glycemic index and glycemic load values: 2008. Diabetes Care *31*, 2281–2283.

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

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods *10*, 996–998.

Foster-Powell, K., Holt, S.H.A., and Brand-Miller, J.C. (2002). International table of glycemic index and glycemic load values: 2002. Am. J. Clin. Nutr. *76*, 5–56.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.

Korem, T., Zeevi, D., Suez, J., Weinberger, A., Avnit-Sagi, T., Pompan-Lotan, M., Matot, E., Jona, G., Harmelin, A., Cohen, N., et al. (2015). Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. Science *349*, 1101–1106.

Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J.R., Prifti, E., Nielsen, T., et al. (2014). An integrated catalog of reference genes in the human gut microbiome. Nat. Biotechnol. *32*, 834–841.

Marco-Sola, S., Sammeth, M., Guigó, R., and Ribeca, P. (2012). The GEM mapper: fast, accurate and versatile alignment by filtration. Nat. Methods *9*, 1185–1188.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine Learning in {P}ython. J. Mach. Learn. Res. *12*, 2825–2830.

Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., and Segata, N. (2015). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat. Methods *12*, 902–903.

Wolever, T.M., and Jenkins, D.J. (1986). The use of the glycemic index in predicting the blood glucose response to mixed meals. Am. J. Clin. Nutr. *43*, 167–172.