Clinical and Translational Report

Cell Metabolism

Bread Affects Clinical Parameters and Induces Gut Microbiome-Associated Personal Glycemic Responses

Graphical Abstract

Highlights

- Crossover trial shows no differential clinical effect of white versus sourdough bread
- The microbiome composition was generally resilient to dietary intervention of bread
- The glycemic response to the two types of bread varies greatly across people
- Microbiome-based classifier accurately predicts glycemic-response-inducing bread type

Authors

Tal Korem, David Zeevi, Niv Zmora, ..., Avraham A. Levy, Eran Elinav, Eran Segal

Correspondence

avi.levy@weizmann.ac.il (A.A.L.), eran.elinav@weizmann.ac.il (E.E.), eran.segal@weizmann.ac.il (E.S.)

In Brief

Korem et al. performed a crossover trial of industrial white or artisanal sourdough bread consumption and found no significant difference in clinical effects, with the gut microbiome composition remaining generally stable. They showed the glycemic response to bread type to be person specific and microbiome associated, highlighting the importance of nutrition personalization.
Bread Affects Clinical Parameters and Induces Gut Microbiome-Associated Personal Glycemic Responses

Tal Korem,1,2,7 David Zeevi,1,2,7 Niv Zmora,3,4,5 Omer Weissbrod,1,2 Noam Bar,1,2 Maya Lotan-Pompan,1,2 Tali Avnit-Sagi,1,2 Noa Kosower,1,2 Gal Malka,1,2 Michal Rein,1,2 Jotham Suez,3 Ben Z. Goldberg,6 Adina Weinberger,1,2 Avraham A. Levy,6,* Eran Elinav,3,* and Eran Segal1,2,8,*

1Department of Computer Science and Applied Mathematics
2Department of Molecular Cell Biology
3Department of Immunology
4Internal Medicine Department, Tel Aviv Sourasky Medical Center, Tel Aviv 6423906, Israel
5Research Center for Digestive Tract and Liver Diseases, Tel Aviv Sourasky Medical Center, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 6423906, Israel
6Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 7610001, Israel
7These authors contributed equally
8Lead contact
*Correspondence: avi.levy@weizmann.ac.il (A.A.L.), eran.elinav@weizmann.ac.il (E.E.), eran.segal@weizmann.ac.il (E.S.)
http://dx.doi.org/10.1016/j.cmet.2017.05.002

SUMMARY

Bread is consumed daily by billions of people, yet evidence regarding its clinical effects is contradicting. Here, we performed a randomized crossover trial of two 1-week-long dietary interventions comprising consumption of either traditionally made sourdough-leavened whole-grain bread or industrially made white bread. We found no significant differential effects of bread type on multiple clinical parameters. The gut microbiota composition remained person specific throughout this trial and was generally resilient to the intervention. We demonstrate statistically significant interpersonal variability in the glycemic response to different bread types, suggesting that the lack of phenotypic difference between the bread types stems from a person-specific effect. We further show that the type of bread that induces the lower glycemic response in each person can be predicted based solely on microbiome data prior to the intervention. Together, we present marked personalization in both bread metabolism and the gut microbiome, suggesting that understanding dietary effects requires integration of person-specific factors.

INTRODUCTION

Bread is a key ingredient of the human diet, is consumed by billions of people worldwide (FAO, 1998), and makes up roughly 10% of the adult caloric intake (Zeevi et al., 2015). Its relative ease of preparation from dry ingredients contributes to the centrality of different forms of bread in many cultural cuisines.

Wheat is the most commonly used cereal for baking bread. It has been cultivated roughly 10,000 years ago as part of the Neolithic agricultural revolution (Salamini et al., 2002) and nowadays is the most commonly grown grain worldwide, with approximately 750 million metric tons produced annually (USDA, 2016). The high penetrance of wheat bread into most human cultures and its long-lasting consumption suggest that humans and their gut microbiome have adapted to metabolizing wheat bread. However, commonly consumed present-day bread greatly differs from traditional wheat bread in many ways, including the degree of refinement, type of leavening agent, preparation process, and additives used.

Whole grains contain germ and bran, which contain dietary fiber, B vitamins, iron, magnesium, and zinc (Slavin et al., 2001) and are therefore considered “healthy,” but they are removed in the milling process of refined flour. Scientific evidence regarding the health benefits of whole-grain consumption are contradictory. Several studies regard consumption of whole grains as advantageous to human health and suggest that it is associated with a significantly lower all-cause mortality (Aune et al., 2016); a reduced risk of cancer (Jacobs et al., 1998), cardiovascular disease (Mellen et al., 2008; Mente et al., 2009), type II diabetes mellitus (T2DM; de Munter et al., 2007), and the metabolic syndrome (Lutsey et al., 2007); and an improvement in glycemic control (Pereira et al., 2002), cholesterol levels (Giacco et al., 2010; Jensen et al., 2006; Sofi et al., 2010), blood pressure (Tighe et al., 2010), inflammation (Katcher et al., 2008; Montonen et al., 2013; Vanegas et al., 2017; Vitaglione et al., 2015), and liver functions (Montonen et al., 2013). Conversely, some of the above studies show improvement in only a handful of clinical markers (Giacco et al., 2010; Jensen et al., 2006; Sofi et al., 2010; Tighe et al., 2010; Vitaglione et al., 2015), and additional large-scale intervention trials show no significant effect on these disease-risk markers (Andresson et al., 2007; Brownlee et al., 2010; Costabile et al., 2008; Giacco et al., 2010; Tucker et al., 2010). Yet additional studies claim that whole-grain bread consumption may lead to a decrease in mineral absorption (Baek Kristensen et al., 2005; Reinhold et al., 1976).

The use of baker’s yeast (Saccharomyces cerevisiae) as a leavening agent is a recent widely used addition to bread...
making, dating back only ~150 years (Batt and Tortorello, 2014). In contrast, sourdough, which contains a culture of mostly wild yeast as well as lactic and acetic acid bacteria that naturally inoculate bread dough, has been used as a leavening agent since ancient times (Minervini et al., 2014). Sourdough fermentation releases several compounds that are not found in modern yeast fermentation, and sourdough breads were shown to increase mineral bioavailability (Arendt et al., 2007; Leenhardt et al., 2005; Lopez et al., 2003) and to induce advantageous effects on glucose metabolism (Lappi et al., 2010). Lastly, a very recent addition to bread making is the introduction of additives such as food preservatives and emulsifiers, whose roles are to improve the shelf-life and texture of commercial bread. Some emulsifiers were suggested to alter the gut microbiome in mice in a manner that induces inflammation and obesity (Chassaing et al., 2015).

The contradictory evidence regarding the effects of whole grains in bread on human health and the relatively poor evidence regarding the effects of sourdough fermentation may be a result of a large variability in even seemingly identical bread types, stemming from differences in bread ingredients, additives, the proportions in which they are used, and preparation methods. Moreover, sourdough fermentation depends on environmental bacterial and fungal strains that may vary compositionally and functionally across different bakeries. One way in which bread, a carbohydrate-rich food, exerts its metabolic effects is through the postprandial glycemic response (PPGR) induced by its consumption. PPGRs are an important aspect of human metabolic health. Postprandial hyperglycemia is an independent risk factor for cardiovascular disease (Gallwitz, 2009) and liver cirrhosis (Nishida et al., 2006) and is associated with markers of glycemic control (Zeevi et al., 2015), obesity (Blaa, et al., 2012; Zeevi et al., 2015), and enhanced all-cause mortality in both T2DM (Cavalot et al., 2011) and cancer (Lamkin et al., 2009).

The high interpersonal variability in the PPGR to identical foods, recently described by us and by others (Matthan et al., 2016; Zeevi et al., 2015), may constitute another previously disregarded reason to the different conclusions reached by different studies as to the effects of bread consumption on human physiology. A significant factor associated with this variability in PPGRs is the host gut microbiome (Zeevi et al., 2015), highlighted in the past decade to critically contribute to human health and to the risk of obesity, glucose intolerance, T2DM, hyperlipidemia, insulin resistance, and their metabolic complications (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).

In this study, we performed a randomized crossover trial with 20 healthy subjects comparing the effects of traditionally milled and prepared whole-grain sourdough bread and industrial white bread made from refined wheat on multiple clinical and disease markers and on the composition and function of the gut microbiome.

**RESULTS & DISCUSSION**

### No Difference in Treatment Effects between White and Sourdough Bread

To examine the difference in the effect of bread type on multiple clinical parameters and on the function and composition of the gut microbiome, we performed a randomized crossover trial with two 1-week-long periods of dietary intervention, separated by a 2-week-long washout period. Following a 3-day run-in period in which their dietary intake was characterized using real-time logging to a smartphone application that we developed (Zeevi et al., 2015), subjects received either industrial white bread made from mostly refined wheat flour (locally known as “achid”) from a major brand common in Israel (Table S1; hereinafter “white bread”), or a sourdough-leavened bread made from whole-grain wheat flour with traditional methods in an artisanal bakery (STAR Methods; Table S1; hereinafter “sourdough bread”). Subjects were randomly assigned to two sequence groups: white bread followed by sourdough bread (WS), or sourdough bread followed by white bread (SW). See Table 1 for group characteristics and Figure S1 for a CONSORT flow diagram.

As the primary outcomes of this trial were measures of glycemic control, consumption of different bread types in this trial was matched based on the amount of available carbohydrates, a main effector of the PPGR. Subjects were instructed to consume 50 g of available carbohydrates from bread every morning of the week (145 g and 110 g of sourdough and white bread, respectively, given in the form of standardized meals, STAR Methods) and to supplement the rest of their regular diet with additional bread of the same type. On the first intervention period, additional bread was consumed ad libitum; while on the second intervention period, available carbohydrates of supplemented bread were matched to the first intervention period (Figure 1A). Subjects were also instructed not to consume additional wheat products during the intervention periods, including bread that was not supplied to them, pasta, and any other wheat-based products (Table S2 lists major food items consumed before and throughout the trial).

During the intervention periods, subjects consumed, on a daily average, over 100 g of available carbohydrates from bread, a significant increase from an average of 42.2 ± 24.5 g during

<table>
<thead>
<tr>
<th>Table 1. Cohort Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>SW Group</strong></td>
</tr>
<tr>
<td>(Mean ± SD)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>WS Group</strong></td>
</tr>
<tr>
<td>(Mean ± SD)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>KS</strong></td>
</tr>
<tr>
<td>p Value</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Number of subjects (n)</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>Sex (% female)</td>
</tr>
<tr>
<td>60%</td>
</tr>
<tr>
<td>Age (y)</td>
</tr>
<tr>
<td>39.1 ± 14.3</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
</tr>
<tr>
<td>26.5 ± 5.6</td>
</tr>
<tr>
<td>HbA1c (%)</td>
</tr>
<tr>
<td>5.19 ± 0.23</td>
</tr>
<tr>
<td>Total cholesterol (mg/dL)</td>
</tr>
<tr>
<td>191 ± 40</td>
</tr>
<tr>
<td>Waist-to-hip circumference ratio</td>
</tr>
<tr>
<td>0.81 ± 0.08</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
</tr>
<tr>
<td>119 ± 17</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
</tr>
<tr>
<td>77 ± 8</td>
</tr>
</tbody>
</table>

*WS, white bread consumption followed by sourdough bread consumption. SW, sourdough bread consumption followed by white bread consumption. KS, Kolmogorov-Smirnov test.*
the run-in period (mean ± SD; Wilcoxon signed-rank p < 0.001, Table 2), representing an increase from 12.5% ± 6.6% (mean ± SD) to an average of over 22% of overall caloric consumption (Table 2). Prior to study initiation, subjects consumed comparable amounts of white (6.5% of calories; Table S2) and non-white bread (5.9%; Wilcoxon signed-rank p > 0.97; Table S2). No significant differences were found in the total caloric consumption between the run-in and either intervention period, and between caloric and macronutrient consumption both between interventions and between periods (Table 2). As consumption of available carbohydrates was purposely matched between the intervention periods, there was a significant difference in the fraction of calories obtained from bread between the two inter-

Using linear mixed models, we compared the treatment effects of a week-long consumption of white bread to that of sourdough bread on 20 clinical variables, using measurements taken as baseline (days 0, 21, Figure 1A) and outcome (days 7, 28, Figure 1A) of each clinical variable (STAR Methods). Notably, we found no significant difference between the two treatments both for the primary outcome measure of this trial, glycemic control, which we quantified using the response to an oral glucose tolerance test (OGTT; Figures 1B and 1C) and wakeup glucose levels of individuals (D), the primary outcome measures of the trial, at the beginning and end of each intervention period. The median responses of all subjects are depicted by a bold black line.

(E and F) The statistical significance (y axis) and effect size (x axis) of species estimated with metagenomic sequencing (E) or genera estimated with 16S rRNA gene sequencing (F). In red are instances that are significant after correcting for multiple testing with FDR of 0.1. See also Figure S1 for a CONSORT flow diagram, Figure S2 for all measured effects of the two bread types, Figure S3 for non-significant results regarding other microbiome features, Table S1 for nutritional composition of white and sourdough breads used, and Table S5 for raw measurements of outcome variables. GTT, glucose tolerance test.
capture a change as small as one standard deviation from expected change with no intervention (given 20 subjects; Kenward and Jones, 2003; Figure 1B). For example, the bounds of the 95% confidence interval for the effect of bread type on wakeup glucose were smaller than 7 mg/dL. Similarly, the bounds for the 95% confidence interval for bread-type effect on body weight were smaller than 400 g. Confidence intervals for estimates of bread-type effect on all outcome variables are listed in Figure 1B. We further note that some differences between the bread types could have been masked by the difference in fraction of calories obtained from bread, as well as by the increase in bread consumption that was employed here compared to the regular diet of the subjects.

Next, we sequenced DNA produced from stool samples collected as baseline (days –1, 20) and outcome (days 6, 27) with both metagenomic and 16S rRNA gene sequencing. These were used to calculate microbiota relative abundances at the species (metagenome), genus (16S), and phylum (16S) levels, the relative abundances of KEGG pathways and modules (Kanehisa and Goto, 2000; STAR Methods), and the Shannon \( \alpha \)-diversity index (16S OTUs and metagenomic species). We employed the same linear mixed models to assess the difference in treatment effect and found that the relative abundances of both \textit{Eubacterium ventriosum} species and the Anaerostipes genus significantly increased with white bread consumption compared to sourdough bread consumption (\( p < 0.001 \) FDR corrected at 0.1, Figures 1E and 1F). Notably, both \textit{E. ventriosum} and members of the Anaerostipes genus are butyrate producers (Mun˜oz-Tamayo et al., 2011; Tims et al., 2013), a short-chain fatty acid that was previously suggested to protect against colon cancer (Lupton, 2004) and obesity (Lin et al., 2012) and to abrogate colonic inflammation (Segain et al., 2000). Conversely, \textit{E. ventriosum} was previously associated with obesity in a study of monozygotic twins discordant for it (Tims et al., 2013). No significant differences in treatment effect were found for \( \alpha \)-diversity (\( p > 0.4 \)) for the phylum level relative abundances (Figure S3A) or for functional properties of the microbiome (Figures S3B and S3C).

**One-Week-Long Consumption of Bread Alters Multiple Blood Markers**

Given the lack of differential treatment effects on clinical parameters between white and sourdough bread, we next asked whether consumption of bread, regardless of its type, for 1 week affects metabolic and clinical markers. To this end, we performed a post hoc analysis and compared the value of each parameter at the beginning and the end of the first intervention week, examining all 20 subjects jointly. We note that this is a within-group comparison, limited by the absence of a control group, and that some effects could possibly stem from trial participation per se rather than bread consumption.

Notably, we found that a single week of bread consumption resulted in changes to multiple clinical variables and risk factors (Figures 2A–2K) that were statistically significant, albeit numerically small and not necessarily clinically significant. We found a significant decrease in levels of the essential minerals calcium, iron, and magnesium (Wilcoxon signed-rank \( p < 0.001 \), \( p < 0.01 \), and \( p < 0.01 \), respectively; Figures 2A, 2D, and 2F), in accordance with previous studies (Bach Kristensen et al., 2005; McCance and Widdowson, 1942; Reinhold et al., 1976), and a significant increase in the levels of lactate dehydrogenase (LDH; \( p < 0.01 \); Figure 2E), a marker of tissue damage. We also found a significant decrease in the levels of liver enzymes aspartate aminotransferase (AST; \( p < 0.005 \); Figure 2B); gamma-glutamyl transpeptidase (GGT; \( p < 0.005 \); Figure 2C); alanine aminotransferase (ALT; \( p < 0.05 \); Figure 2I), markers of liver function; C-reactive protein (CRP), a commonly used marker of inflammation (\( p < 0.05 \); Figure 2K); and creatinine (\( p < 0.05 \); Figure 2J), a marker of kidney function. We found a significant decrease in both total and low-density lipoprotein (LDL) cholesterol levels (\( p < 0.05 \) for both; Figure 2G,H) but no significant

---

**Table 2. Nutrient Consumption before and during Intervention Weeks**

<table>
<thead>
<tr>
<th>Run-In</th>
<th>Energy (kcal)</th>
<th>Total Lipid (g)</th>
<th>Protein (g)</th>
<th>Carbs. (g)</th>
<th>Avail. Carbs. from Bread (g)</th>
<th>Energy from Bread (% kcal)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>2,154 ± 431</td>
<td>94.1 ± 21.1</td>
<td>84.8 ± 32.8</td>
<td>219.3 ± 52.1</td>
<td>42.2 ± 24.5</td>
</tr>
<tr>
<td>White-Sourdough (WS)</td>
<td></td>
<td>2,336 ± 588</td>
<td>97.2 ± 26.9</td>
<td>103.5 ± 45.2</td>
<td>250.2 ± 79.4</td>
<td>100.0 ± 41.0</td>
</tr>
<tr>
<td>Week 1 (W; mean ± SD)</td>
<td></td>
<td>2,166 ± 465</td>
<td>90.7 ± 20.5</td>
<td>99.0 ± 39.7</td>
<td>244.7 ± 64.3</td>
<td>101.4 ± 40.7</td>
</tr>
<tr>
<td>Week 2 (S; mean ± SD)</td>
<td></td>
<td>2,187 ± 600</td>
<td>95.3 ± 29.3</td>
<td>89.6 ± 43.0</td>
<td>255.6 ± 56.5</td>
<td>107.2 ± 29.2</td>
</tr>
<tr>
<td>Wilcoxon Signed-Rank Test p Values</td>
<td></td>
<td>2,230 ± 494</td>
<td>91.1 ± 25.2</td>
<td>84.8 ± 31.5</td>
<td>261.0 ± 55.1</td>
<td>107.8 ± 29.8</td>
</tr>
</tbody>
</table>

| Run-in versus Week 1 | 0.146 | 0.655 | 0.041* | 0.006** | 0.0002*** | 0.0003*** |
| Run-in versus Week 2 | 0.093 | 0.794 | 0.057  | 0.003** | 0.0002*** | 0.0002*** |
| Week 1 versus Week 2 | 0.456 | 0.279 | 0.314  | 1.000   | 0.391    | 0.526     |
| Run-in versus White  | 0.053 | 0.941 | 0.048* | 0.005** | 0.0002*** | 0.0003*** |
| Run-in versus Sourdough | 0.156 | 0.941 | 0.048* | 0.002** | 0.0002*** | 0.0002*** |
| Sourdough versus White | 0.093 | 0.709 | 0.794  | 0.766   | 0.794    | 0.002**   |

* \( p < 0.05 \), ** \( p < 0.01 \), *** \( p < 0.001 \). Avail., available; Carbs., carbohydrates. See also Table S2 for fraction of calories obtained from specific food items and Table S3 for a replication of this analysis with linear mixed models.
change in high-density lipoprotein (HDL) cholesterol levels (Figure 2L). Non-significant changes in other outcome variables are shown in Figure S4. Overall, and despite the limitations of this analysis, understanding the effect of different foods on clinical parameters is a first step toward combining different foods in order to personally tailor diets that produce clinical effects.

Overall Microbiota Composition Is Resilient to Bread Consumption

Several studies indicated that even short-term dietary interventions, whether animal-based, plant-based (David et al., 2014), or involving barley bread consumption (Kovatcheva-Datchary et al., 2015), result in significant, rapid, and reproducible alterations to the gut microbiome.

Subjects in our study provided stool samples from the start and end of each intervention period, days 1, 6, 20, and 27 (Figure 1A; STAR Methods). We performed a principal coordinate analysis (PCoA) of microbial species abundances (Figure 3A, derived from metagenome sequencing) and of OTUs abundances (Figure S5A, derived from 16S rRNA gene sequencing) and found a significant clustering of the microbiota compositions of each subject, which remained similar to themselves throughout the crossover trial (ANOVA p < 10^{-20} for species, Figure 3A; p < 10^{-20} for OTUs, Figure S5A). Analysis with ANCOM (Mandal et al., 2015) found 126 out of 175 species and 565 out of 604 OTUs to separate significantly between different individuals throughout the trial (FDR < 0.1).

To evaluate the magnitude of changes in microbiota composition, we compared it to that of the changes in the run-in period of a previous trial in which microbiome samples were collected but no dietary intervention was performed (Zeevi et al., 2015; Table S4 lists baseline characteristics; STAR Methods), using abundances estimated using both 16S rRNA gene and metagenomic sequencing (STAR Methods). We found no significant differences in α-diversity (Shannon) between baseline and outcome measurements of both intervention weeks (Wilcoxon p > 0.4 for species, Figure 3B; p > 0.05 for OTUs, Figure S5B) concurrent with lack of significant changes in the non-interventional cohort. We found no significant differences in the change to α-diversity of metagenomic species between the intervention
weeks and the non-intervention cohort (Mann-Whitney p > 0.5, Figure 3C), although there was a small yet significant difference in the change to OTUs-derived \( a \)-diversity between the first intervention week and the non-interventional cohort (\( p = 0.03 \), Figure S5C). Notably, we found no increase in \( b \)-diversity compared to the non-interventional cohort for both metagenomic species and OTUs (one-sided Mann-Whitney \( p > 0.8 \) for species Figure 3D; \( p > 0.9 \) for OTUs, Figure S5D).

When adding the species composition of the samples from the previous trial with no intervention to the PCoA plot, we again found changes to the microbiota species composition of each subject to be small, such that samples from the same subject clustered closely together (Figure 3E). Sizes of minimal bounding ellipses of samples from the same subject on the PCoA plot showed no significant change (Mann-Whitney \( p > 0.45 \), Figure S5E). Anecdotally, one subject was enrolled in both studies (colored yellow and green, Figure 3E) and retained similar microbiota composition despite the prolonged time passed between studies (over 18 months).

The absence of a significant species compositional change in the microbiota is surprising in light of studies that showed such effects as a result of dietary interventions both in general (David et al., 2014) and following whole-wheat consumption (Garcia-Mazcorro et al., 2014).
et al., 2016). Taken together, our results demonstrate that even though the intervention performed here was radical enough to significantly change clinical variates, microbiota composition underwent only minor alterations, demonstrating that the gut microbiota is resilient to some types of nutritional changes.

High Interpersonal Variability in the Postprandial Glycemic Response to Bread

The lack of differential treatment effects between white and sourdough bread implies that either the two breads exert similar effects within each individual or, more intriguingly, that the effect of each bread is person specific such that, averaged across subjects in the cohort, bread type does not affect the overall average. One of the ways through which bread, a carbohydrate-rich food, may exert its clinical effect is through changes in blood glucose levels that it elicits (PPGR) and the cascade of systemic effects that follow. Examining the difference between the average PPGR to white bread and the average PPGR to sourdough bread strongly supports the latter hypothesis, as numerous individuals exhibit opposite PPGRs to white and sourdough bread with or without butter, in the form of standardized meals (STAR Methods). We devised a statistical framework aimed at testing whether variability in PPGR to either type of bread is higher than expected given the intrapersonal variability (STAR Methods). Our null hypothesis, based on the glycemic index (Jenkins et al., 1981), posits that there is no personalized response, i.e., that the PPGR to different types of meals is an intrinsic property of the food and that any variability in PPGR is strictly intrapersonal. More specifically, our null hypothesis assumes that, for any food, the ratio between its PPGR and the PPGR to glucose should be the same for all individuals in the population. Thus, under this null hypothesis, the PPGR of an individual to a specific food should follow a normal distribution with a mean equal to the PPGR to that individual multiplied by the ratio between the PPGR to the food and to glucose in the population, and with a standard deviation equal to the intrapersonal variability of that individual. Standardizing the PPGR measurements of each person to a given food using these estimators should thus result in a standard normal distribution (mean of 0 and variance of 1).

Figure 4. Significant Interpersonal Variability in the Postprandial Glycemic Response to Bread

(A) Differences between the average PPGR to white and sourdough bread, for subjects in the SW (yellow) and WS (purple) groups. Note that ten subjects had higher responses to sourdough bread.

(B) The distribution of standardized responses to white bread (blue), white bread with butter (light blue), sourdough bread (purple), and sourdough bread with butter (light purple) exhibit significantly higher variance \((F\text{-test } p < 10^{-10})\) compared to the expected distribution based only on intrapersonal variability (black line).

(C) A comparison of the standard deviation of standardized postprandial responses to different meals, showing that each meal type separately exhibits significantly higher variability in the response than is expected based only on intrapersonal variability \((F\text{-test } p < 10^{-10}; *p < 0.001)\).

(D) Receiver operating characteristic (ROC) curve of classification of the PPGR-inducing type of bread for each subject based only on initial microbiome data, \(AUC = 0.83\). Predictions were generated using a gradient boosting regression algorithm in a leave-one-out cross validation. S.D., sourdough. WS, white bread consumption followed by sourdough bread consumption. SW, sourdough bread consumption followed by white bread consumption.

(Figure 4A). This result is further supported by previous evidence regarding the high interpersonal variability in PPGRs to real-life meals (Zeevi et al., 2015).

To test whether there is a statistically significant personalized response to bread type, we used the PPGRs of our subjects to 75 g of glucose and to the different types of bread consumed, with or without butter, in the form of standardized meals (STAR Methods). We devised a statistical framework aimed at testing whether variability in PPGR to either type of bread is higher than expected given the intrapersonal variability (STAR Methods). Our null hypothesis, based on the glycemic index (Jenkins et al., 1981), posits that there is no personalized response, i.e., that the PPGR to different types of meals is an intrinsic property of the food and that any variability in PPGR is strictly intrapersonal. More specifically, our null hypothesis assumes that, for any food, the ratio between its PPGR and the PPGR to glucose should be the same for all individuals in the population. Thus, under this null hypothesis, the PPGR of an individual to a specific food should follow a normal distribution with a mean equal to the PPGR of that individual to glucose, multiplied by the ratio between the PPGR to the food and to glucose in the population, and with a standard deviation equal to the intrapersonal variability of that individual. Standardizing the PPGR measurements of each person to a given food using these estimators should thus result in a standard normal distribution (mean of 0 and variance of 1).
Notably, when performing this standardization process and comparing the distribution of standardized meals with the expected standard normal, we observe significantly higher variances both when combining all meals (F-test for equality of two variances \( p < 10^{-10}; \) Figure 4B) and when each meal type is considered separately (\( p < 0.05; \) Figure 4C).

This result indicates that individuals exhibit personalized PPGRs to bread, reinforcing our previous finding that a major source of variability in response to food is interpersonal (Zeevi et al., 2015). Thus, to the extent that the effect of a food on metabolic parameters is partly mediated by the PPGR induced by that food, our results suggest that such personalized glycemic responses should be measured and accounted for when evaluating the effects of food. We hypothesize that this may explain several contradictions in the literature regarding the effects of different foods.

**Prediction of Personal Glycemic-Response-Inducing Bread Using Microbiome Features**

Following the significant interpersonal variability that we detected in the PPGRs to the different meal types and the manifestation of opposite PPGRs to the two bread types, we next asked whether we could predict from baseline measures in advance whether it is white or sourdough bread that will induce lower glycemic responses for each individual.

As a reference, we note that the published glycemic index for similar bread types was 70 and 54 for white bread and sourdough bread, respectively (Foster-Powell et al., 2002). As the glycemic index inherently ignores inter-individual differences, it would result in wrong classification for the ten subjects who had lower glycemic responses to white bread than to sourdough bread (Figure 4A).

We devised a classification algorithm based on gradient boosting regression (Friedman, 2001), as these algorithms can model complex non-linear relationships. We used only features derived from the baseline microbiome sample (collected at \( T = -1 \) days, Figure 1A), namely the relative abundances of microbiome species derived using MetaPhlAn (Truong et al., 2015); relative abundances of microbial genes, KEGG modules, and KEGG pathways (Kanehisa and Goto, 2000); and the extent that the effect of a food on metabolic parameters is partly mediated by the PPGR induced by that food, our results suggest that such personalized glycemic responses should be measured and accounted for when evaluating the effects of food. We hypothesize that this may explain several contradictions in the literature regarding the effects of different foods.

We evaluated the performance of our model using the standard leave-one-out cross validation (CV) scheme, whereby a model is trained on all subjects except one and then produces a prediction for the left-out subject. In each CV fold, we performed extensive feature selection to ensure that the number of features is not excessive considering the number of training examples, thereby reducing the risk of overfitting. To reduce the immense feature space of microbial gene relative abundances, we performed a principal component analysis (PCA) on the training set of each CV and selected four PCs. For species, pathways, and modules abundances, we removed sparse features and those correlated to other features across the training set of each CV and then retained only the top six features best correlated with the outcome in the training set of each CV fold.

Notably, the model predictions were highly accurate with an area under the receiver operating characteristic (ROC) curve (AUC) of 0.83 (Figure 4D), demonstrating that the glycemic-response-inducing bread can be accurately classified for each subject using only microbiome data. We note that further research is needed to validate these results on additional data. Fitting the prediction algorithm on all subjects combined revealed that some of the informative features were the MetaPhlAn-derived relative abundances of *Coprobacter fastidiosus* and *Lachnospiraceae bacterium 3_1_46FAA*.

While exact mechanisms for the effect of the gut microbiota on glycemic control are yet to be elucidated, our highly accurate predictor demonstrates that personalized bread responses can be predicted and that such predictors can bear clinical significance. Moreover, as this prediction was based on data derived from stool samples, it demonstrates the adequacy of this proxy measurement in such a prediction.

**CONCLUSIONS**

In this study, we found no significant differences on a broad array of clinical parameters between two 1-week-long dietary interventions, one including consumption of industrial white bread made from refined wheat and the other consumption of sourdough-leavened bread made from whole grains using traditional methods. This was despite statistically significant, albeit small, changes in the levels of the same parameters following the first week of bread consumption, regardless of its type. Gut microbiome analysis revealed differential treatment effects between the two types of bread in only two taxa and showed that the microbiota composition remained generally stable and person specific throughout this trial. Finally, using a rigorous statistical framework, we show marked and highly significant interpersonal variation in the glycemic response to the two types of bread, with some subjects featuring a higher response to one bread and some to the other. We further show that the type of bread that induces higher glycemic responses could be predicted for each subject using only microbiome data.

Understanding the interpersonal variation in the effect of bread, one of the most-consumed staple foods, would allow the personalization of bread-related nutritional recommendations and optimization of food choices worldwide. More broadly, our study underlines the importance of personalization in dietary recommendations, as even the straightforward comparison of breads commonly considered “healthy” and “unhealthy” revealed personal effects on PPGR, suggesting that universal dietary recommendations may have limited efficacy.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Main study cohort
  - Reference cohort
- **METHOD DETAILS**
  - Clinical trial
  - Exclusion and inclusion criteria
SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cmet.2017.05.002.

AUTHOR CONTRIBUTIONS

T.K. and D.Z. conceived the project, designed the study, conducted all analyses, interpreted the results, and wrote the manuscript. T.K. and D.Z. equally contributed to this work and are listed in random order. N.Z. designed the study and interpreted the results. O.W. and N.B. assisted in analyses. A.W. designed the study, directed sample sequencing, and together with M.L.-P. and T.A.-S. performed metagenomic extraction and sequencing. M.L.-P. performed 16S sequencing. G.M., N.K., and M.R. designed the study and coordinated data collection. J.S. designed the study. B.Z.G., A.A.L., E.E., and E.S. conceived the project and designed the study. A.A.L., E.E., and E.S. designed and conducted the analyses, interpreted the results, and wrote the manuscript.

CONFLICTS OF INTEREST

E.S. and E.E. are paid scientific consultants for DayTwo Inc.

ACKNOWLEDGMENTS

We thank the Segal and Elinav group members for fruitful discussions, Ohad Manor for his help with analyses, Naomi Avivi-Ragolsky for technical help, Anomar Ogen for baking the sourdough bread, and Amonon Berezin from Kemach Haaretz for milling the flour. T.K. and D.Z 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; the Leona M. and Harry B. Helmsley Charitable Trust; the Gurwin Family Fund for Scientific Research; the Crown Endowment Fund for Immunological Research; the estate of Jack Gitlitz; the estate of Lydia Hershkovich; the Benoziyo Endowment Fund for the Advancement of Science; the Adelis Foundation; John L. and Vera Costello, Sherman Oaks; grants funded by the European Research Council; the Crown Human Genome Center; the Else Kroener Fresenius Foundation; Donald L. Schwarz, Sherman Oaks, CA; Jack N. Halpern, New York, NY; Leesa Steinberg, Canada; and grants funded by the European Research Council and the Israel Science Foundation.

Received: November 8, 2016
Revised: March 12, 2017
Accepted: May 10, 2017
Published: June 6, 2017

REFERENCES


STAR METHODS

KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological Samples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stool samples of reference cohort used for 16S rDNA amplification and sequencing</td>
<td>Zeevi et al., 2015</td>
<td>N/A</td>
</tr>
<tr>
<td>Critical Commercial Assays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PowerMag Soil DNA isolation kit</td>
<td>MoBio Laboratories</td>
<td>N/A</td>
</tr>
<tr>
<td>Deposited Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw sequence files, 16S sequencing, reference cohort</td>
<td>This paper</td>
<td>ENA: PRJEB17643</td>
</tr>
<tr>
<td>Raw sequence files, metagenomic sequencing, reference cohort</td>
<td>Zeevi et al., 2015</td>
<td>ENA: PRJEB11532</td>
</tr>
<tr>
<td>Raw sequence files, main study cohort</td>
<td>This paper</td>
<td>ENA: PRJEB17643</td>
</tr>
<tr>
<td>Human reference genome</td>
<td>UCSC (<a href="https://genome.ucsc.edu">https://genome.ucsc.edu</a>)</td>
<td>hg19</td>
</tr>
<tr>
<td>Integrated reference catalog of the human gut microbiome</td>
<td>Li et al., 2014</td>
<td><a href="http://meta.genomics.cn">http://meta.genomics.cn</a></td>
</tr>
<tr>
<td>Database of complete bacterial genomes</td>
<td>Korem et al., 2015</td>
<td>Available upon request from <a href="mailto:eran.segal@weizmann.ac.il">eran.segal@weizmann.ac.il</a></td>
</tr>
<tr>
<td>Oligonucleotides</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16S PCR primer: GTGCCAGCMGCGCGGTAA</td>
<td>IDT (<a href="http://www.idtdna.com/site">http://www.idtdna.com/site</a>)</td>
<td>515F</td>
</tr>
<tr>
<td>16S PCR primer: GGACTACHVGGGTWTCTAAT</td>
<td>IDT (<a href="http://www.idtdna.com/site">http://www.idtdna.com/site</a>)</td>
<td>806R</td>
</tr>
<tr>
<td>Software and Algorithms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEM mapper build 1.376</td>
<td>Marco-Sola et al., 2012</td>
<td><a href="https://sourceforge.net/projects/gemlibrary/">https://sourceforge.net/projects/gemlibrary/</a></td>
</tr>
<tr>
<td>USEarch8.0</td>
<td>Edgar, 2013</td>
<td><a href="http://www.drive5.com/usearch/">http://www.drive5.com/usearch/</a></td>
</tr>
<tr>
<td>QIIME 1.8.0</td>
<td>Caporaso et al., 2010</td>
<td><a href="http://qiime.org/">http://qiime.org/</a> ; RRID: SCR_008249</td>
</tr>
<tr>
<td>MetaPhlAn2</td>
<td>Truong et al., 2015</td>
<td><a href="https://bitbucket.org/biobakery/metaphlan2">https://bitbucket.org/biobakery/metaphlan2</a>; RRID: SCR_004915</td>
</tr>
<tr>
<td>SAS v. 9.4</td>
<td>SAS institute</td>
<td><a href="https://www.sas.com">https://www.sas.com</a>; RRID: SCR_004635</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurements of outcome variables</td>
<td>This paper</td>
<td>Table S5</td>
</tr>
</tbody>
</table>

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Eran Segal (eran.segal@weizmann.ac.il).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Main study cohort
Twenty healthy adults were included in this study, including nine males (four in SW, five in WS) and 11 females (six in SW, five in WS), 27-66 years of age (38.2 ± 11.3 [mean ± SD], Table 1 for group characteristics). Study protocol was approved by Tel Aviv Sourasky Medical Center Institutional Review Board (IRB), approval numbers TLV-0658-12, TLV-0050-13 and TLV-0522-10. All subjects signed written informed consent forms.
**Reference cohort**
Analyzed here are samples from a previous trial (Zeevi et al., 2015). Baseline characteristics of all subjects analyzed are given in Table S4.

**METHOD DETAILS**

**Clinical trial**
This study was a single-center randomized crossover trial, performed at the Weizmann Institute of Science, Israel. No changes were done to the study protocol and methods after the trial commenced.

Thirty-four subjects were recruited for this study throughout February 2016. 12 potential subjects withdrew due to personal reasons or study requirements and two were excluded due to exclusion criteria, all before randomization. Randomized group allocations were computer generated with no restrictions. Recruitment and follow-up were performed by certified clinical research associates. Twenty healthy subjects were eventually randomized without any restrictions, and the trial took place between February 21st and April 5th, 2016 (Figure S1).

The trial was completed as planned. All twenty subjects completed the trial and there were no drop-outs or withdrawals. No adverse effects were reported by any of the study subjects. Adherence to the interventions was assessed using real-time logging of all meals consumed, performed on a proprietary smartphone application that we developed (Zeevi et al., 2015).

**Exclusion and inclusion criteria**
All subjects fulfilled the following inclusion criteria: males and females, aged 18-70, who are currently not following any diet regime or diettian consultation and are able to provide informed consent and technically operate a glucometer for oral glucose tolerance test (OGTT) and the calibration of the CGM. Exclusion criteria included: (i) pregnancy or fertility treatments; (ii) usage of antibiotics or anti-fungals 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 CGM; (vi) active neuropsychiatric disorder; (vii) myocardial infarction or cerebrovascular accident in the 6 months prior to participation; (viii) coagulation disorders; (ix) chronic immunosuppressive medication usage; (x) pre-diagnosed type I or type II diabetes mellitus or treatment with anti-diabetic medication. Adherence to inclusion and exclusion criteria was validated by medical doctors.

**Outcome variables**
We measured outcome variables at baseline, after the first intervention period, after the washout period, and at the end of the second intervention period (Figure 1A). Predetermined outcome variables consisted of: (a) blood tests, performed after fasting at days 0, 7, 21 and 28, measuring levels of: triglycerides, LDL cholesterol, HDL cholesterol, total cholesterol, ALT, AST, GGT, iron, calcium, creatinine, urea, thyroid stimulating hormone (TSH), LDH and CRP; (b) blood pressure measurements, taken with an automated blood pressure monitor (M6 model, Omron, Hoofddorp, the Netherlands) at days 0, 7, 21 and 28; (c) weight and basal metabolic rate (BMR) measurements, taken with the BC-418 Segmental Body Composition Analyzer (Tanita, Tokyo, Japan) at days 0, 7, 21 and 28; (d) average blood glucose levels in the 15 min following wakeup (referred as “wakeup glucose”), taken as the average of four continuous glucose monitor (CGM; iPro2 by Medtronic, MN, USA) measurements at days 1 & 2, 5 & 6, 22 & 23, and 26 & 27; (e) blood glucose response to OGTT quantified as the incremental area under the glucose curve (iAUC with respect to the median glucose in the 30 min prior to the meal; Zeevi et al., 2015), after consumption of 75 g of glucose, as measured with a blood glucose monitor (Contour by Bayer AG, Leverkusen, Germany), and calculated as the average of two tests performed at days – 6 & –5, 8 & 9, 15 & 16, and 29 & 30; (f) and microbiome analyses, performed on stool samples taken within 24 hr of days –1, 6, 20 and 27. All measurements are provided in Table S5.

**Bread preparation**
Flour was freshly stone milled from hard red bread wheat (Triticum aestivum var. aestivum). The resulting flour was sifted in order to remove only the largest particles of bran, resulting in a 98% “extraction rate” (1000 g of wheat yielded 980 g of flour). Loaves were prepared using the following four ingredients: freshly stone milled flour as described above, water, salt and a mature sourdough starter without any other additive. The overall formula for the bread was (in bakers’ percentage) 100% flour, 90% water, and 1.8% salt with the sourdough starter portion representing 37% of total flour weight (i.e., 20% of the flour in the formula had been pre-fermented). The dough was kneaded in a planetary mixer, “bulk” fermented for 1 hr at 24°C and “retarded” at 4°C for 8 hr. The dough was portioned into 1150 g pieces, shaped and transferred into loaf pans, where they underwent “proofing” (final rising) for 2 hr at 24°C. The loaves were baked in a stone hearth oven at 245°C for approximately 60 min to obtain final loaf weight of 1kg after baking.

The composition of the sourdough bread was determined by the food and water laboratory of Bactochem (Ness-Ziona, Israel), a company certified by the Israeli Ministry of Health, except for the fiber content that was estimated from the USDA National Nutrient
Database values “Basic Report 2008, Wheat flour, whole-grain” after subtracting 2% of the sifted bran. For the industrial bread, we used manufacturer labeling. The composition of both bread types is depicted in Table S1.

**Standardized meals**
We supplied our subjects with standardized meals at the start of each intervention week, according to the intervention type. Standardized meals were calculated to have 50 g of available carbohydrates. During the white bread intervention period, subjects consumed 3 meals of 110 g white bread, and 3 meals of 110 g white bread + 30 g butter. During the sourdough bread intervention period, subjects consumed 3 meals of 145 g sourdough bread, and 3 meals of 145 g sourdough bread + 30 g butter. Subjects 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.

**Drug usage**
In the WS group, two subjects sporadically reported consumption of NSAIDs. One subject reported regular consumption of multivitamins and omega 3. In the SW group, one subject reported regular consumption of vitamin D3, Eltroxin and Rampiril, and sporadic consumption of antacids; one subject reported regular consumption of iron supplementation; one subject reported consumption of birth control pills and multivitamins, and sporadic consumption of NSAIDs; and one subject reported regular consumption of birth control pills and sporadic consumption of NSAIDs.

**Gut microbiome sampling and sequencing**
Subjects self-sampled their stool using a swab following detailed printed instructions. Collected samples were immediately stored in a home freezer (−20 °C) and transferred in a provided cooler to our facilities, where they were stored at −80 °C. We purified DNA 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 (Zeevi et al., 2015), and sequenced on the Illumina NextSeq platform with a read length of 80bp to a depth of 8,429,912 ± 5,568,845 reads (mean ± SD). Two samples with < 1M reads were removed from further analyses. For 16S rRNA gene sequencing, PCR amplification of the V3/4 region using the 515F/806R 16S rRNA gene primers was performed followed by 500bp paired-end sequencing on the Illumina Miseq platform to a depth of 77,872 ± 61,464 reads (mean ± SD). 2 samples were not analyzed due to technical issues in the 16S rRNA gene amplification process. 2 samples with fewer than 30,000 reads were removed from further analyses. All 16S analyses were performed following subsampling to 30,000 reads.

**Gut microbiome analysis**
We filtered metagenomic reads containing Illumina adapters, filtered low quality reads and trimmed low quality read edges using Trimmmomatic (Bolger et al., 2014). We detected host DNA by mapping with GEM (Marco-Sola et al., 2012) to the Human genome (hg19) with inclusive parameters, and removed those reads. 16S analysis was performed using USearch8.0 (Edgar, 2013) using the commands fastq_mergepairs, derep_fulllength, sortbysize, out_radius_pct, uchime_ref, usearch_global, followed by the script assign_taxonomy.py from QIIME (Caporaso et al., 2010).

We obtained relative abundances (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 sum to 1. Metagenomics-derived RAs were capped at 10⁻⁵, while 16S rRNA-derived RAs were capped at 10⁻⁴.

PCoA was calculated on species (metagenomics) or OTUs (16S) RA, using Bray-Curtis dissimilarity. α-diversity for each individual was calculated on species RA (metagenomics) or OTUs (16S) using the Shannon index. β-diversity was calculated using the Bray-Curtis divergence between baseline and outcome samples in each intervention week and at the start and end of the run-in period of a previous trial (Zeevi et al., 2015). Only taxa present in at least 10 samples were included in analyses, except when calculating α-diversity.

**Reference cohort**
Analyzed here are samples from the run-in period of a dietary intervention trial from (Zeevi et al., 2015), in which no dietary intervention was performed. For metagenomic samples previously generated sequences (Zeevi et al., 2015) were used. For 16S rRNA gene sequencing, samples were processed anew as described above. Only subjects who had more than three samples collected were included in the analysis in Figures 3A, 3E and S5A. Baseline characteristics of all subjects analyzed are given in Table S4.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**Statistical analyses**
Analysis of treatment effect in the crossover trial and calculation of 95% confidence interval was done using linear mixed models as recommended (Mehrotra, 2014), using SAS version 9.4, estimating treatment effect with fixed effects for period, sequence group, the
difference between baseline measurements at both intervention periods, and interaction between that difference and the period. No other co-variates were included in the model. For the effect of bread on the gut microbiome (Figures 1E, 1F, and S3), only instances (e.g., genus, module) with abundance above capping level observed in both weeks in at least 15 subjects were included in the analysis. Results were corrected for multiple hypothesis testing using FDR of 0.1 for each phylogenetic or functional level separately. Analysis of general bread consumption effect was performed using Wilcoxon signed-rank test, with results corrected for multiple hypothesis testing using FDR of 0.1.

**Classification of PPGR inducing bread type**

We employed a stochastic gradient boosting regression algorithm which was trained on the difference between the average responses to white and sourdough bread divided by the average OGTT response, and tested against the true classification to two categories of lower PPGRs to white or sourdough bread using leave-one-out cross validation (CV). The prediction used: (a) metagenomics-derived species (capped at 10^5), pathways and module abundances (corrected to mapping percentage to the gene catalog of (Li et al., 2014)), of which, 6 non-sparse features were selected based on Pearson correlation to the target value in the training set of each CV fold; (b) four principal components of composition of genes obtained by mapping to (Li et al., 2014), with principal component analysis calculated only on the training set of each CV fold; (c) the number of said genes present in the samples; and (d) mapping percentage to the human genome, the gene catalog, and a database of complete bacterial genomes (Korem et al., 2015).

**Measurement of interpersonal variability**

To calculate the expected PPGR ratio for each meal, for each subject, we average the PPGR to every standardized meal and divided this average by their average PPGR to glucose:

\[ R_{f(i)} = \frac{PPGR_{fglu(i)}}{PPGR_{fglu}} \]

Where \( R_{f(i)} \) is the ratio of the response of subject \( i \) to food \( f \), and \( PPGR_{fglu(i)} \) is the average PPGR of subject \( i \) to food \( f \). Next, we calculate this average ratio in the population, excluding said subject, i.e., this calculation is performed for every subject, by only taking into account other subjects:

\[ \bar{R}_{f(-i)} = \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} R_{f(j)} \]

Where \( n \) is the total number of subjects participating in the study. \( \bar{R}_{f(-i)} \) is the PPGR ratio for food \( f \) used for future calculations regarding subject \( i \).

To calculate the expected response of each person to every food, for each subject we calculate \( PPGR_{fglu} \), the average response of subject \( i \) to glucose, measured over 6-8 measurements. We estimate the expected response to food \( f \) by multiplying the average response to glucose for each subject by the PPGR population-level ratio for food \( f \) excluding the subject in question:

\[ PPGR_{fglu(i)} = PPGR_{fglu} \cdot \bar{R}_{f(-i)} \]

We estimate the intra-individual standard deviation of the response to food \( f \) by constructing a random sample consensus (RANSAC) linear model of subjects’ standard deviation of PPGR to glucose given their median PPGR to glucose. In this linear model we do not fit an intercept (i.e., require that for a PPGR of 0, the estimated standard deviation is 0). Using this linear model we estimate:

\[ SD_{f(i)} = Im \left( PPGR_{fglu(i)} \right) = Im \left( PPGR_{fglu} \cdot \bar{R}_{f(-i)} \right) \]

Where \( Im \) is the linear model in question. The total expected standard deviation is therefore given by:

\[ totSD_{f(i)} = SD_{f(i)} \cdot \sqrt{1 + \frac{1}{N_{fglu(i)}}} \]

Where \( N_{fglu(i)} \) is number of measurements of glucose PPGR for subject \( i \).

Standardization of the glucose response to every food \( f \) was performed by subtracting the estimated expected response to the food and dividing by the estimated standard deviation:

\[ SPPGR_{f(i)} = \frac{PPGR_{fglu(i)} - PPGR_{fglu}}{totSD_{f(i)}} \]

Where \( SPPGR_{f(i)} \) is the standardized postprandial glycemic response of person \( i \) to food \( f \). Under the null hypothesis of no interpersonal variability, the distribution of \( SPPGR_{f(i)} \) over all foods across all subjects is expected to be a standard normal distribution, with mean 0 and variance 1.
DATA AND SOFTWARE AVAILABILITY

Sequencing data has been deposited at the European Nucleotide Archive with accession ENA: PRJEB17643. All measurements of outcome variables are available in Table S5.

ADDITIONAL RESOURCES

Trial was reported to clinicaltrials.gov, NCT: NCT02936362.