

Online Supporting Material

Supplemental Table 1 – Features used for creation of the predictive model

Supplemental Figure 1 – Comparison of species level microbial abundances between cohorts

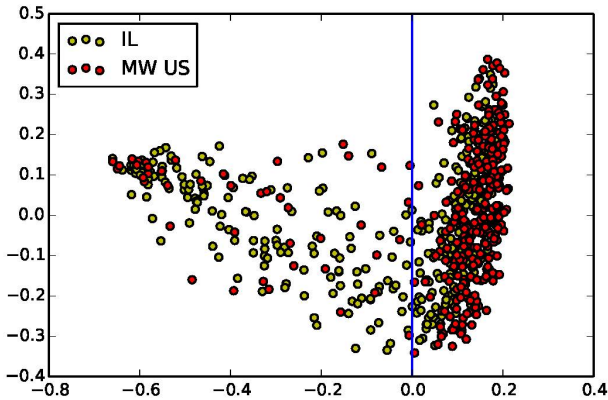
Supplemental Figure 2 - Partial Dependence Plots for all 72 features used depicting the influence of each feature in the predicted postprandial response level.

Supplemental Table 1 – Features used for creation of the predictive model

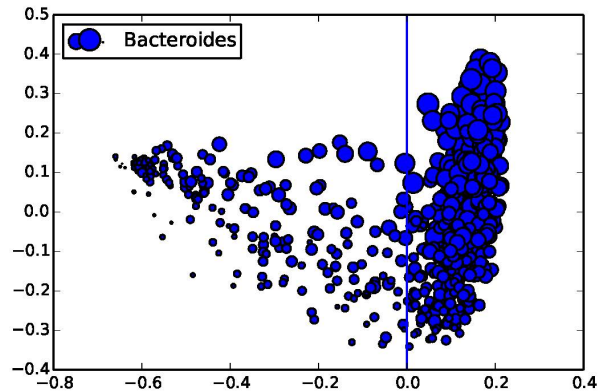
Features					
Meal	Meal context	Blood tests	Personal	Derived from CGM measurements	Microbiome
Carbohydrates (g)	Caloric sum of 6 hours (kcal)	HbA1c%	Physical activity (mins)	CGM (dynamic range)	<i>Subdoligranulum</i> (relative abundance)
Sodium (mg)	Caloric sum of 3 hours (kcal)		Physical activity frequency	1hour ago relative to global median	<i>Bacteroides</i> (relative abundance)
Carbohydrates to fat	Carbohydrates sum of 6 hours (g)		Morning hunger	Delta AUC starts 2 hours ago	<i>Lachnospiraceae</i> (relative abundance)
Dietary fiber (g)	Fiber sum in last 24 hours (g)		Sleep quality	Long trend 1 hour ago	<i>Eubacterium rectale</i> (relative abundance)
Energy (kcal)	Carbohydrates sum of 12 hours (g)		Stress	Long trend 2 hours ago	<i>Alistipes putredinis</i> (relative abundance)
Alcohol (g)	Caloric sum of 12 hours (kcal)		Evening hunger	Long trend 4 hours ago	Verrucomicrobia (relative abundance)
Lipid (g)	Carbohydrates sum of 3 hours (g)		Currently smokes	Short trend 1 hour ago	<i>Prevotella</i> (relative abundance)
Sugar (g)	Fiber sum in last 12 hours (g)		Ever smoked	Short trend 2 hours ago	<i>Eubacterium eligens</i> (relative abundance)
Caffeine (mg)	Caloric sum of 2 hours (kcal)		Midday hunger	Delta AUC starts 4 hours ago	<i>Bacteroides vulgatus</i> (relative abundance)
Protein (g)	Time of day (h)		General hunger	Short trend 4 hours ago	<i>Bacteroides dorei</i> (relative abundance)
Water (g)			Weight		<i>Bacteroides thetaiotaomicron</i> (relative abundance)
			BMI		<i>Bacteroides xylanisolvens</i> (relative abundance)
			Age		Actinobacteria (relative abundance)
			Gender		<i>Alistipes finegoldii</i> (relative abundance)
			Regular defecation		Proteobacteria (relative abundance)
					<i>Roseburia hominis</i> (relative abundance)
					Firmicutes (relative abundance)
					<i>Roseburia inulinivorans</i> (relative abundance)
					Bacteroidetes (relative abundance)
					<i>Parabacteroides distasonis</i> (relative abundance)
					<i>Bacteroides uniformis</i> (relative abundance)
					<i>Eubacterium siraeum</i> (relative abundance)
					<i>Ruminococcus bromii</i> (relative abundance)
					<i>Alistipes onderdonkii</i> (relative abundance)
					<i>Akkermansia muciniphila</i> (relative abundance)

Supplemental Figure 1 – Comparison of species level microbial abundances between cohorts (A-C) PCoA analysis of species Bray-Curtis distances between participants. Vertical line at zero is used for visualization purposes. Analysis was performed using PC-ord (A) points are color coded by cohort. (B) size of points corresponds to total abundance of *Bacteroides* species (C) size of points corresponds to total abundance of *Prevotella* species. This analysis (panels A-C) show a significant uneven distribution of samples from the Midwestern (red circles) and Israeli (yellow circles) cohorts (χ^2 contingency p-value= $2e-10$), which is at least partly explained by the difference in *Prevotella* and *Bacteroides* species abundances between the cohorts. The species that most significantly differ in abundance between the left cluster in the PCoA plot, which is enriched in the Israeli cohort participants, and the right part of the plot, are *Prevotella copri* and to a lesser extent *Bacteroides uniformis* (Bonferonni corrected Welch's t-test p values of $2e-172$ and $1e-19$, respectively). (D) Differential abundance of species in the two cohorts. The scatter plot shows average log abundance of species in the two cohorts. Only species with an average abundance $> 0.1\%$ in at least one of the cohorts were considered. Log abundances were clipped to the $[-4, 4]$ range. Species for which there is > 2 -fold difference between cohorts were tested for significance using t-test. Those species with significantly different means (< 0.01 Bonferonni-corrected p-value using t-test) are highlighted (yellow - high abundance in Israeli cohort, red - high abundance in Midwestern cohort).

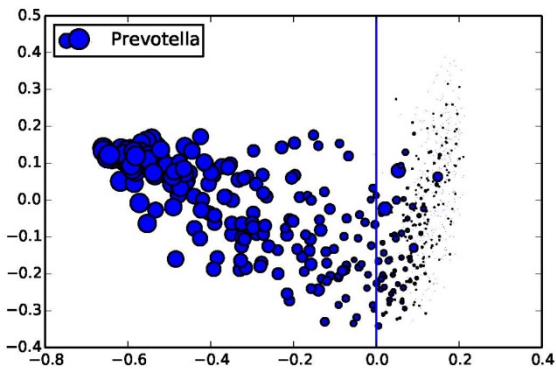
A)



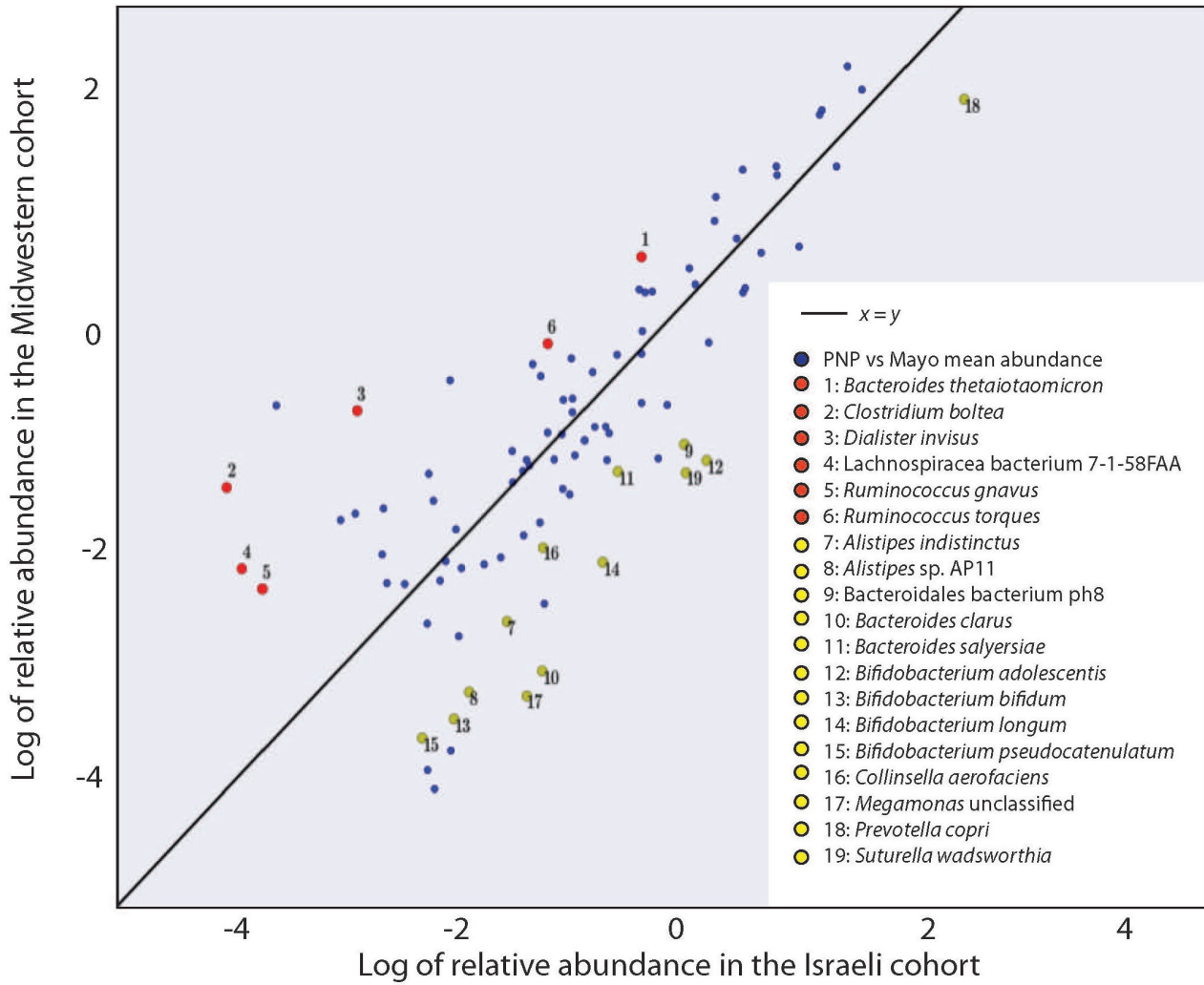
B)



C)

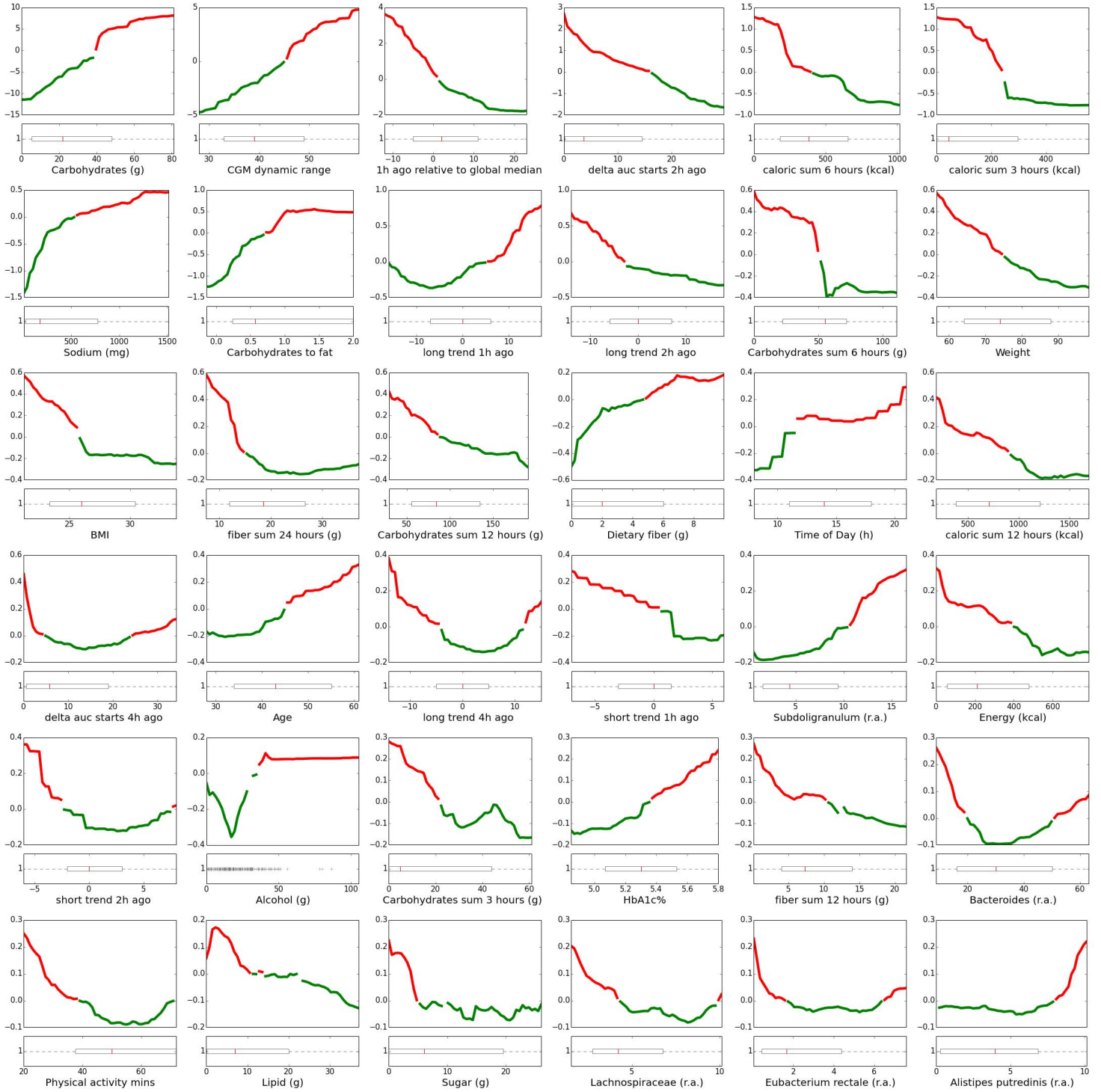


D)



Supplemental Figure 2 – Partial dependence plots showing the marginal contribution of all 72 features (x-axis) to the predicted PPGR (y-axis). Red and green indicate above and below zero contributions, respectively. Boxplots (bottom) indicate the feature distribution across the cohort. Boxes extend from the lower to upper quartile of the data, with a line at the median. Whiskers show the range of the data and flier points ($> Q3 + 1.5(Q3-Q1)$ or $< Q1 - 1.5(Q3 - Q1)$) are considered outliers.

Online Supporting Material



Online Supporting Material

