A picture containing text

Description automatically generated

**Figure S1. The baseline gut microbiota composition at family level. Bars represent individual mice.**

A screenshot of a computer

Description automatically generated with low confidence

**Figure S2. Effects of inulin or resistant starch on (A) body weight, (B) daily food intake, (C) daily energy intake, and (D) 48-hr fecal sample weight of mice receiving diet supplementation used in this study.** Each symbol represents the mean body weight in panel A or a single data point in panels B-D. The body weight data were analyzed by ordinary one-way ANOVA (Analysis of variance) with Turkey post hoc test between inulin or resistant starch and cellulose group. \* *P* < 0.05. The daily energy intake was converted from food intake through XXX

A picture containing text, indoor

Description automatically generated

**Figure S2. Dynamics (A) and time-averaged levels (B) of fecal short-chain fatty acids (SCFAs) concentration following dietary fiber intervention.** In panel A, dots/lines represent mean concentrations across mice within the same vendor and shading areas represent standard error of the mean. In panel B, each colored dot means the time-averaged level of SCFAs (calculated by area under the concentration curve divided by the observation time) in a mouse and gray lines are the best linear regression fit.

A screenshot of a video game

Description automatically generated

**Figure S3. Dynamics of evenness (A), number of observed ASVs (B), and cellulose-group microbiota composition (C) following dietary fiber intervention.** The height of lines (panels A, B) or stacked bands (panel C) represent mean values across mice within the same vendor and shading areas (panels A, B) represent standard error of the mean. In panel C, taxonomic labels w/ “Un.” group bacteria that are unclassified or uncultured at lower taxonomic ranks.

**Metagenome reulsts**

**D**. Gene family profiles at day 0 (baseline), 5 (short-term) and 31 (long-term) following the inulin intervention. R2 and *P*-value were obtained from Adonis analysis. Small dots represent individual mice. The eclipse around the cluster center (large dot) indicates the 95% confidence interval.

A screenshot of a video game

Description automatically generated with medium confidence

**Figure S4.** **Time trajectories of gut microbiota responses in PCoA (principal coordinate analysis) coordinates (A) and temporal changes in the distance of microbiota composition between vendors (B).** Panel A shows responses to resistant starch and panel B shows responses to inulin, resistant starch and cellulose. In panel A, dots represent the mean PCoA coordinate score across mice within the same vendor and error bars represent standard error of the mean. In panel B, dots/lines are the mean pairwise Aitchison distance between samples from different vendors and shading areas represent standard error of the mean.

Graphical user interface, application

Description automatically generated

**Figure S5.** **Bacterial taxa with significant difference in relative abundance between the inulin group and the cellulose group.** Relative abundance changes were calculated between day 0 and day 1 (A), day 0 and day 5 (B), day 0 and day 31 (C). *P*-values were obtained from Wilcoxon rank-sum test after multiple test correction via false discovery rate (FDR) estimation. \*, FDR < 0.05; \*\*, FDR < 0.01; \*\*\*, FDR < 0.001.

Graphical user interface

Description automatically generated

**Figure S6. Analysis of previously published data from Chijiiwa et al.** [33]**.** In this study, the shift in murine gut microbiota was tracked for two weeks following inulin treatment. **A**. Diversity. **B**. Temporal trajectory of gut microbiota in PCoA (principal coordinate analysis) plot. Each dot represents the mean principal coordinate score across all mice and the corresponding error bar represents the standard error of the mean.

A screenshot of a video game

Description automatically generated with medium confidence

**Figure S7. Comparison of temporal changes in relative abundance (A) between inulin responders and other bacteria and (B,C) between different inulin responders.** The dynamical responses were organized by responder in panel B or by vendor in panel C.

A picture containing text, light

Description automatically generated

**Figure S8. Inferring bacterial responders to resistant starch intervention and associated ecological interaction network.** The same figure legend applies as in the main text Fig. 3E-G (the same order).

**Graphical user interface

Description automatically generated with medium confidence**

**Figure S9. Quantification and visualization of baseline-dependent responses of (A) total bacterial density, (B) individual bacterial taxa, and (C) short-chain fatty acids (SCFAs) to resistant starch intervention.** The same figure legend applies as in the main text Fig. 4A-C (the same order).

**Background pattern

Description automatically generated with low confidence**

**Figure S10. Poor performance of Random Forest (RF) regression model in predicting short-chain fatty acids (SCFAs) concentration (see Fig. 5B of the main text for the results) cannot be rescued by using (A) alternative predictors, (B) alternative regression models, and (C,D) weighting of training samples. A**. Prediction accuracy of a RF model trained on different taxonomic- (ASV, Species, Genus, Family) or functional- (Gene, Pathway, Phenotype) predictors. For each taxonomic level, unclassified or uncultured taxa at this level were grouped by the lowest classified rank above this level. The abundances of genes, pathways and phenotypes were predicted using PICRUSt2. **B**. Prediction accuracy of the MelonnPan algorithm [45] trained on the same predictors as used in panel A. **C**. Weights assigned to the training data. The gut microbiota composition of all samples was shown in a reduced two-dimensional UMAP (Uniform Manifold Approximation and Projection) space [49]. The bigger the weights, the larger circle sizes. See Methods in the main text for details of weight calculation. **D**. Prediction accuracy of an RF model built from weighted training data. The absolute abundance of bacterial taxa (grouped by the lowest classified taxonomic level) was used as predictors.

**A picture containing chart

Description automatically generated**

**Figure S11. Prediction of short-chain fatty acid (SCFA) concentration from gut microbiota using data from resistant starch-treated mice**. The same figure legend applies as in the main text Fig. 5B-D (the same order).

**A screenshot of a video game

Description automatically generated**

**Figure S12. Inference of SCFA producers in inulin-treated mice by different approaches** (Random forest regression vs. Repeated correlation analysis)**. A.** Random forest (RF) regression. For each SCFA, we showed top 10 bacterial taxa with highest Gini importance score in model training using all data. The absolute abundances of bacterial taxa were standardized and filtered (threshold 10-5) by LASSO (least absolute shrinkage and selection operator) regression before passing to RF model. Several key hyperparameters in LASSO and RF were optimized using grid search cross-validation with R2 as the score metric. The vendor-level prevalence scores were obtained from Fig. 5C in the main text. **B**. Repeated correlation analysis [50]. Longitudinal data and correlation trend lines are color-coded on a per-mouse basis. Repeated measures correlation coefficients (*r*rm) and FDR-corrected P-values are indicated in the plot.

Graphical user interface

Description automatically generated

**Figure S13. Inference of inulin responders in human gut microbiome. A.** Principal coordinate analysis (PCoA) of baseline human gut microbiota (Bray-Curtis distance matrix of 16S or shallow shotgun metagenomics) in four cohorts of literature studies with inulin intervention. Inulin responders inferred from the four literature studies are showed to the right of the PCoA plot. We used the same generalized Lotka-Volterra model and Bayesian inference framework as we used for analyzing our mouse data (see Methods in the main text for details). Cross (x) marks an exception that the inferred responder can be classified to the species level (Anaerostipes hadrus). Taxonomic labels w/ “Un.” group bacteria that are unclassified or uncultured at lower taxonomic ranks. **B**. Dynamics of relative abundance (rel. abun.) of unclassified Bifidobacterium and unclassified Anaerostipes in our dataset. Lines represent mean concentrations across mice within the same vendor and shading areas represent standard error of the mean.

A screenshot of a video game

Description automatically generated with medium confidence

**Figure S14. Prediction of time-averaged growth response of gut microbiota by the relative abundance of dietary fiber responders in the baseline community**. **A**. Inulin intervention. **B**. Resistant starch intervention. The time averaged growth response is obtained by the area under the curve of total bacterial density divided by the observation time**.** The combination of responders highlighted in red has the highest Pearson correlation coefficient. In both scatter plots, gray lines represent the best fitting line.

A computer screen capture

Description automatically generated with low confidence

**Figure S15**. **The relationship between microbiome and metabolome is time-dependent.** **A**. Dynamics of gut microbiota composition (x-axis) and total SCFA concentration (y-axis) plotted on the same graph. We used the first principal coordinate score from PCoA (principal coordinate analysis) ordination to represent changes in gut microbiota composition (relative abundance) along the direction of maximum variance. Note that SCFAs were substantially produced between day 0 and 1 while gut microbiota composition only changes slightly. Points represent the mean PCoA coordinate score across mice within each vendor and error bars represent the standard error of the mean. **B**. Correlation of baseline unclassified (Un.) Parabacteroides absolute abundance with initial propionate production rates on day 0 (upper panel) and rates in later days (lower panel). Gray line: linear regression.

**Text

Description automatically generated**

**Figure S16. Rarefaction analysis of 16S rRNA amplicon sequencing data.** Rarefaction curves were generated using the iNEXT package [81]. Solid lines represent the observed alpha diversity with the number of reads sampled, and dashed lines represent the extrapolation of the solid lines until 25% more reads. To avoid sample-to-sample bias due to variable sequencing depth (different number of reads per sample), all samples were rarefied to 38,980 sequences (black dashed line) per sample before downstream analysis.

A picture containing text, outdoor object

Description automatically generated

**Figure S17. Reconstructed time series of (A) five inulin responders and (B) short-chain fatty acids by sequential Non-negative matrix factorization.** Dots represent observations. Both lines and dots are color-coded on a per-mouse basis.