Supplementary Information for

Ecological memory of prior nutrient exposure in the human gut microbiome

Jeffrey Letourneau, Zachary C Holmes, Eric P Dallow, Heather K Durand, Sharon Jiang, Verónica M Carrion, Savita K Gupta, Adam C Mincey, Michael J Muehlbauer, James R Bain, and Lawrence A David

Corresponding author: Lawrence A David

Email: lawrence.david@duke.edu

This PDF file includes:

Figures S1 to S12

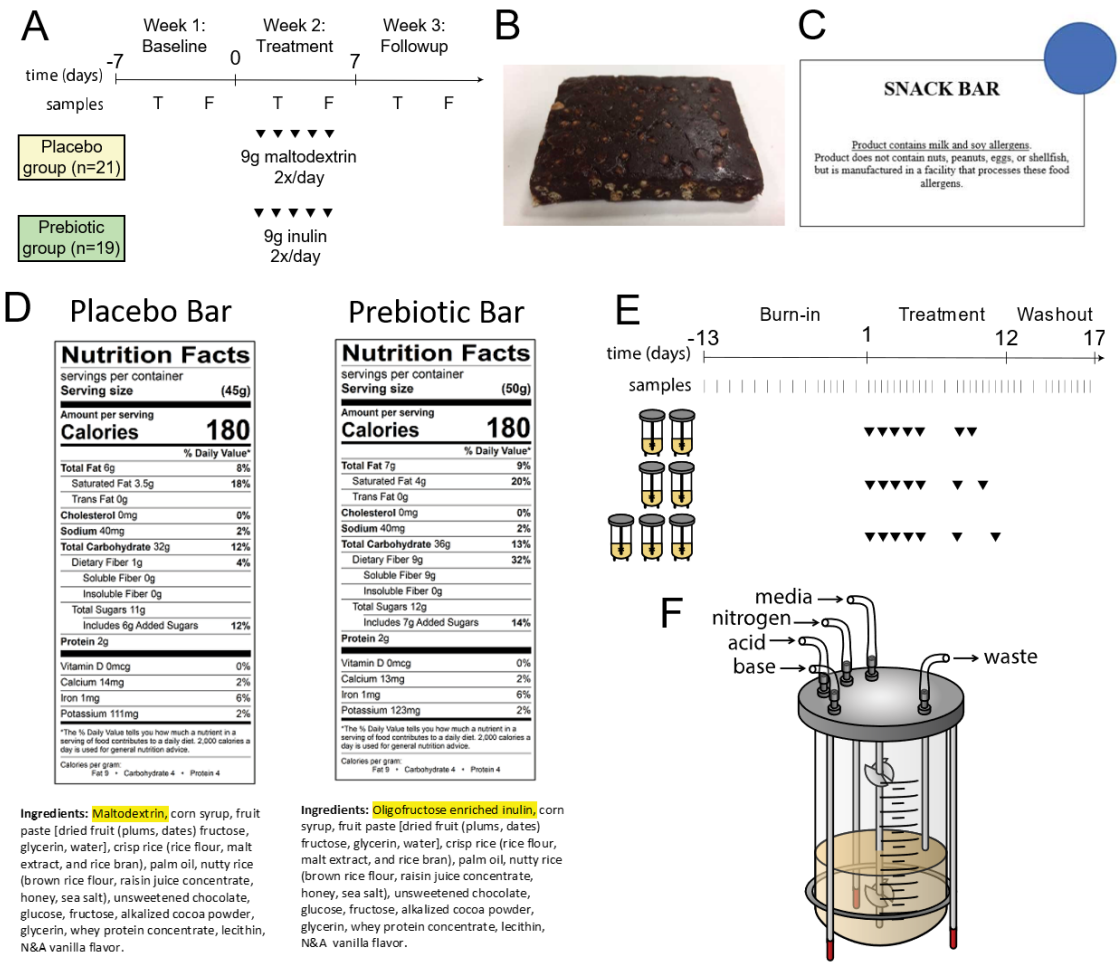

Figure S1. Study design.

A, Human participant study design. Participants consumed snack bars twice daily, which contained either 9 g inulin (prebiotic group) or 9 g maltodextrin (placebo group), and provided stool samples twice weekly, on Tuesdays and Fridays. **B**, Photo of snack bar consumed by research participants. **C**, Approximate label used for snack bars. A colored dot, either red or blue, was coded to the identity of the bar, to which participants were blinded. **D**, Nutrition facts and ingredients for the placebo and prebiotic bars. E, Schematic of artificial gut run. After a 13-day burn-in period, discrete inulin dosing (denoted by ▼) was started in the seven artificial guts used for the experiment. Samples were taken once or twice daily, denoted by |. **f**, Diagram of the “artificial gut”, a continuous-flow bioreactor.

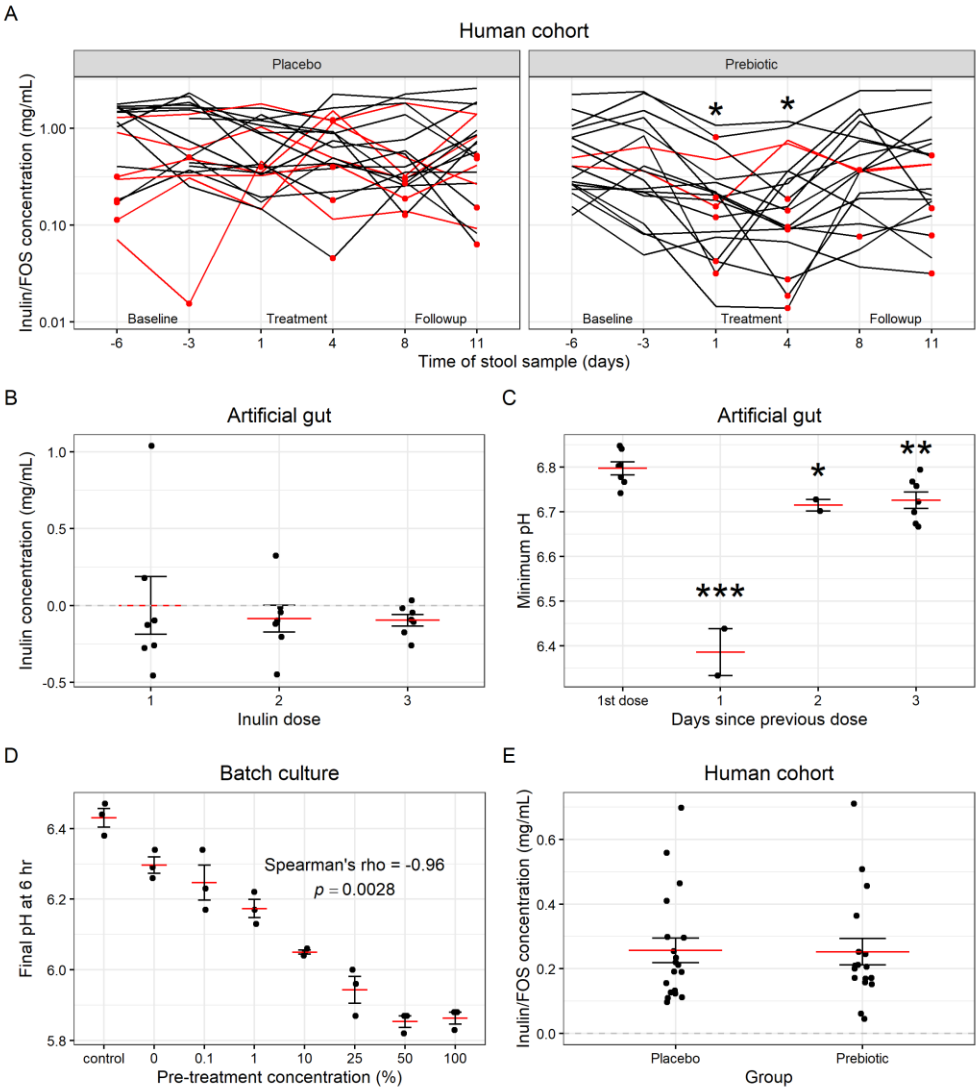

Figure S2. Enhancement of inulin metabolism within 24 hours of first dose.

A, Inulin/FOS concentrations (DP 3+) after incubation of slurried stool samples with inulin for 24 hours *ex vivo* on a log scale, with lines connecting samples from the same participant. Red lines indicate participants for which there was an average increase between Baseline and treatment values; red dots indicate the minimum value observed for each participant. (placebo group $n = 21$; prebiotic group $n = 19$). **B**, Concentration of inulin remaining in each artificial gut vessel 2 hours before dosing. Prior to dose 1, there had previously been no inulin in the system. ($n = 7$ artificial gut vessels.) **C**, Minimum pH reached over the 24 hours following each inulin dose, plotted by time since previous dose. (Mixed effects linear model with 1st dose as intercept; $n = 7$, 2, 2, and 7 vessels.) **D**, Final pH after 6 hours incubation on inulin, preceded by pre-treatment with an inulin dose of varying concentration. ($n = 3$ cultures; for Spearman correlation, average values used, excluding control, so $n = 7$ conditions.) **E**, Baseline fecal inulin/FOS concentrations at day -3 as measured by HPAEC-PAD. **B**-**E**, Mean and standard error plotted. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

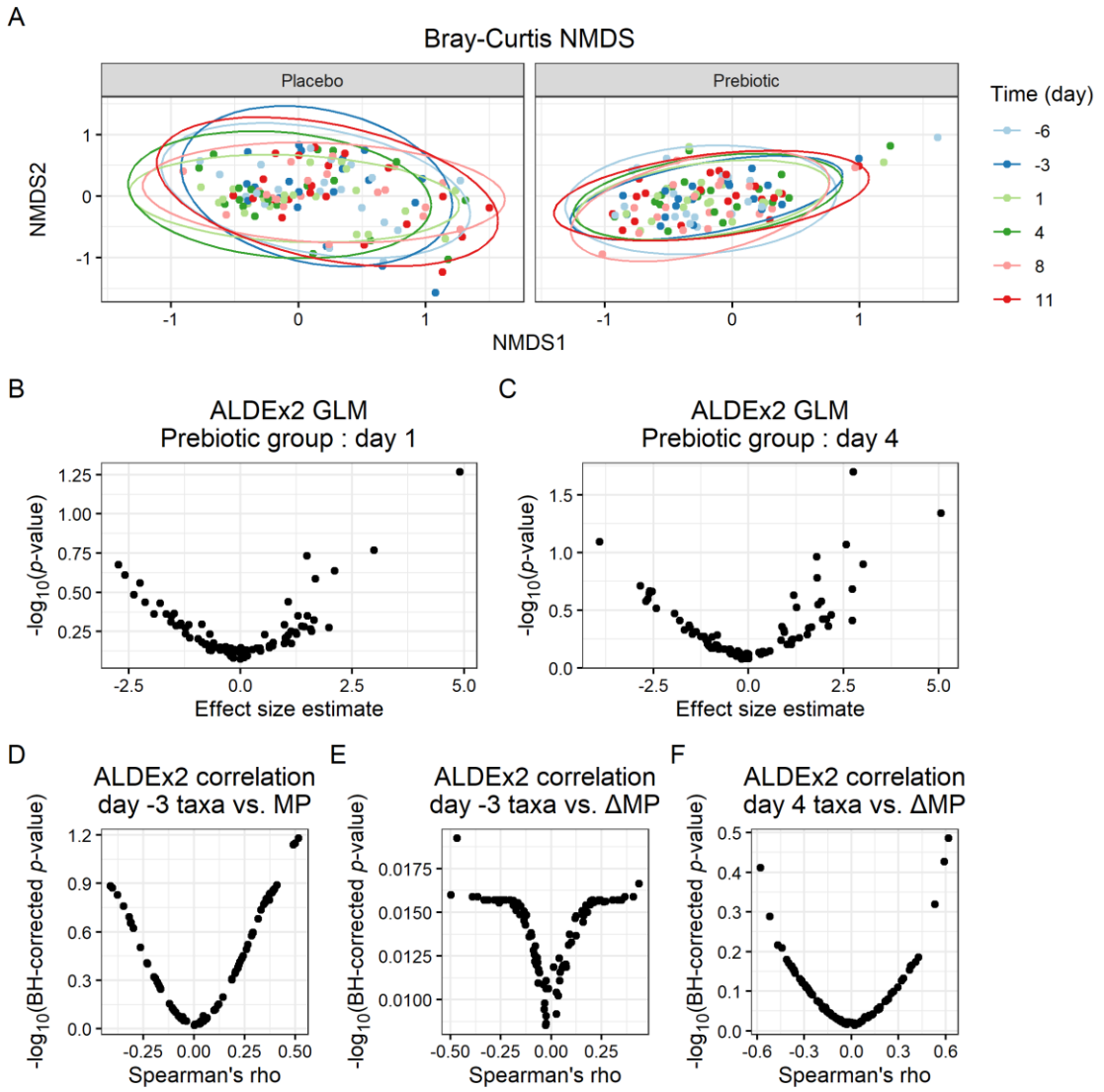

Figure S3. Human cohort 16S gene amplicon sequence data.

A, Bray-Curtis NMDS plot of stool samples, colored by time point. All terms for a PERMANOVA test of group*time with participants as strata had $p > 0.05$. (n = 21 participants in placebo group, 19 in prebiotic group.) **B-C**, Volcano plots for an ALDEx2 GLM model of the group*time interaction, showing effect size estimates and (uncorrected) p-values for the two model terms associated with treatment week. For all taxa in both interaction terms, BH-corrected $p > 0.05$. **D-F**, Volcano plots depicting Spearman's rho and BH-corrected p-values for three separate ALDEx2 correlation test models investigating the relationship between metabolic potential (MP) and taxonomy. For (**D**), the day -3 metabolic potential, as inulin/FOS content remaining at the end of *ex vivo* culture, was supplied as a continuous variable. For (**E-F**)**,** the change in metabolic potential was calculated as the ratio of the treatment week mean to the baseline week mean.

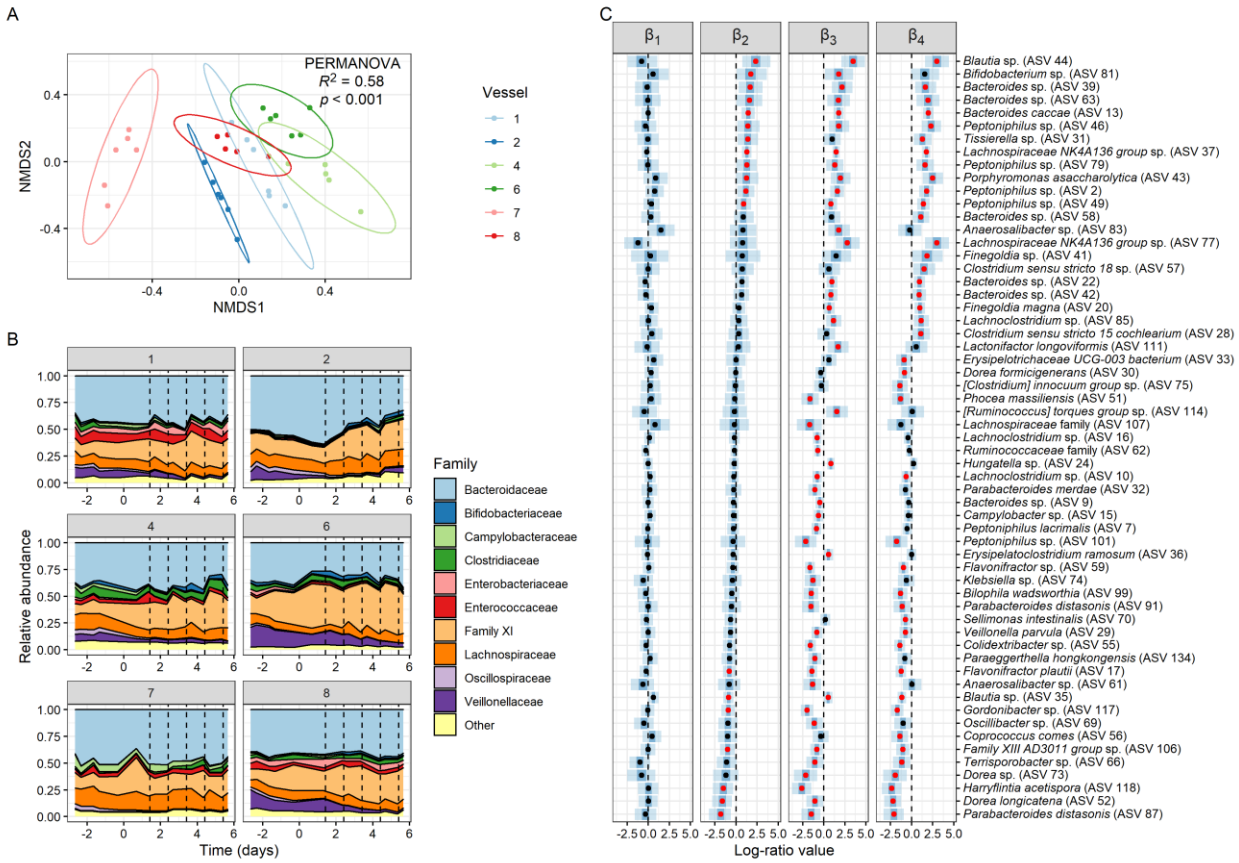

Figure S4. Changes in artificial gut community composition.

A, Bray-Curtis NMDS plot of artificial gut vessels during the week prior to inulin dosing. **B**, Relative abundances of the top ten families over time in each vessel. Dashed lines indicate times of inulin doses. **C**, Bayesian multinomial logistic-normal model parameter estimates for all ASVs where the 95% credible interval for at least one parameter excludes zero (indicated by dashed line), sorted by β_2 values. Light blue indicates the 95% credible interval, dark blue indicates the 50% credible interval, dots are mean values, and red dots indicate that the 95% credible interval excludes zero. The parameter β_1 indicates changes on afternoon 1 (dose 1 hour +6), β_2 indicates changes on morning 1 (dose 2 hour -2), β_3 indicates changes subsequent hour +6 samples, and β_4 indicates changes on subsequent hour -2 samples. (n = 6 vessels.)

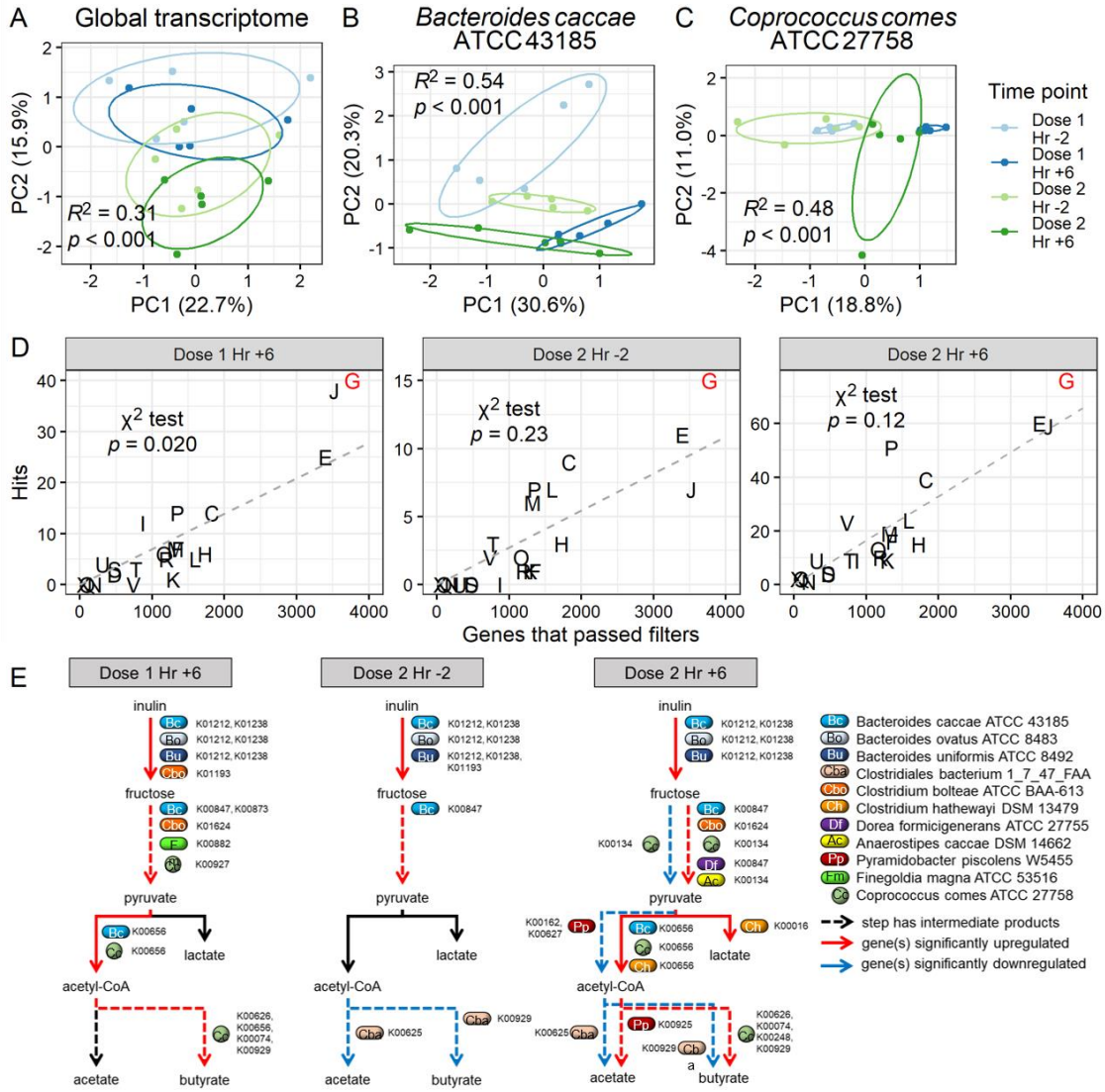

Figure S5. Changes in artificial gut metatranscriptome following inulin exposure.

A-C, PCA plots of global transcriptome (**A**) and individual transcriptomes of select taxa, a primary inulin degrader (**B**) and a degrader of secondary metabolites (**C**), by time point. PERMANOVA R^2 and *p*-values shown. **D**, Representation of COG categories in transcripts found to be significantly differentially expressed by ALDEx2 GLM at each time point in the within-taxon analysis. The dashed line describes where all COG categories would lie if they were represented in the hits equally to their proportion in the overall gene set. Chi-squared test for enrichment of category G (carbohydrate metabolism and transport) shown. (*n* = 5 vessels.) **E**, Simplified pathway diagram depicting the conversion of inulin to relevant secondary metabolites by time point. Taxa found to significantly (Benjamini-Hochberg adjusted $p < 0.05$) up- or downregulate genes encoding enzymes carrying out steps in this pathway shown. (ALDEx2 GLM; *n* = 5 vessels.)

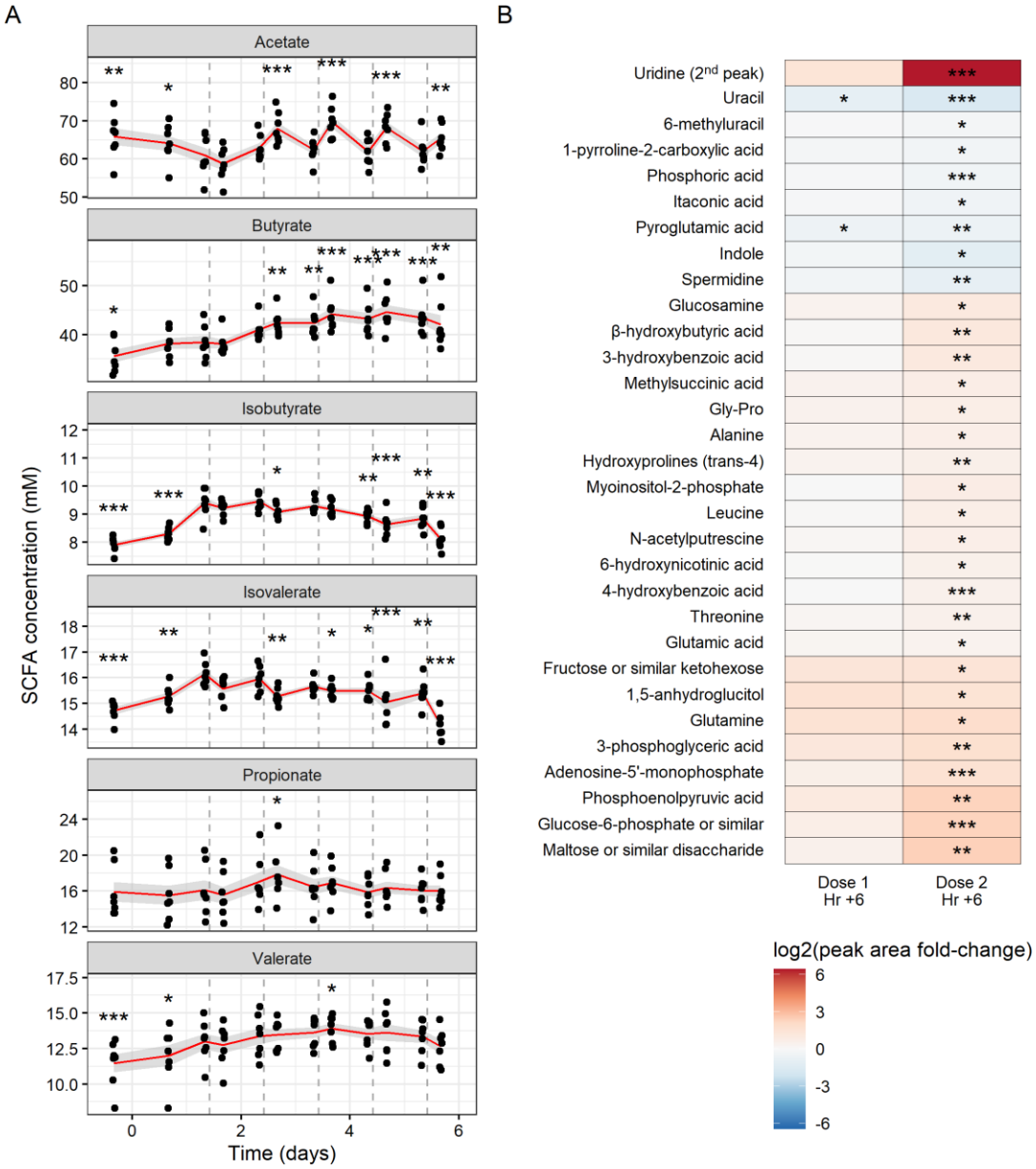

Figure S6. Changes in artificial gut metabolome.

A, All measured SCFAs in the artificial gut during dosing week and the two days prior. (Mixed-effects linear model with dose 1 hour -2 (day 1.42 on x-axis) as intercept; n = 7 vessels.) Dashed lines represent times of inulin dosing. Mean and standard error plotted. **b**, GC/MS metabolites found to be significantly different from baseline by mixed-effects linear model in the artificial gut (n = 6 vessels). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

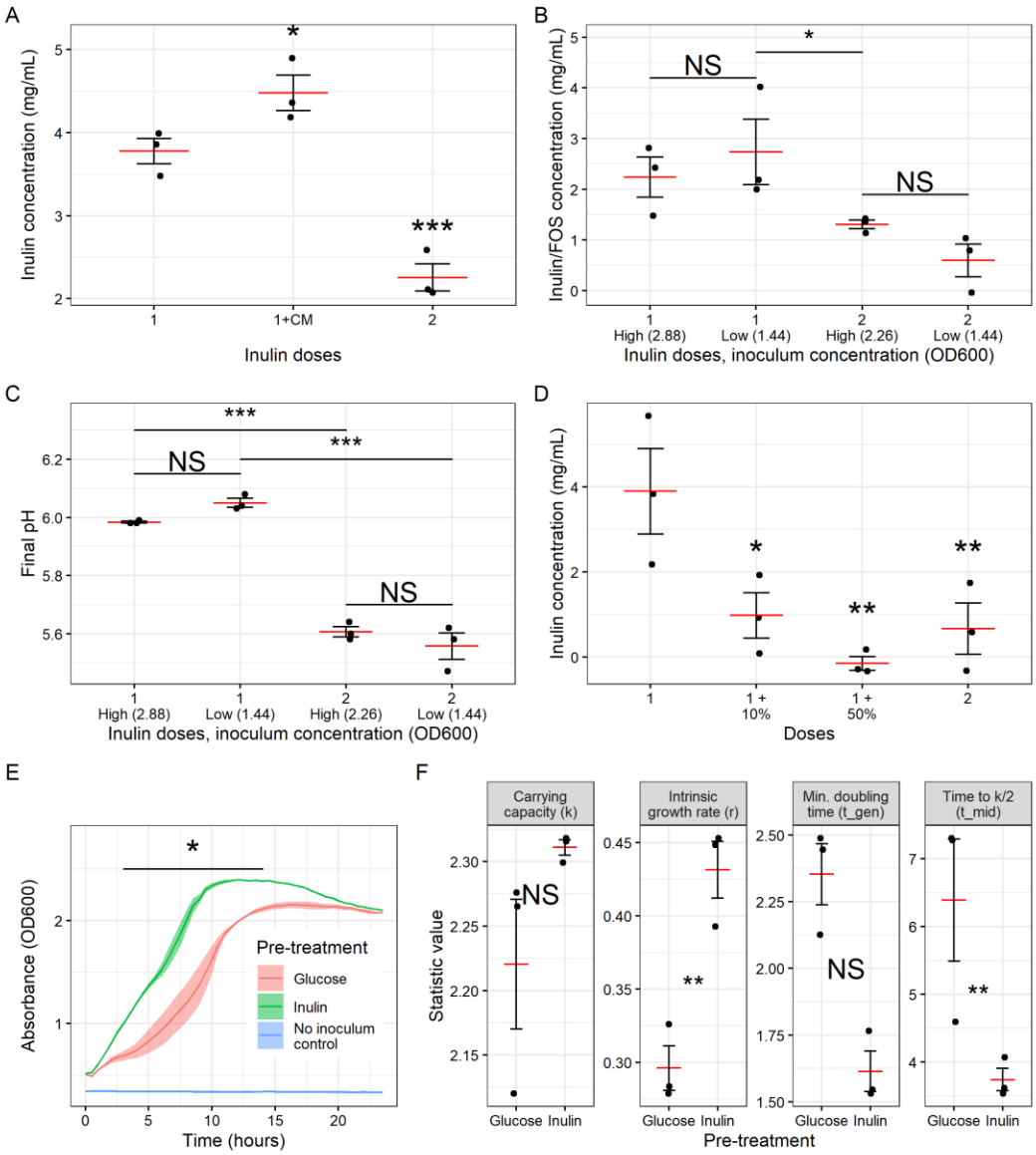

Figure S7. Factors driving ecological memory.

A, Conditioned media (CM) from inulin-exposed cultures did not increase metabolism of inulin in treatment-naïve cultures (linear model; $n = 3$ batch cultures). **B-C**, Effect of starting cell concentration on inulin/FOS breakdown (**B**) and final pH (**C**). (Tukey test; $n = 3$ cultures.) **D**, Effects of 10 or 50% *Bacteroides caccae* addition to complex community cultures. (Linear model; $n = 3$ batch cultures). **E**, Effects of glucose or inulin pre-treatment on growth curves of *B. caccae* monocultures grown on inulin in conditioned media (from inulin-naïve cultures). The no inoculum control well had only mGAM media and serves as a control for contamination. (Linear model; interaction term of time and pre-treatment $p < 0.05$ for all points below the line; $n = 3$ batch cultures). **F**, Metrics from growth curves shown in (**E**) calculated by the growthcurver R package. (Two-sided unpaired t-tests; $n = 3$ cultures.) Mean and standard error shown. **d**, NS $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

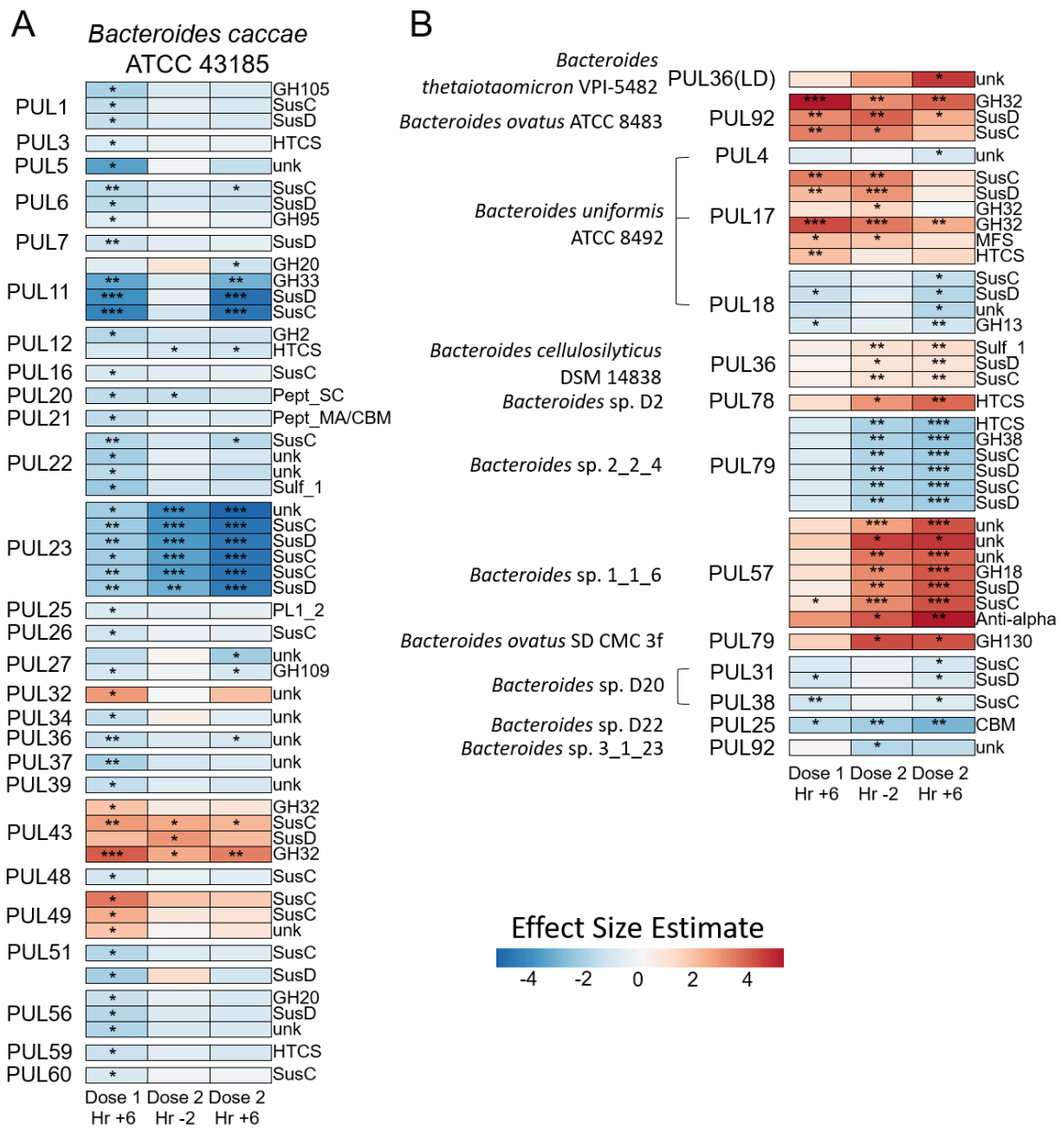

Figure S8. Activation of PULs in *Bacteroides*.

A-B, Polysaccharide utilization loci (PULs) in *Bacteroides caccae* ATCC 43185 (**A**) and all other *Bacteroides* species (**B**) for which at least one gene was significantly differentially expressed following inulin treatment. Conserved functional prediction from PULDB shown. (ALDEx2 GLM; n = 5 vessels.) * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

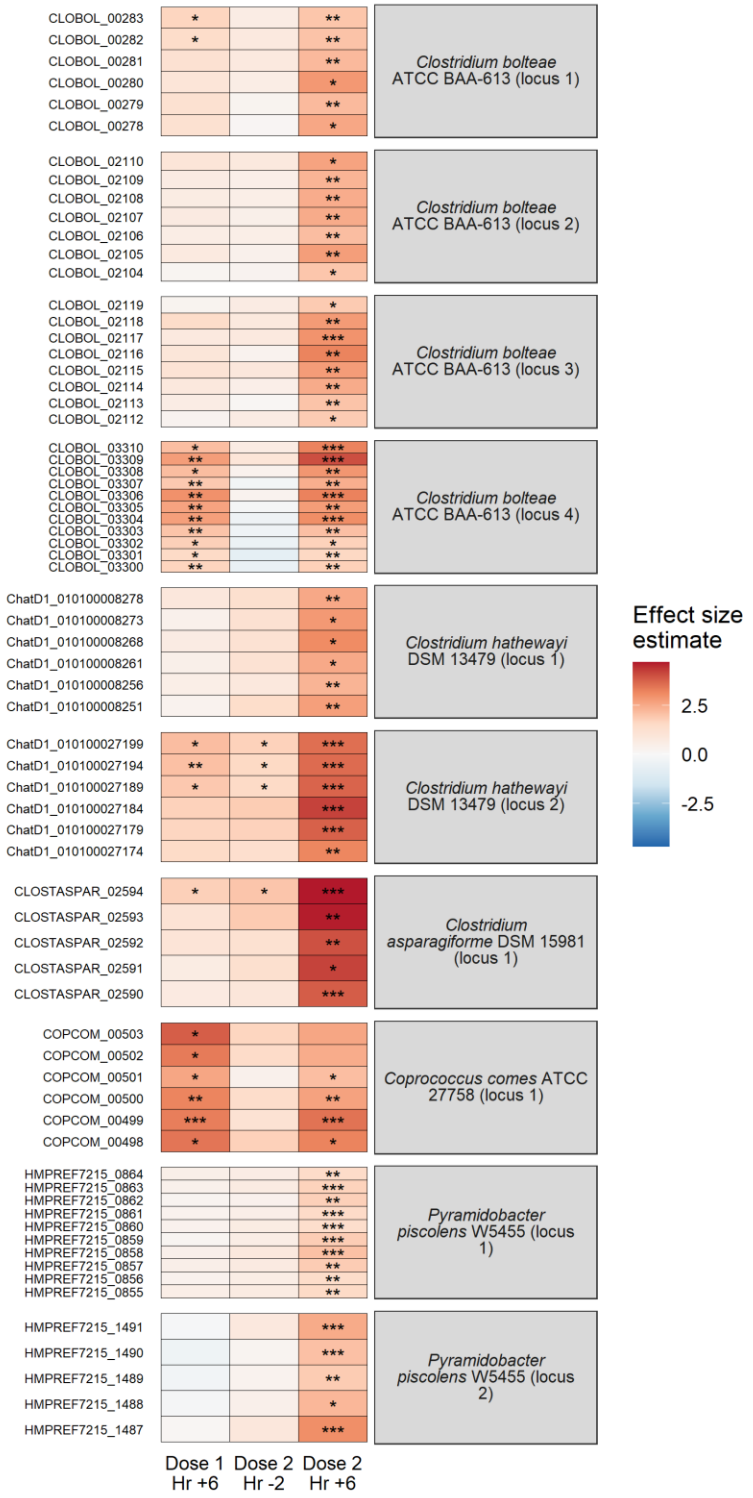

Figure S9. Activation of putative PUL-like operons in non-*Bacteroides* taxa.
Heatmap of genes comprising putative operons (sets of adjacent genes) with significantly increased expression on inulin treatment by ALDEx2 GLM.

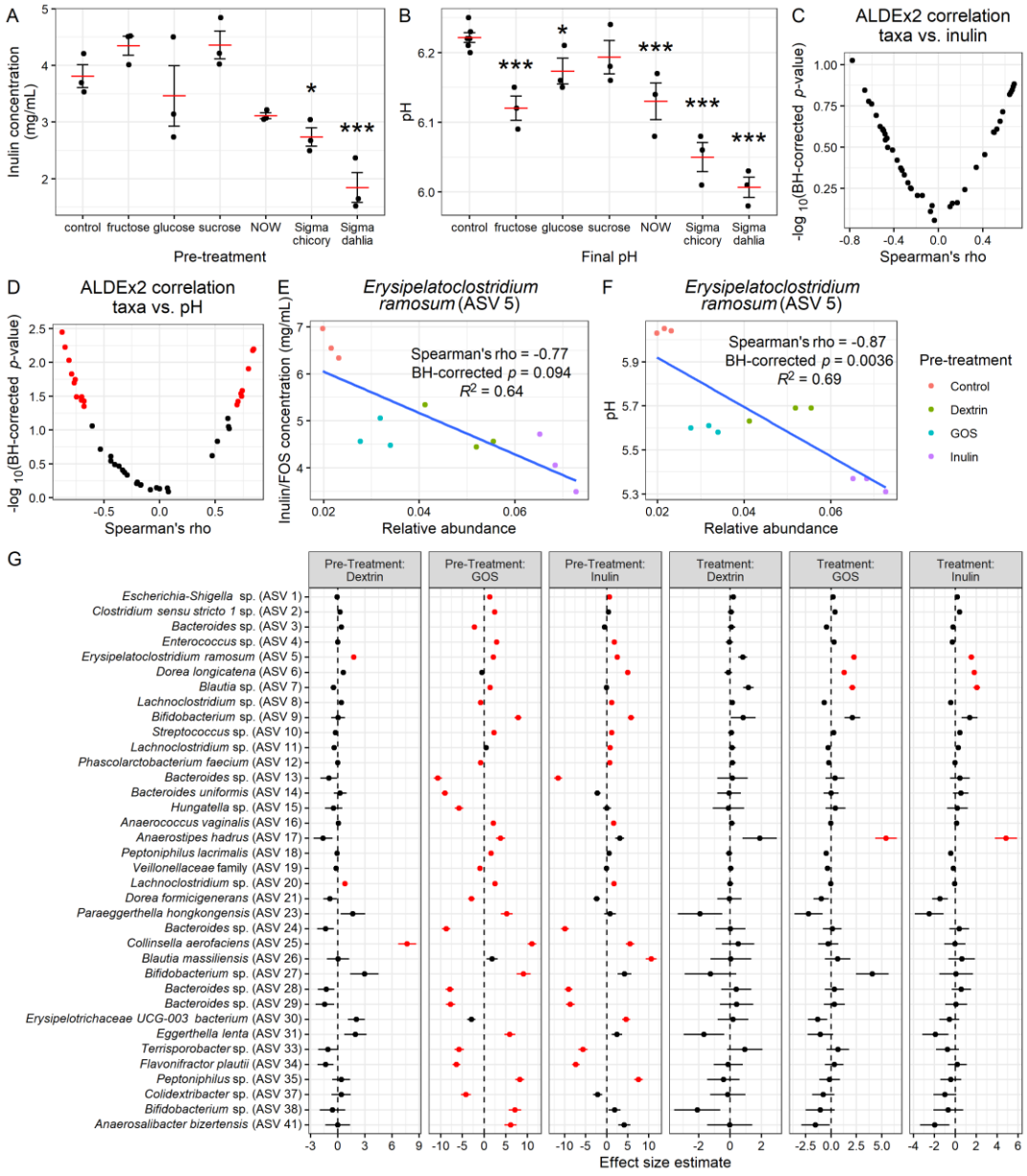

Figure S10. Supplementary data on cross-reactivity of ecological memory.

A-B, Final inulin concentration (**A**) and pH (**B**) when pre-treated with inulin of various chain length or component mono/di-saccharides. (Linear models with control as intercept; n = 3 cultures.) Mean and standard error shown. **C-G**, Statistical results for the prebiotic cross-reactivity experiment shown in Fig. 4D-F. **C-D**, Volcano plots of ALDEx2 Spearman correlation tests between ASVs and final inulin/FOS concentration (**C**) or pH (**D**). Red points indicate post-Benjamini-Hochberg-corrected $p < 0.05$ for that taxon. **E-F**, Correlations between an ASV mapping to *E. ramosum* and final inulin/FOS concentration (**E**) or pH (**F**). This was the top correlated taxon for both tests. Relative abundance is shown here for interpretability, but CLR-transformed counts were used for the correlation test. **G**, ALDEx2 GLM results for all ASVs, testing the effect of first and second treatments with three prebiotics. Mean effect size estimate and standard error shown. Red indicates a statistically significant ($p < 0.05$) result for a given ASV-covariate combination. Note that *E. ramosum* and two other ASVs had positive and significant values for pre-treatment with all three prebiotics. (n = 3 cultures.) * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

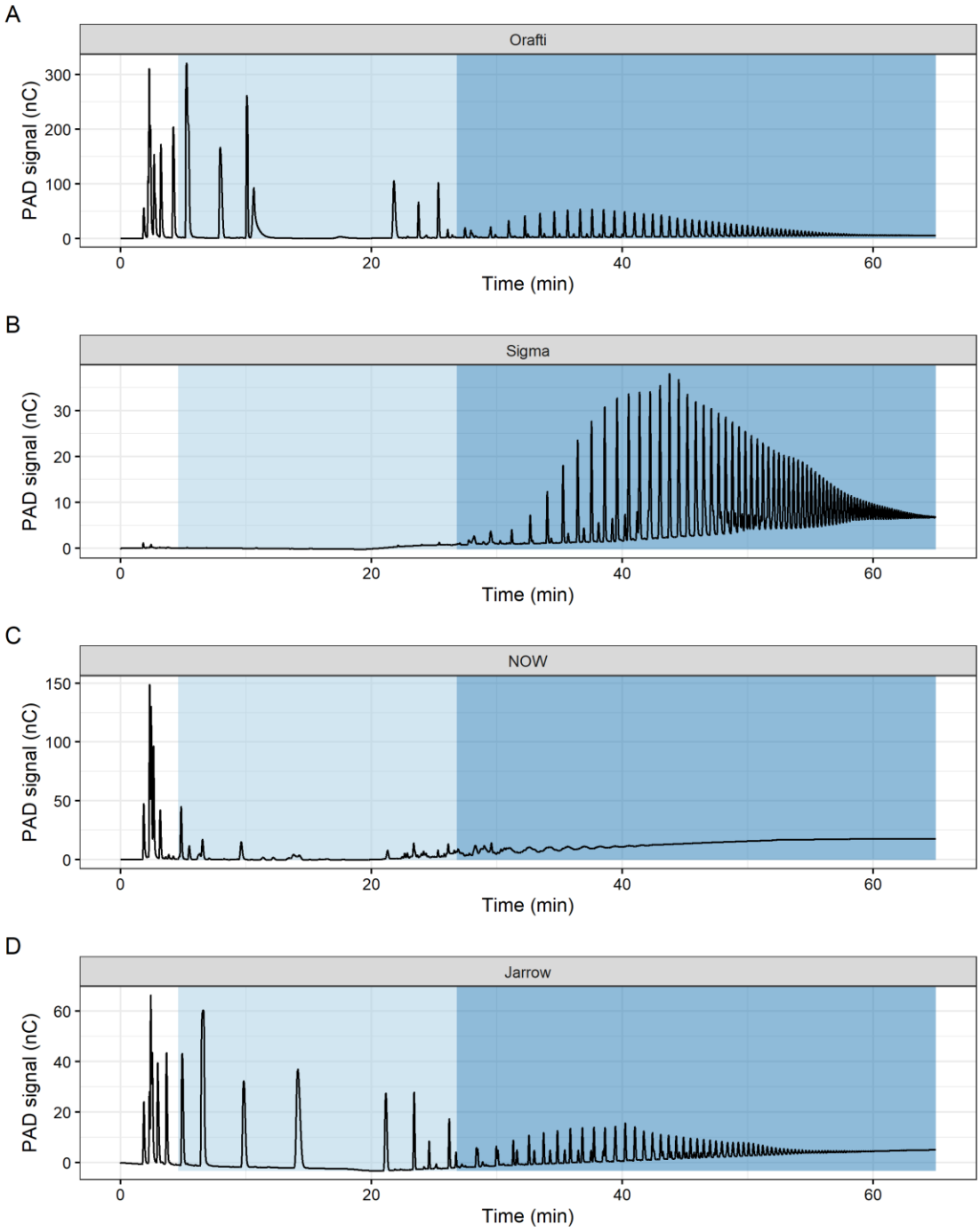

Figure S11. Spectra of primary inulin sources used in this study.

Raw HPAEC-PAD spectra for inulin samples. Light blue shading denotes the approximate region containing FOS peaks (DP 3-10), and dark blue denotes inulin peaks (DP 11+). **A,** Orafti Synergy1 inulin, a mixture of short- and long-chain inulin, used for the human study. **B**, Sigma inulin from dahlia tubers, a very pure source of long-chain inulins, used for most *in vitro* work including the artificial gut experiment. **C**, NOW inulin, a source with mostly shorter chain FOS. **D**, Jarrow inulin, another mixed chain length inulin.

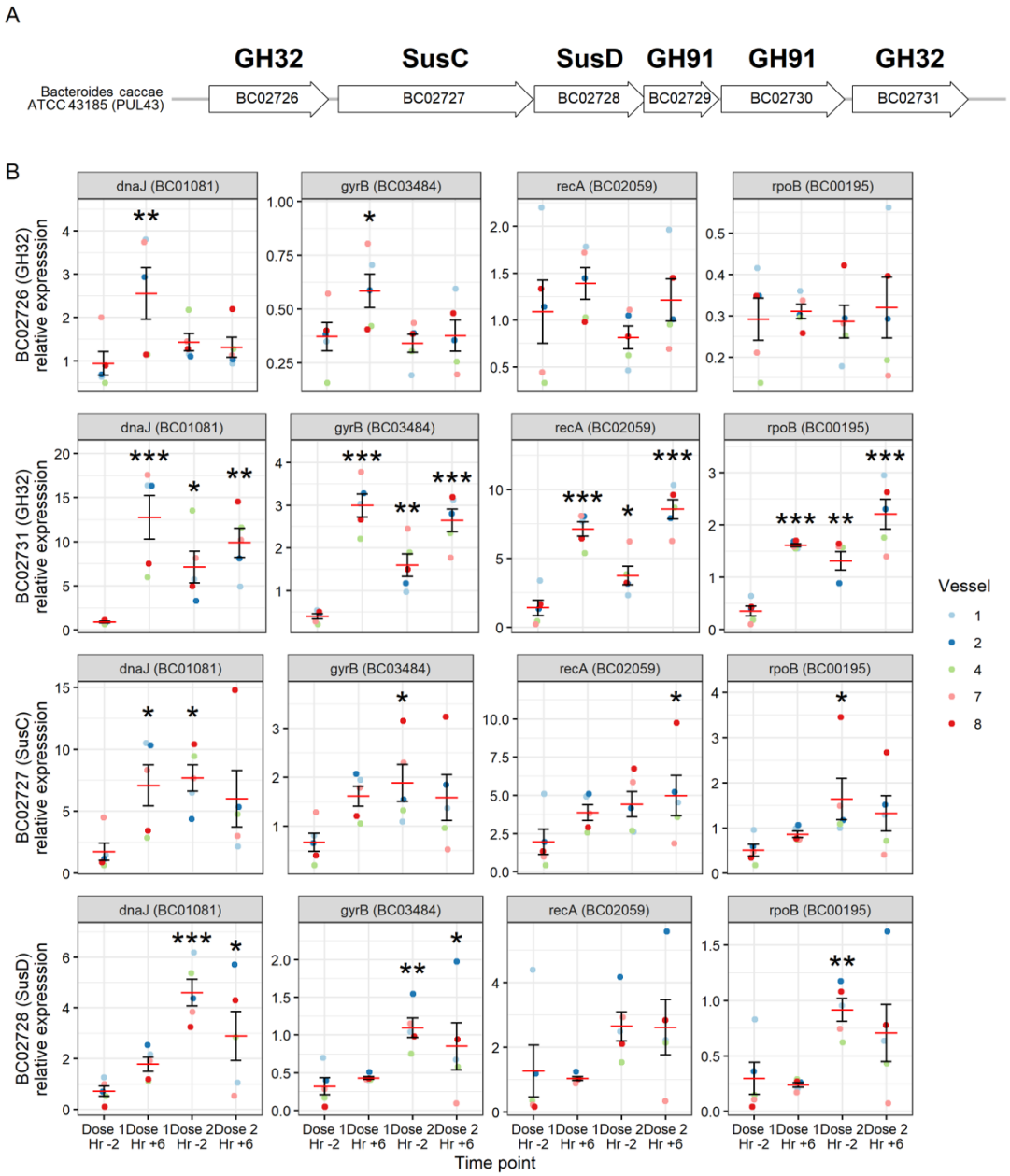

Figure S12. Confirmation of changes to *B. caccae* PUL43 gene transcripts by comparison to housekeeping genes.
A, Gene diagram depicting the *Bacteroides caccae* PUL of interest shown in Fig. 3a, with gene names and functional categories. **B**, Expression levels of the four *B. caccae* genes in PUL43 found to be significantly upregulated following inulin exposure in the artificial guts (Figs. 2C, 3A), reanalyzed by computing the ratios of these transcripts to each of four housekeeping genes: dnaJ (BC01081), gyrB (BC03484), recA (BC02059), and rpoB (BC00195). The top two rows shown the GH32 genes, and the bottom two rows show the SusC/SusD components. (Mixed-effects GLM with Dose 1 Hr -2 as intercept; n = 5 artificial gut vessels.) * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.