Supplementary data for "Gut Microbiome-Host Homeostasis upon exposure to PFOS and GenX in male mice

Faizan Rashid [1,2,†] , Veronika Dubinkina [3, 4,†] , Saeed Ahmad [3], Sergei Maslov [3, 4] , and Joseph Maria Kumar Irudayaraj [1,2,3,5,6*]

[1]Biomedical Research Center in Mills Breast Cancer Institute, Carle Foundation Hospital, Urbana, IL 61801, USA.

[2]Department of Comparative Biosciences, College of Veterinary Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

[3]Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

[4]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

[5]Micro and Nanotechnology Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

[6] Cancer Center at Illinois, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

* Correspondence: jirudaya@illinois.edu; Tel.: +1-765-404-0499

† Equal credit co-first authors.
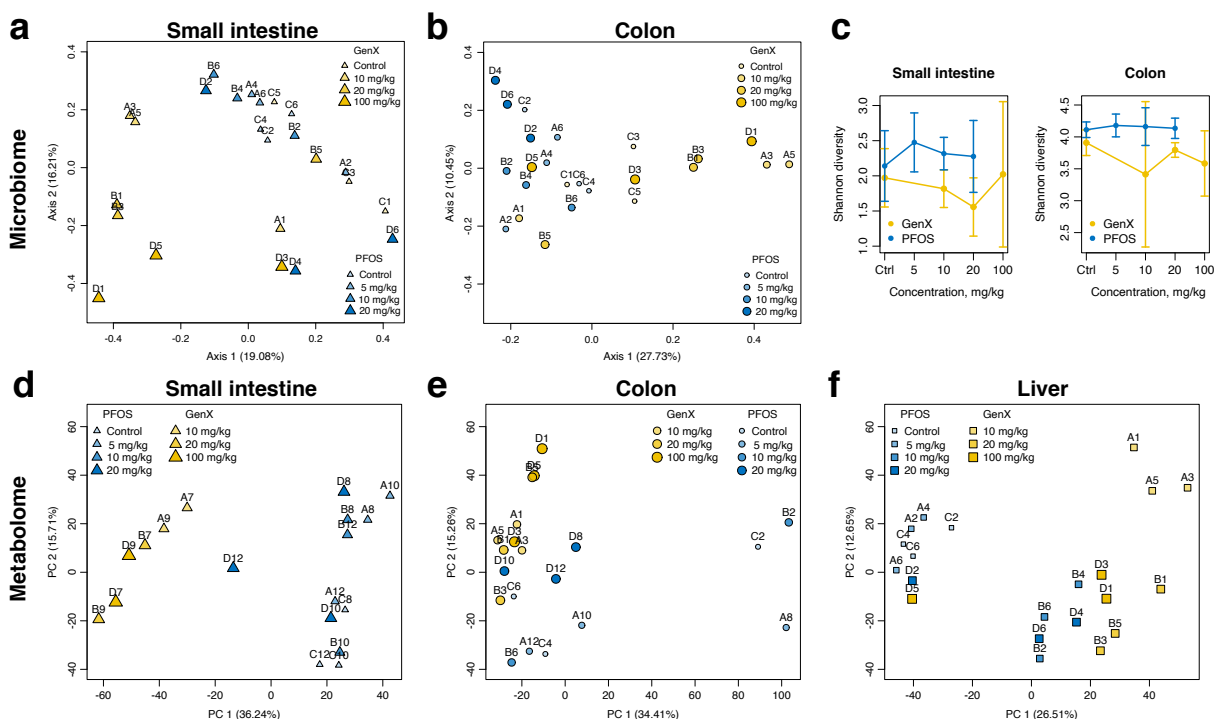
Supplementary Figures

Figure S1. Beta-diversity plots for multi-comics data collected in the study. (a-b) Principal coordinates analysis (PCoA) based on Bray-Curtis dissimilarity of 16S rRNA samples. PCoA for small intestine (a) and colon (b). Blue points correspond to PFOS, yellow points to GenX, circles and triangles represent colon and small intestine respectively. The size and color intensity of the points correspond to the concentration of chemicals in the diet. We encoded animals as A1, A2, etc., and indicated respective samples on each plot. (c) Shannon diversity index for 16S rRNA data collected for different compound concentrations. (d-f) Principal component analysis (PCA) for all metabolomics variables. (d) small intestine, (e) colon and (f) for liver. Percentages on axes show the proportion of variance explained.
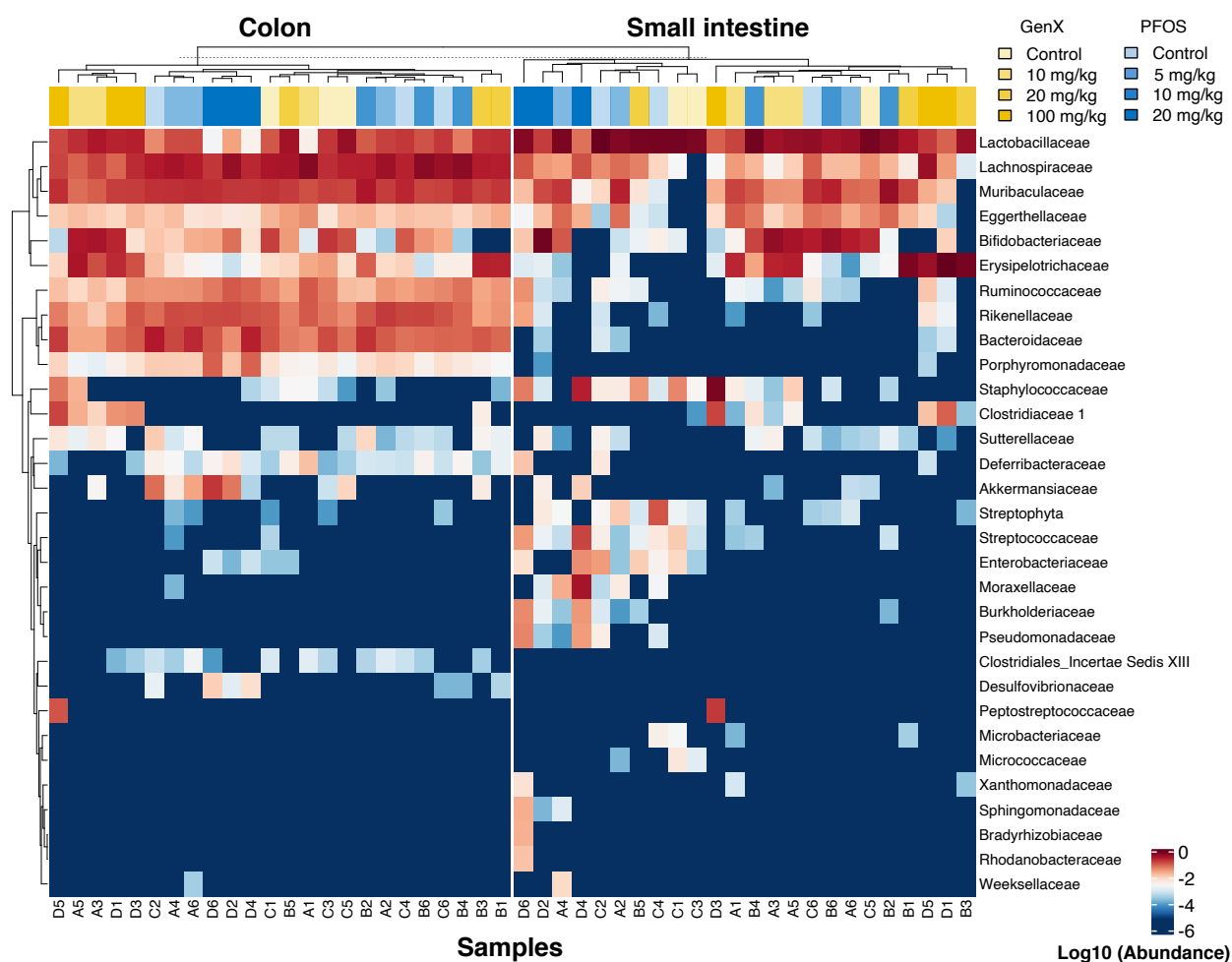
Figure S2. Heatmap of the taxonomic composition of all 16S samples on family level (according to rdp classification). The columns correspond to the samples; compound concentration in the diet is denoted with a top color bar. Figure shows the taxa with a total relative abundance of ≥ 0.5% across all samples. Hierarchical clustering was performed using the Euclidean metric and complete linkage.

Figure S3. Changes in genus levels of 16S samples for small intestine and PFOS diets. Shown are genera with >80% confidence in rdp classification, >0.5% total abundance across all small intestine samples, and significantly non-zero slope of linear regression.

Figure S4. Changes in genus levels of 16S samples for small intestine and GenX diets. Shown are genera with >80% confidence in rdp classification, >0.5% total abundance across all small intestine samples, and significantly non-zero slope of linear regression.

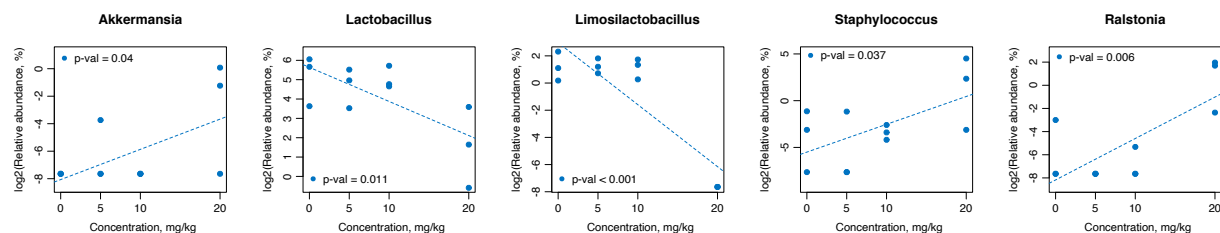**Figure S5.** Changes in genus levels of 16S samples for colon and PFOS diets. Shown are genera with >80% confidence in rdp classification, >0.5% total abundance across all colon samples, and and significantly non-zero slope of linear regression.

Figure S6. Changes in genus levels of 16S samples for colon and GenX diets. Shown are genera with >80% confidence in rdp classification, >0.5% total abundance across all colon samples, and significantly non-zero slope of linear regression.
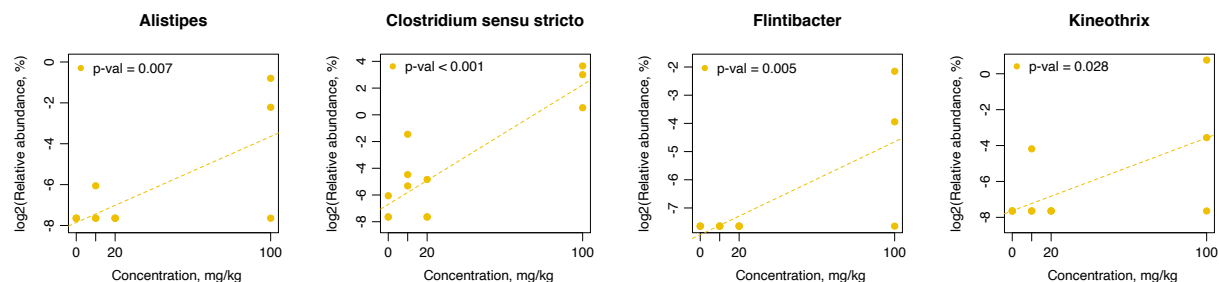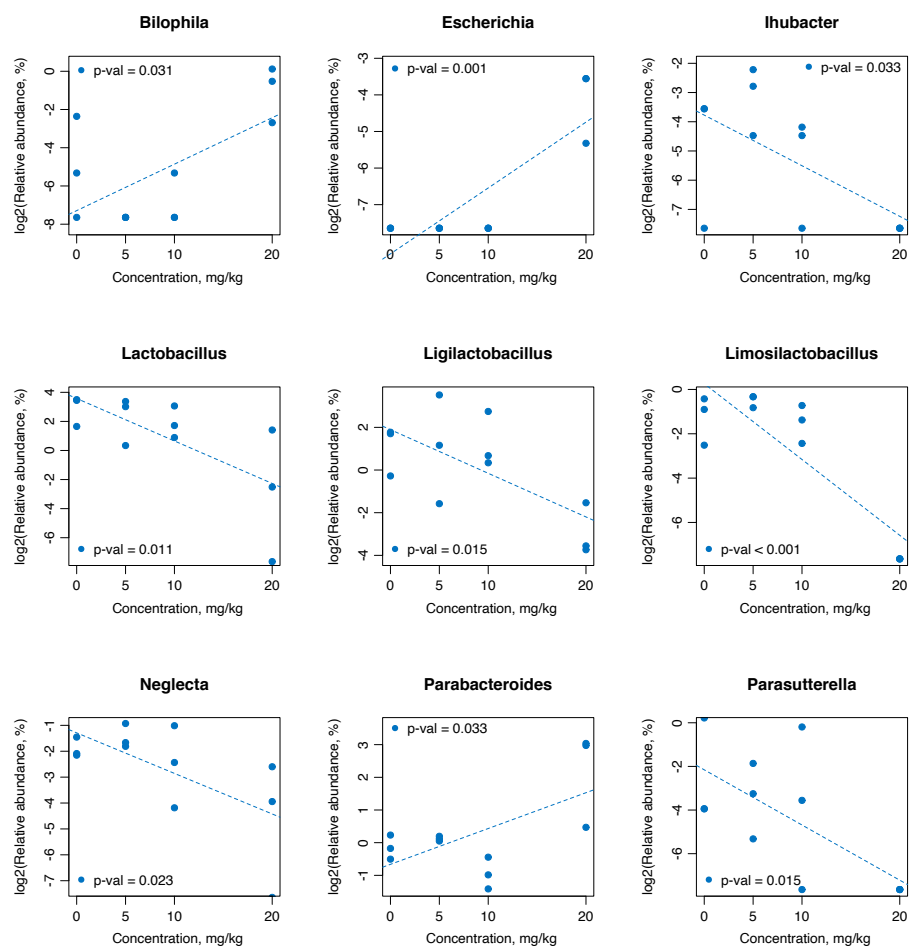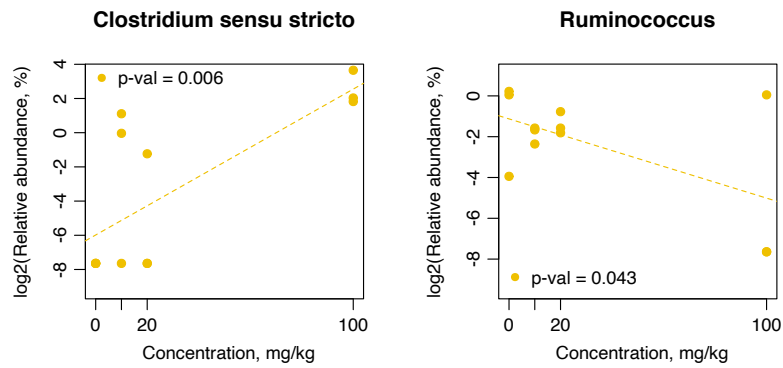

Figure S7. Changes in phylum levels of 16S samples for the small intestine (a) and colon (b) (>80% confidence in rdp classification). No significant changes were noted in these.

Figure S8. Small intestine metabolites significantly vary between different diets. (a) Heatmap of metabolites selected by ANOVA test (p-value<0.0001), peaks with low interquartile variability were filtered out before this analysis. Only metabolites with assigned potential chemical structures are shown (38/383). Color represents z-score value of log10 normalized metabolite concentrations. (b) Upregulated and downregulated pathways based on metabolite enrichment analysis in the small intestine. Only metabolites that have a significant association with the concentration of the toxic compounds (p-value of linear regression model < 0.05) were used for this analysis. Blue bars indicate pathways enriched in PFOS diets, red - in GenX, color intensity shows the number of metabolites present in the enriched/lowered metabolite set that belong to a given pathway. Pathway fold enrichment is calculated by the QEA analysis. Only ones with a non-zero influence score or multiple hits and FDR < 0.05 are shown. Negative fold enrichment values indicate a set of pathways containing downregulated metabolites, positive - elevated ones. (c) Metabolites that systematically decrease with the concentration of toxic compounds. (d) Metabolites that are systematically elevated in GenX samples, but constant in PFOS. (e) Metabolites are constant in PFOS but systematically decreased in GenX. (f) Metabolites with non-linear response to the concentration of toxic compounds.

Figure S9. Colon metabolites significantly vary between different diets. (a) Heatmap of metabolites selected by ANOVA test (p-value<0.01), peaks with low interquartile variability were filtered out before this analysis. Only metabolites with assigned potential chemical structures are shown (36/137). Color 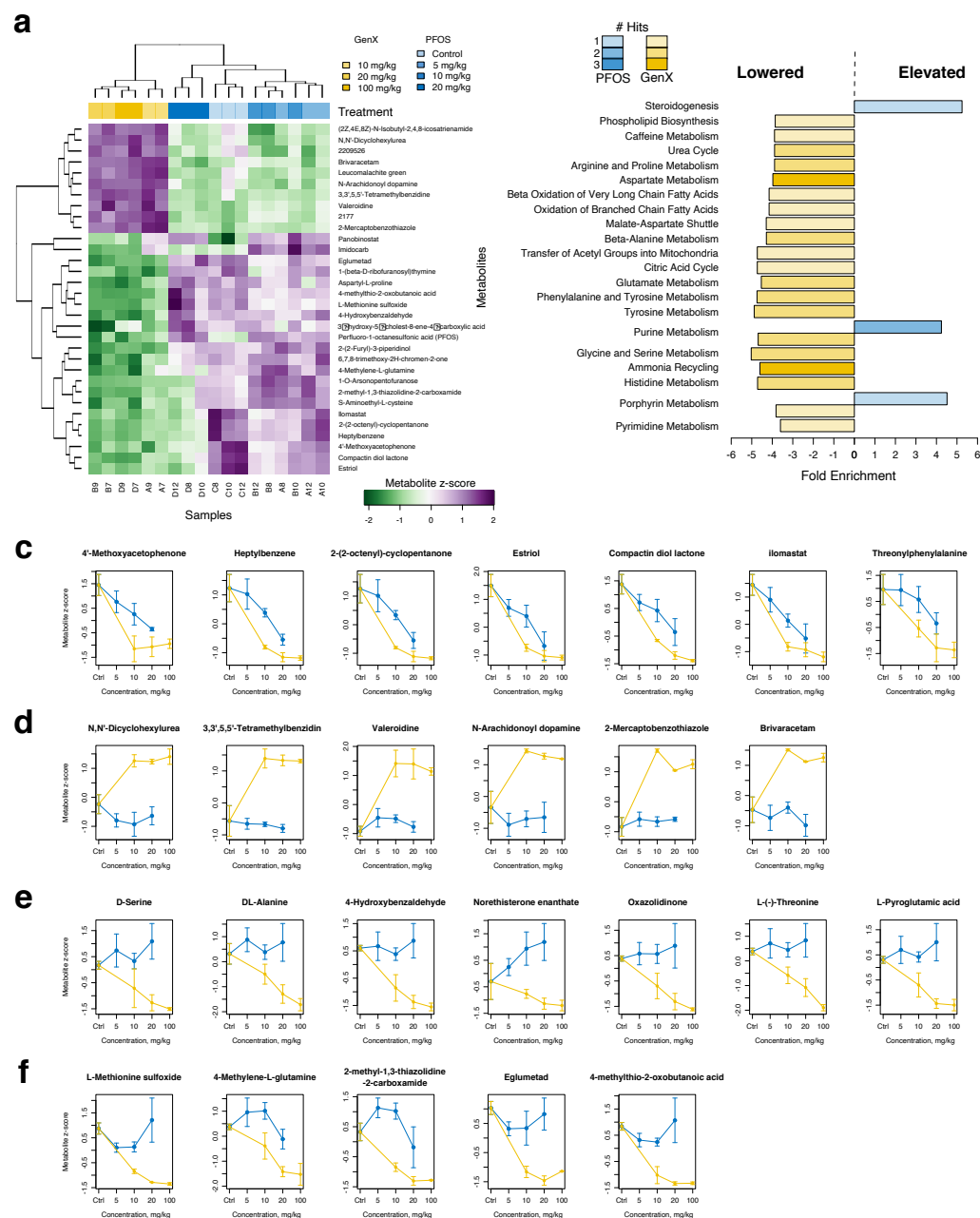represents z-score value of log10 normalized metabolite concentrations. (b) Selected metabolites that show some systematic trends with the concentration of toxic compounds. (c) Metabolites that were significantly elevated/lowered both in PFOS and GenX diets (p-value of linear regression model < 0.05). (d) Upregulated and downregulated pathways based on metabolite enrichment analysis in colon. Only metabolites that have a significant association with the concentration of toxic compounds (p-value of linear regression model < 0.05) were used for this analysis. Blue bars indicate pathways enriched in PFOS diets, red denotes GenX, color intensity shows the number of metabolites present in the enriched/lowered metabolite set that belong to a given pathway. Pathway fold enrichment is calculated by the QEA analysis. Only ones with a non-zero influence score or multiple hits and FDR < 0.05 are shown. Negative fold enrichment values indicate a set of pathways containing downregulated metabolites, positive - elevated ones.

L-glutamate and L-glutamine biosynthesis
purine nucleotides degradation II (aerobic)
D-fructuronate degradation
adenosine nucleotides degradation II
adenosylcobalamin salvage from cobinamide I
adenosylcobalamin salvage from cobinamide II
adenosylcobalamin biosynthesis from cobyrinate a,c-diamide I
guanosine nucleotides degradation III
purine nucleobases degradation I (anaerobic)
thiazole biosynthesis I (E. coli)
superpathway of Clostridium acetobutylicum acidogenic fermentation
pyruvate fermentation to butanoate
cob(II)yrinate a,c-diamide biosynthesis I (early cobalt insertion)
isopropanol biosynthesis
GDP-D-glycero-α-D-manno-heptose biosynthesis
L-rhamnose degradation I
pyruvate fermentation to acetone
L-isoleucine biosynthesis IV
tetrapyrrole biosynthesis II (from glycine)
superpathay of heme biosynthesis from glutamate
tetrapyrrole biosynthesis I (from glutamate)
L-histidine degradation I
L-ornithine biosynthesis
L-tryptophan biosynthesis
pyruvate fermentation to isobutanol (engineered)
heme biosynthesis II (anaerobic)
preQ0 biosynthesis
queuosine biosynthesis
TCA cycle VIII (helicobacter)
flavin biosynthesis I (bacteria and plants)
phosphopantothenate biosynthesis I
superpathway of tetrahydrofolate biosynthesis
L-isoleucine biosynthesis II
L-valine biosynthesis
L-isoleucine biosynthesis I (from threonine)
6-hydroxymethyl-dihydropterin diphosphate biosynthesis I
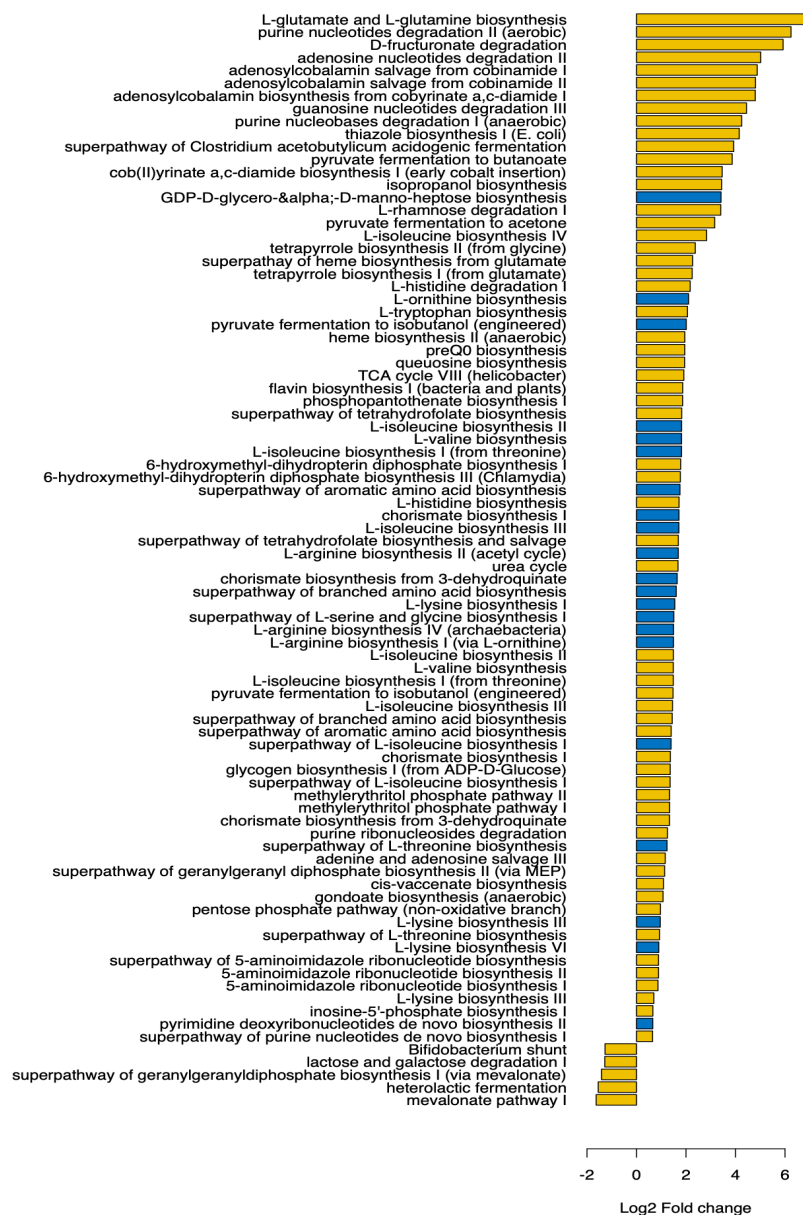6-hydroxymethyl-dihydropterin diphosphate biosynthesis III (Chlamydia)
superpathway of aromatic amino acid biosynthesis
L-histidine biosynthesis
chorismate biosynthesis I
L-isoleucine biosynthesis III
superpathway of tetrahydrofolate biosynthesis and salvage
L-arginine biosynthesis II (acetyl cycle)
urea cycle
chorismate biosynthesis from 3-dehydroquinate
superpathway of branched amino acid biosynthesis
L-lysine biosynthesis I
superpathway of L-serine and glycine biosynthesis I
L-arginine biosynthesis IV (archaebacteria)
L-arginine biosynthesis I (via L-ornithine)
L-isoleucine biosynthesis II
L-valine biosynthesis
L-isoleucine biosynthesis I (from threonine)
pyruvate fermentation to isobutanol (engineered)
L-isoleucine biosynthesis III
superpathway of branched amino acid biosynthesis
superpathway of aromatic amino acid biosynthesis
superpathway of L-isoleucine biosynthesis I
chorismate biosynthesis I
glycogen biosynthesis I (from ADP-D-Glucose)
superpathway of L-isoleucine biosynthesis I
methylerythritol phosphate pathway II
methylerythritol phosphate pathway I
chorismate biosynthesis from 3-dehydroquinate
purine ribonucleosides degradation
superpathway of L-threonine biosynthesis
adenine and adenosine salvage III
superpathway of geranylgeranyl diphosphate biosynthesis II (via MEP)
cis-vaccenate biosynthesis
gondoate biosynthesis (anaerobic)
pentose phosphate pathway (non-oxidative branch)
L-lysine biosynthesis III
superpathway of L-threonine biosynthesis
L-lysine biosynthesis VI
superpathway of 5-aminoimidazole ribonucleotide biosynthesis
5-aminoimidazole ribonucleotide biosynthesis II
5-aminoimidazole ribonucleotide biosynthesis I
L-lysine biosynthesis III
inosine-5'-phosphate biosynthesis I
pyrimidine deoxyribonucleotides de novo biosynthesis II
superpathway of purine nucleotides de novo biosynthesis I
Bifidobacterium shunt
lactose and galactose degradation I
superpathway of geranylgeranyldiphosphate biosynthesis I (via mevalonate)
heterolactic fermentation
mevalonate pathway I

-2   0   2   4   6

Log2 Fold change

Figure S10. Differentially abundant pathways in the small intestine microbiome. We used PICRUSt2 predicted pathway abundances and linear regression model to identify pathways that are significantly enriched in PFOS (blue) / GenX (red) diets (p-value < 0.01). Only pathways with |Log2 Fold change|>1 are shown.

Figure S11. Differentially abundant pathways in the colon microbiome. We used PICRUSt2 predicted pathway abundances and linear regression model to identify pathways that are significantly enriched in PFOS (blue) / GenX (red) diets (p-value < 0.01). Only pathways with |Log2 Fold change|>1 are shown.
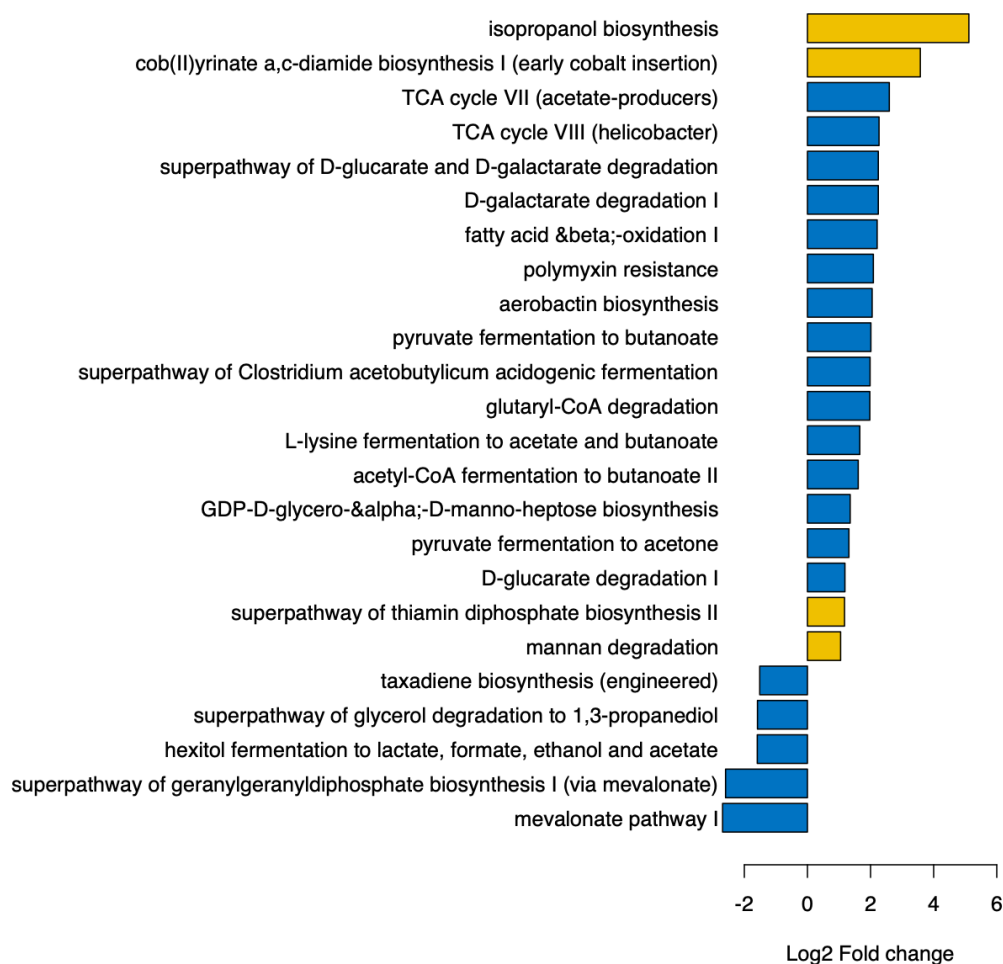
Figure S12. Heatmap of Liver metabolites which are significantly varying between different diets. Metabolites were selected based on the ANOVA test (FDR adjusted p-value < 0.01), peaks with low interquartile variability were filtered out before this analysis. Only 196 metabolites with names are shown (out of 491). Color represents z-score value of log10 normalized metabolite concentrations.