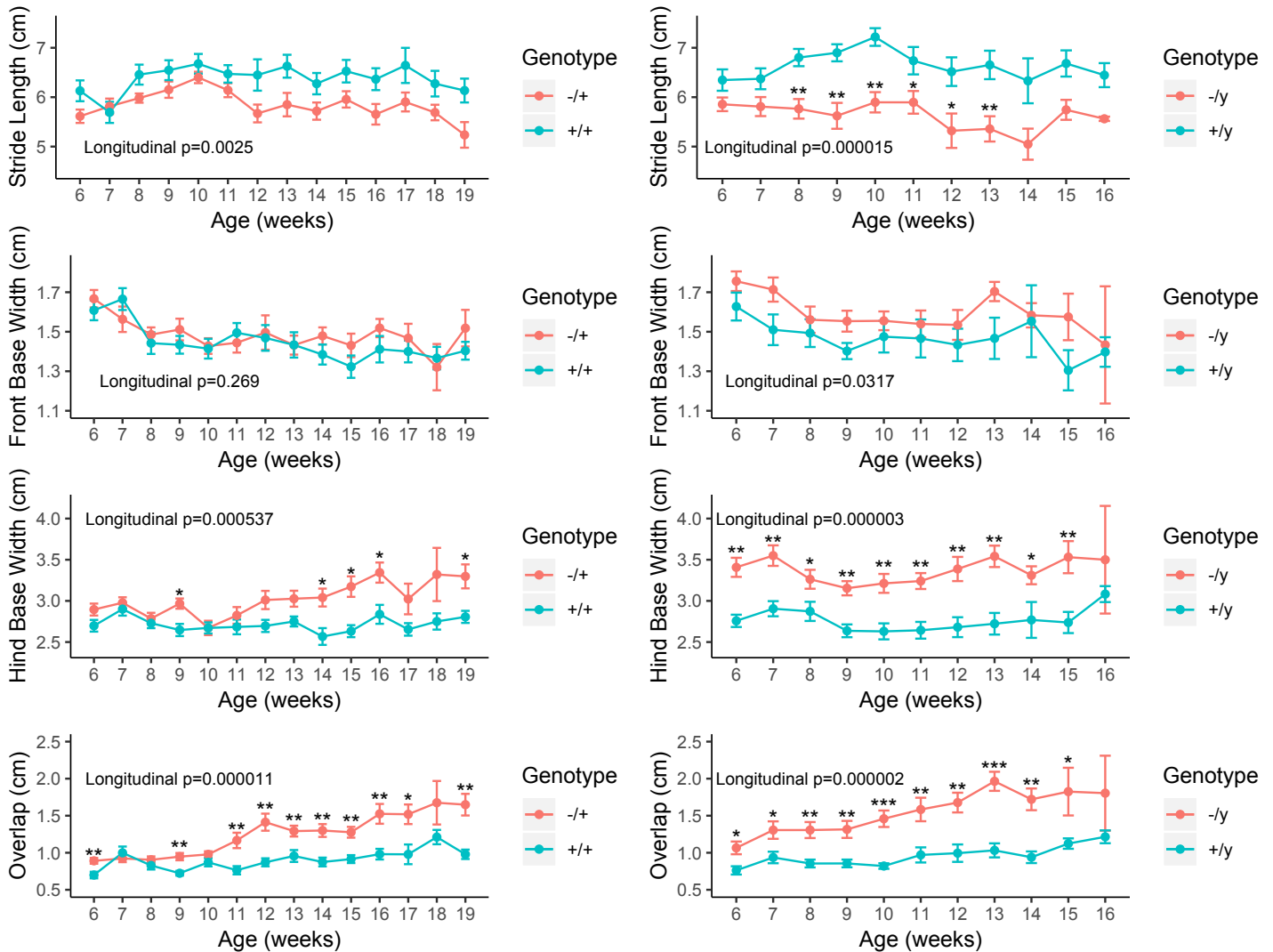
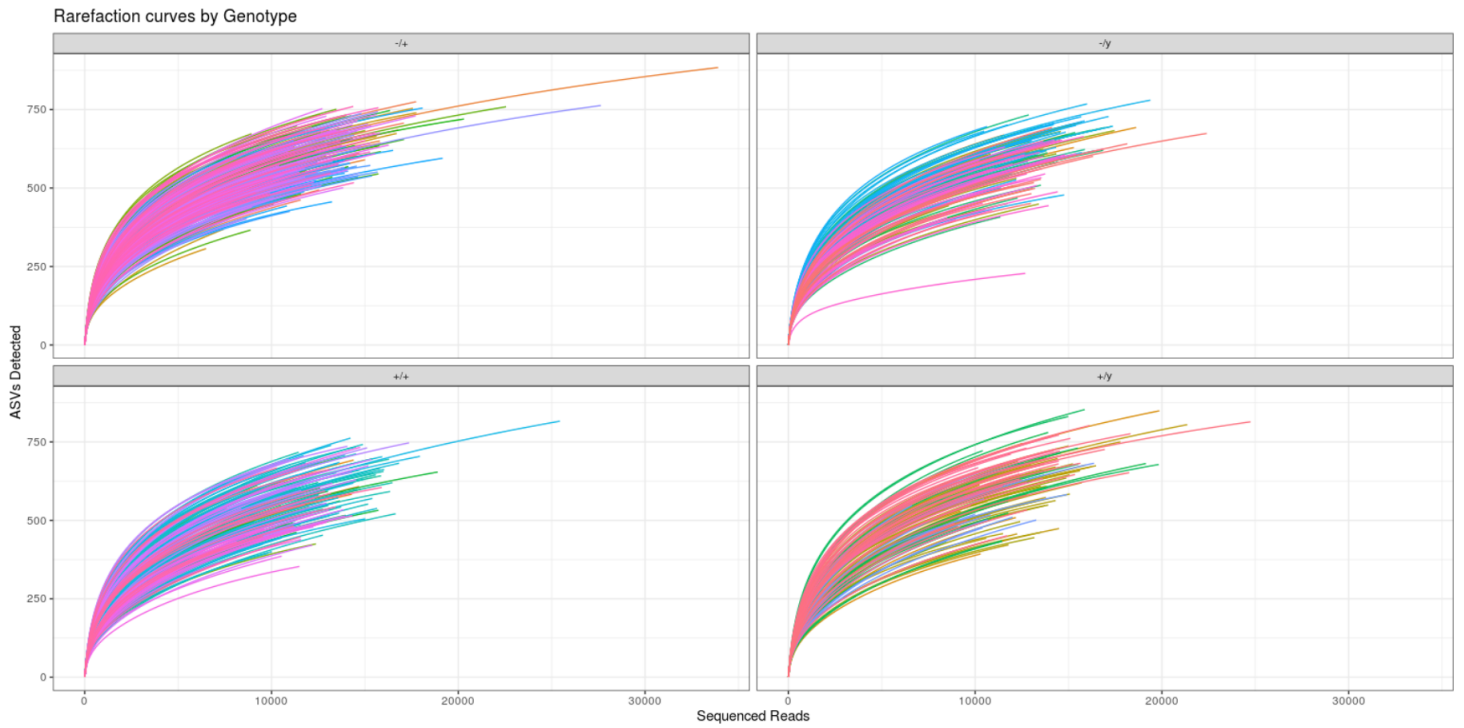
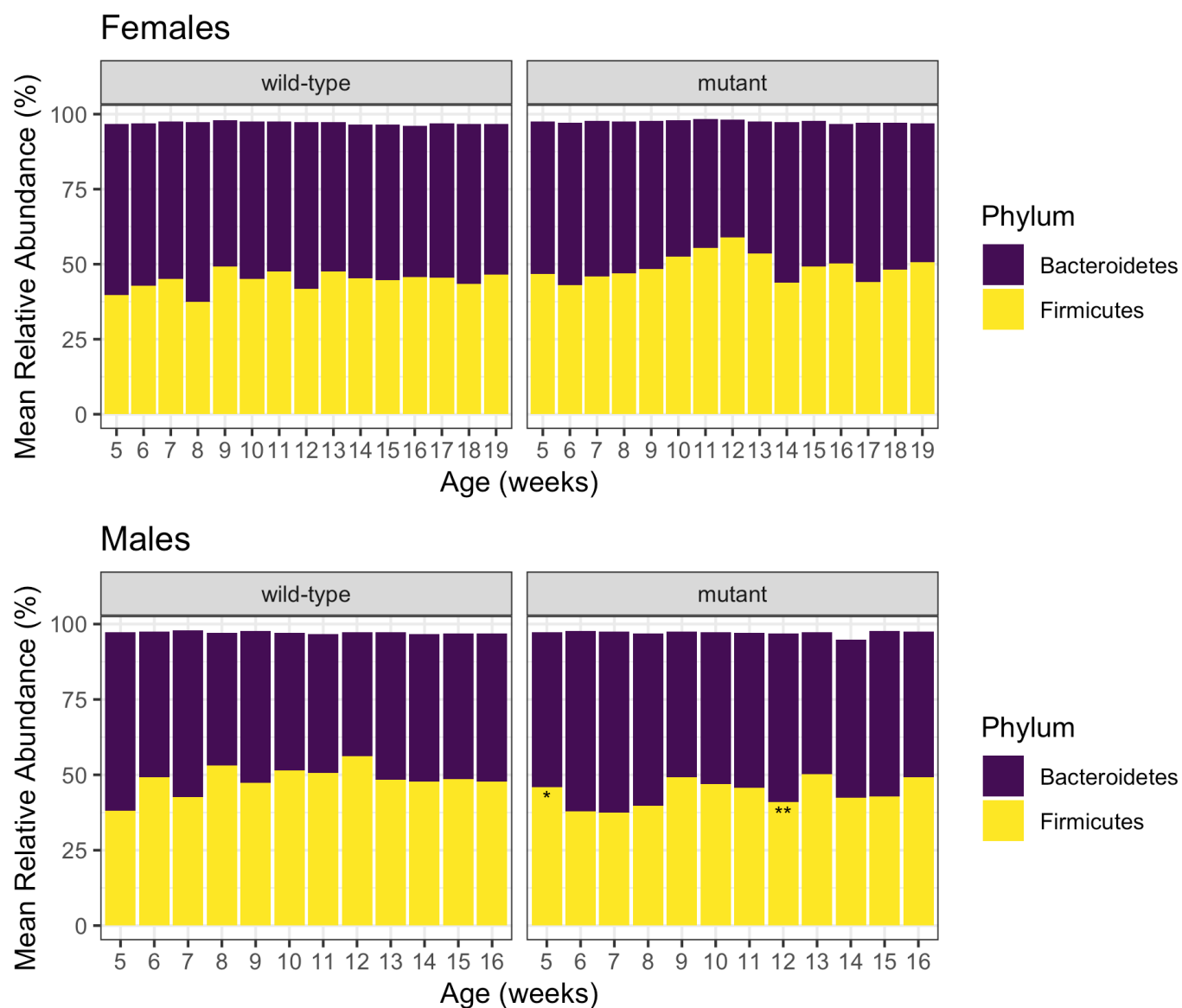




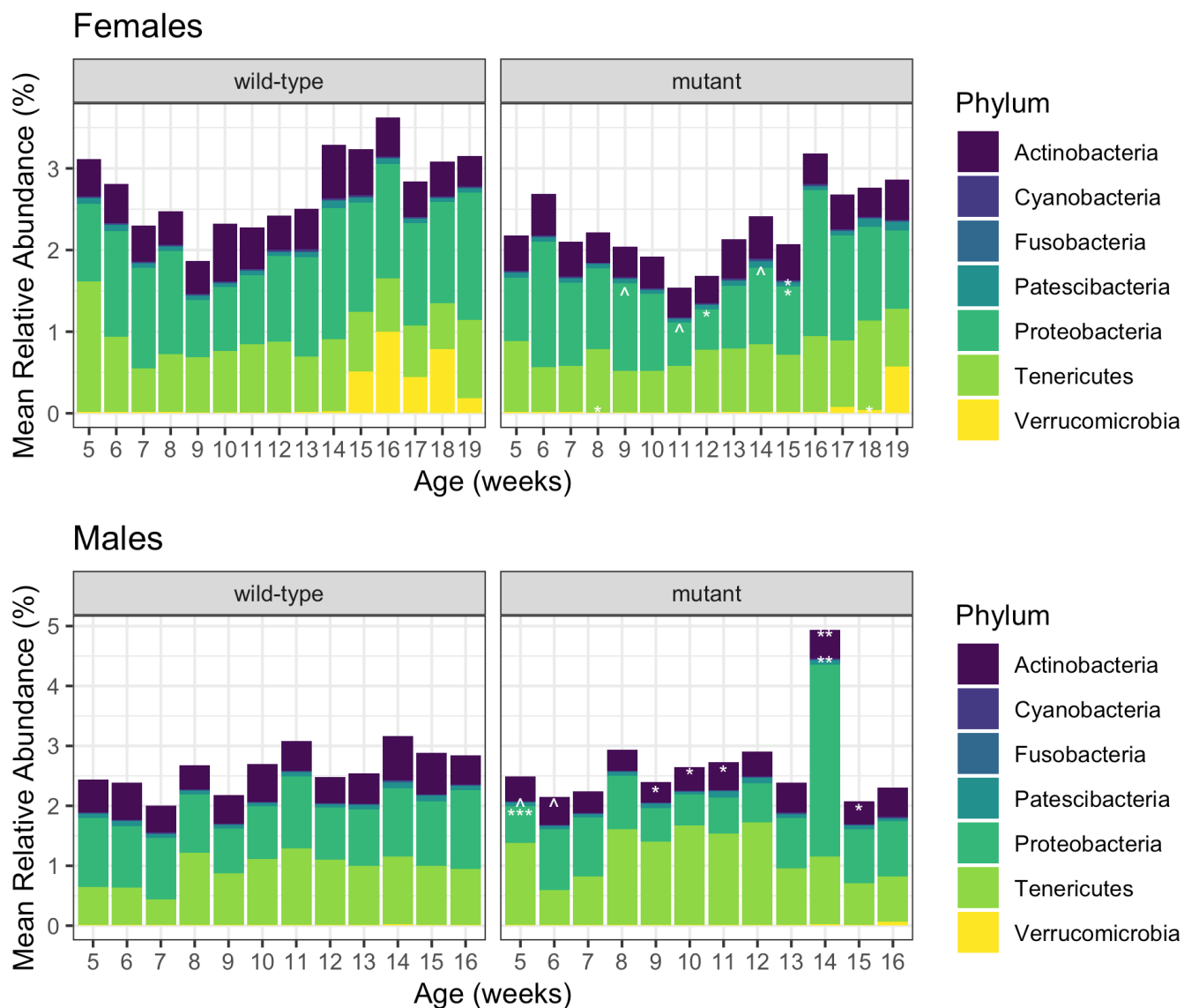
Supplemental Figure 1. Longitudinal gait analysis. Comparison of four different measures of gait across genotypes: 1) stride length, 2) front base width (distance between the two front legs while walking in cm), 3) hind base width (distance between the two hind legs while walking in cm), and 4) overlap distance (distance between front and hind paws while walking). Dots represent the mean and error bars are +/- standard error of the mean (SEM). $N=11-19/\text{genotype}/\text{sex}$. *FDR<0.05, **FDR<0.01, ***FDR<0.001 in mutant vs. control, cross-sectionally. Longitudinal p-values represent the overall association between mutant and wild-type mice across disease course using linear mixed effects models.



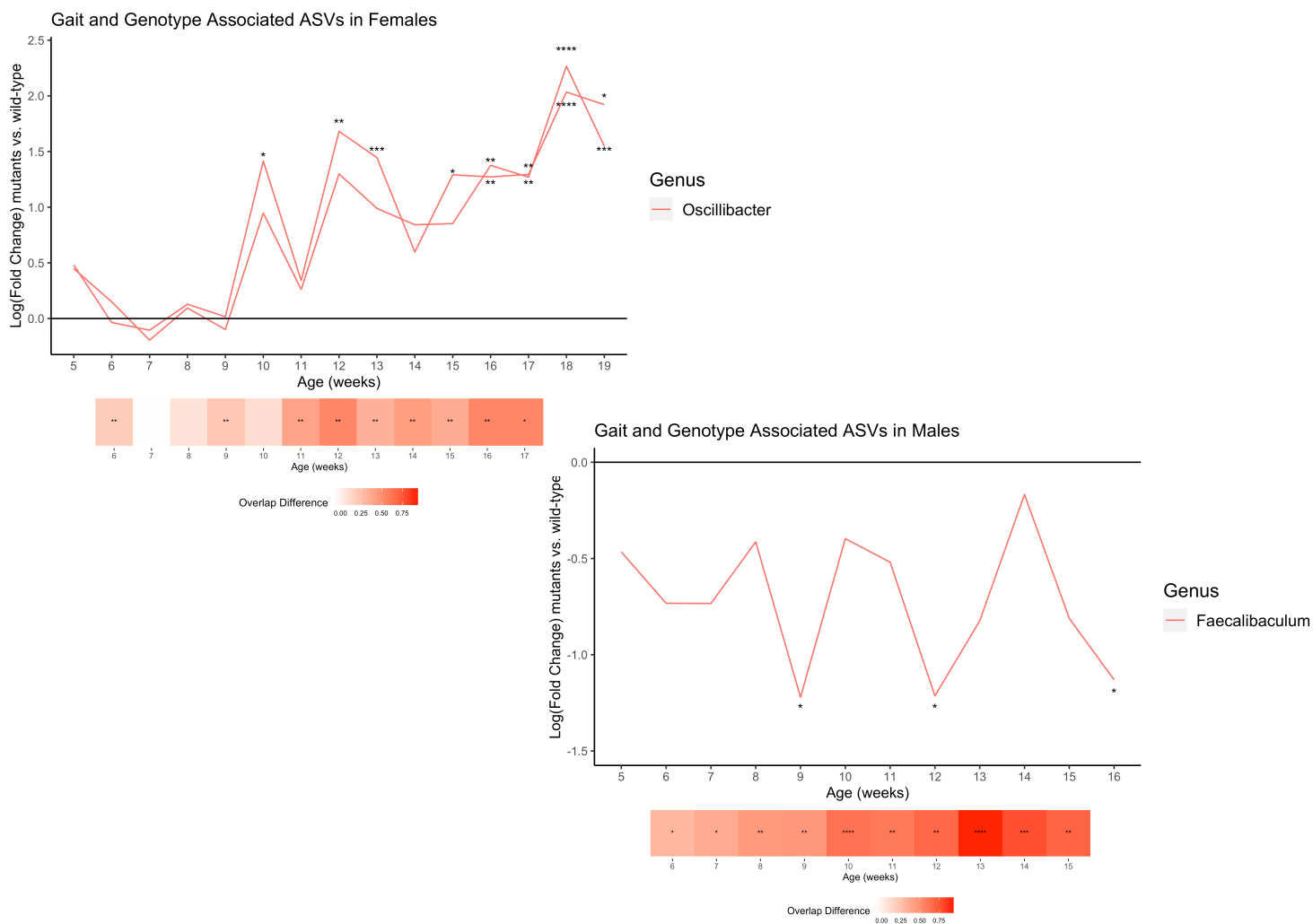
Supplemental Figure 2. Rarefaction curves by sex and genotype. Rarefaction curves show the number of Amplicon Sequence Variants (ASVs) detected as a function of sequencing depth. Each line represents an individual sample. N=11-19/genotype/sex/time point.



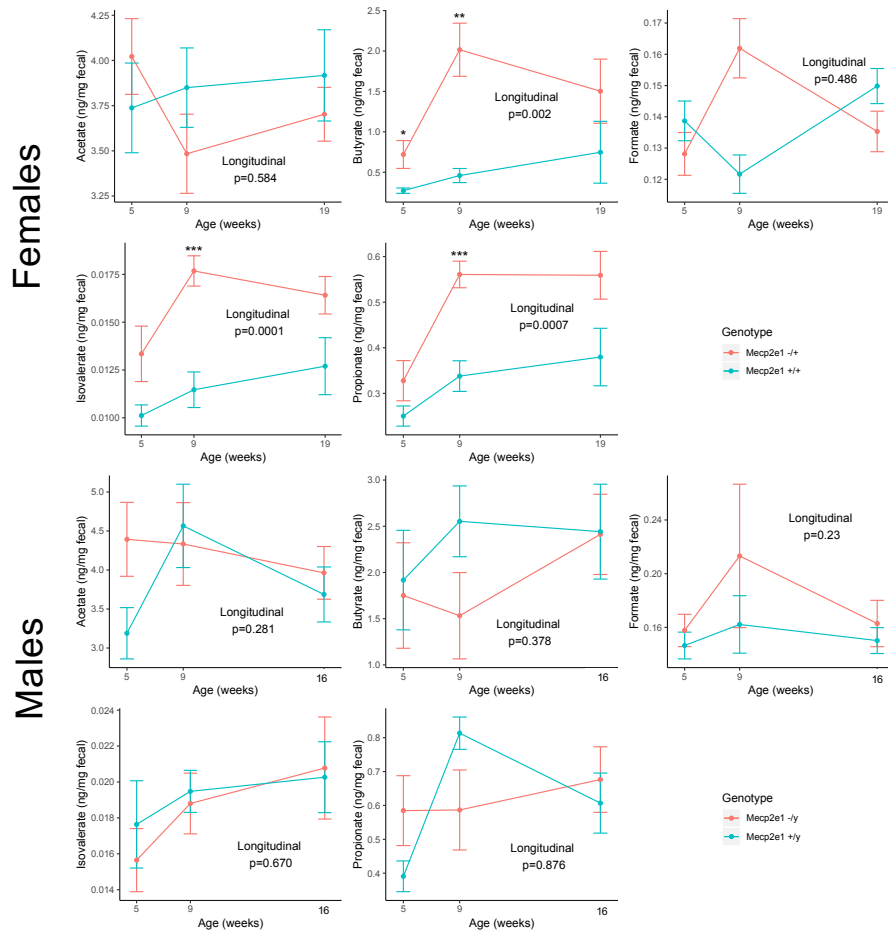
Supplemental Figure 3. Relative abundance of major phyla in fecal samples by sex and genotype. Mean relative abundance for the two major phyla in the mouse gut microbiome (*Firmicutes* and *Bacteroidetes*) represented by stacked bar plots at each collection time point, stratified by sex and genotype. N=11-19/genotype/sex/time point. *FDR<0.05, **FDR<0.01 in mutant vs. wild-type cross-sectionally.



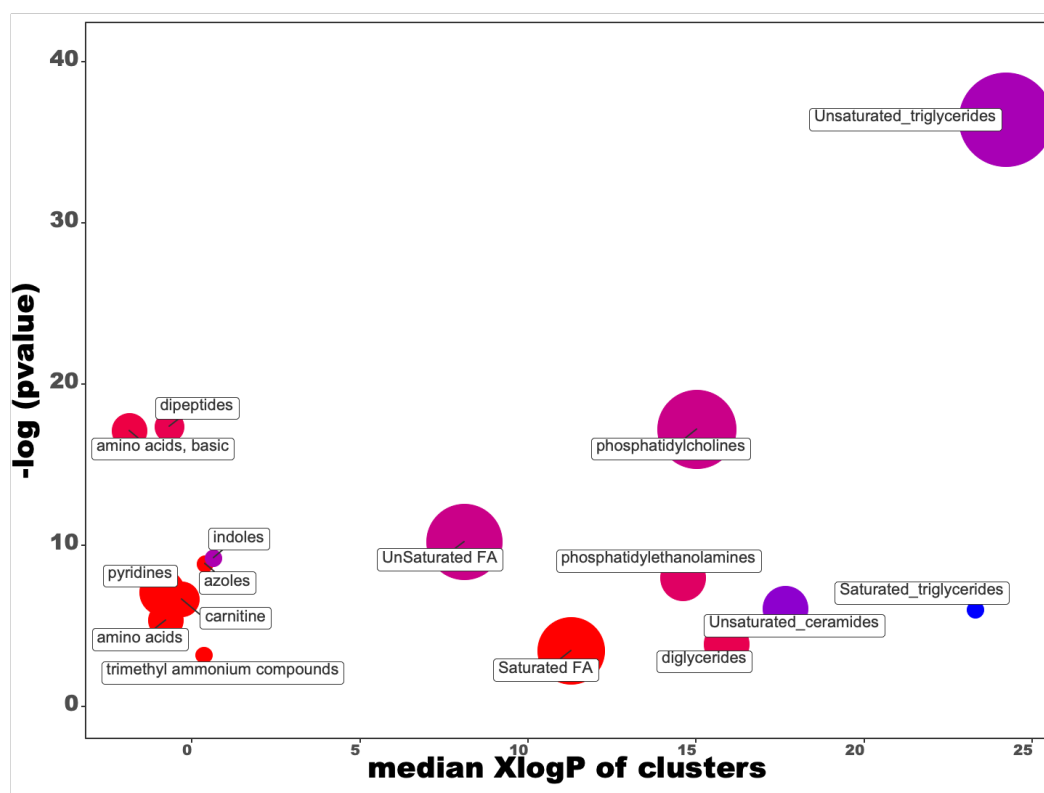
Supplemental Figure 4. Relative abundance of minor phyla in fecal samples by sex and genotype. Mean relative abundance for minor phyla in the mouse gut microbiome (*Actinobacteria*, *Cyanobacteria*, *Fusobacteria*, *Patescibacteria*, *Proteobacteria*, *Tenericutes*, and *Verrucomicrobia*) represented by stacked bar plots at each collection time point, stratified by sex and genotype. N=11-19/genotype/sex/time point. ^FDR<0.10, *FDR<0.05, **FDR<0.01 in mutants vs. wild-type.



Supplemental Figure 6. Differences in genotype- and gait-associated ASVs in mutant vs. wild-type *Mecp2-e1* mice across disease course. Log(fold change) of mutant vs. wild-type (wt) mice at each time point in ASVs that were significantly associated with gait (overlap distance) and genotype at a minimum of one-quarter of the time points in both females and males. Each line represents an individual ASV and lines are colored by Genus. Heatmaps at the bottom of each graph are colored red with increasing intensity based on the difference in phenotype measure by genotype at each time point as reported in Figure 1. N=11-19/genotype/sex. *FDR<0.05, **FDR<0.01, ***FDR<0.001 in *Mecp2-e1* mutant vs. wild-type mice cross-sectionally at each time point.

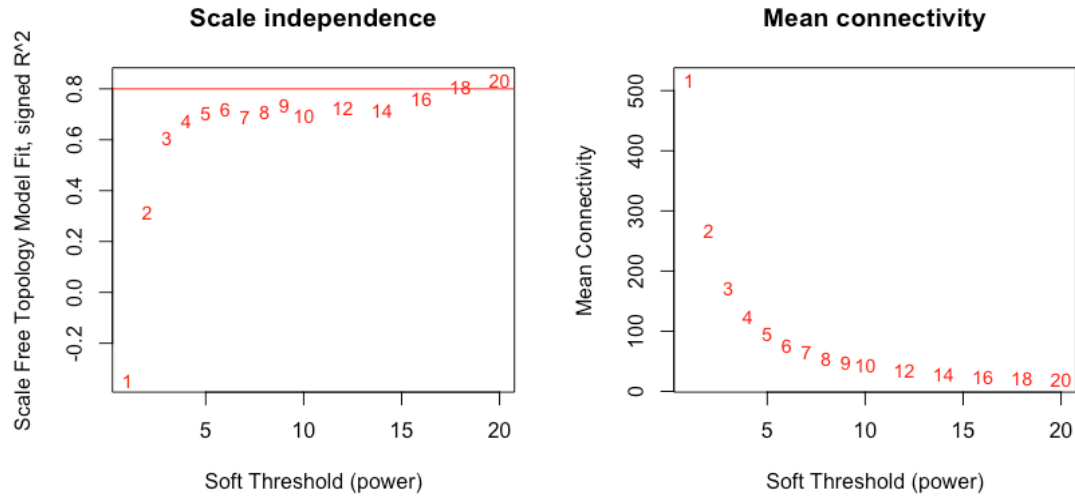


Supplemental Figure 7. Fecal SCFA levels across age in *Mecp2-e1* mutant vs. wild-type females and males. Each dot represents the mean of each group at each time point and error bars are +/- standard error of the mean (SEM). N=8/genotype/sex. *FDR<0.05, **FDR<0.01 in *Mecp2-e1* mutant vs. wild-type cross-sectionally. Longitudinal p-values represent the overall association between mutant and wild-type mice across disease course using linear mixed effects models controlling for within-litter effects and repeated measures.

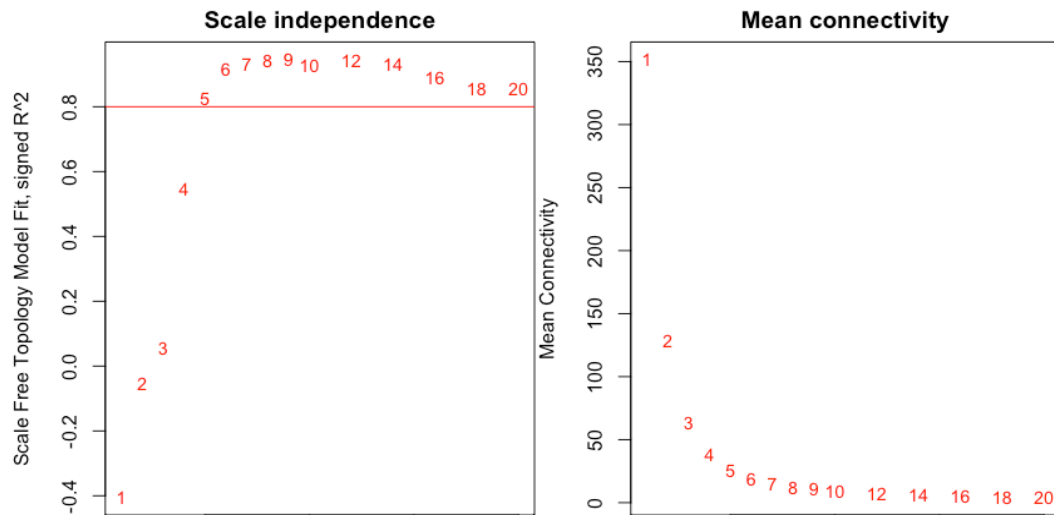


Supplemental Figure 8. Fecal metabolite clusters enriched in female *Mecp2-e1* $\pm/\pm$ vs. $\pm/\pm$ mice. The y-axis is $-\log(p\text{-value})$ resulting in the most statistically significant clusters being placed higher on the graph. Median XlogP (x-axis) of clusters is the cluster order on the cluster similarity tree generated by ChemRICH. Cluster colors give the proportion of increased or decreased compounds (red = increased, blue = decreased), and the size of each circle is determined by the total number of metabolites in each cluster.

Females



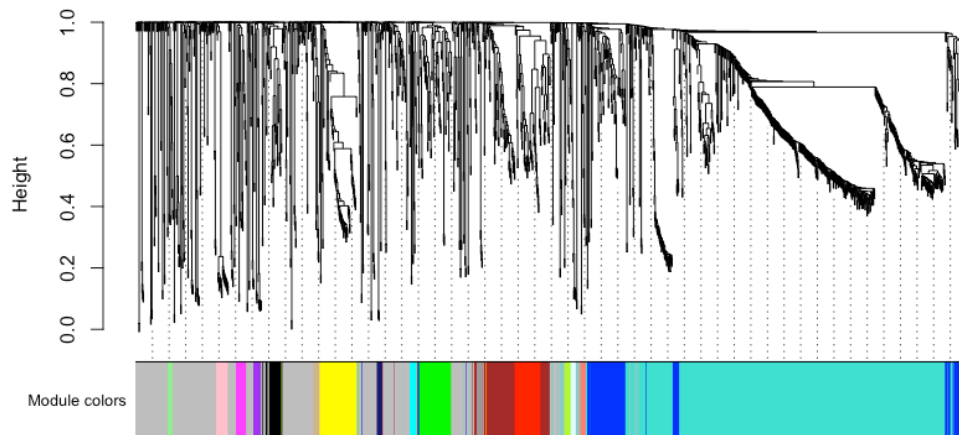
Males



Supplemental Figure 9. Fecal metabolite scale independence and mean connectivity plots for WGCNA analysis in females and males. The soft threshold plotted against scale free topology model fit (R^2) is used to determine a soft power threshold for network construction. The smallest value at which the scale free topology model R^2 is at least 0.8 is selected as the soft power threshold. Mean connectivity plots examine the network connectivity plotted against different soft power thresholds.

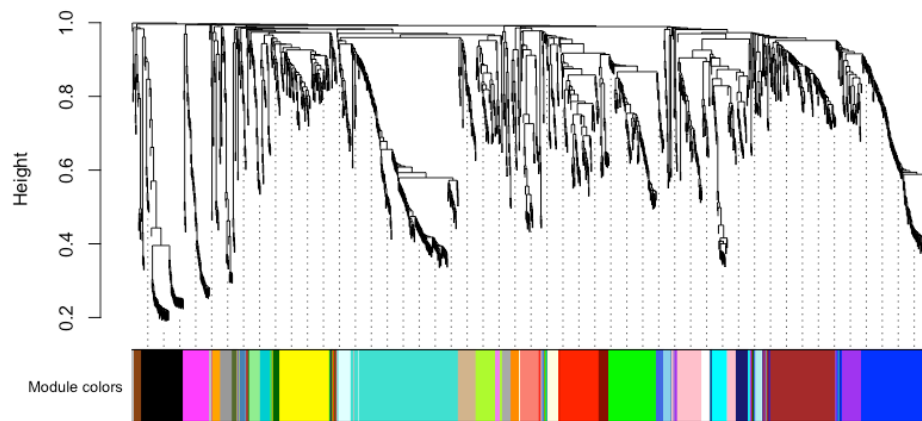
Females

Cluster Dendrogram



Males

Cluster Dendrogram



Supplemental Figure 10. Fecal metabolite modules represented as cluster dendrograms. Cluster dendrograms for WGCNA modules depict the clustering of metabolites and how they fit within each module.