Supplementary online material of Beneficial Microbes
DOI: https://doi.org/10.3920/BM2017.0191

Eczema-protective probiotic alters infant gut microbiome functional capacity but not composition: sub-sample analysis from a RCT

R. Murphy, X.C. Morgan, X.Y. Wang, K. Wickens, G. Purdie, P. Fitzharris, A. Otal, B. Lawley, T. Stanley , C. Barthow, J. Crane, E.A. Mitchell and G.W. Tannock

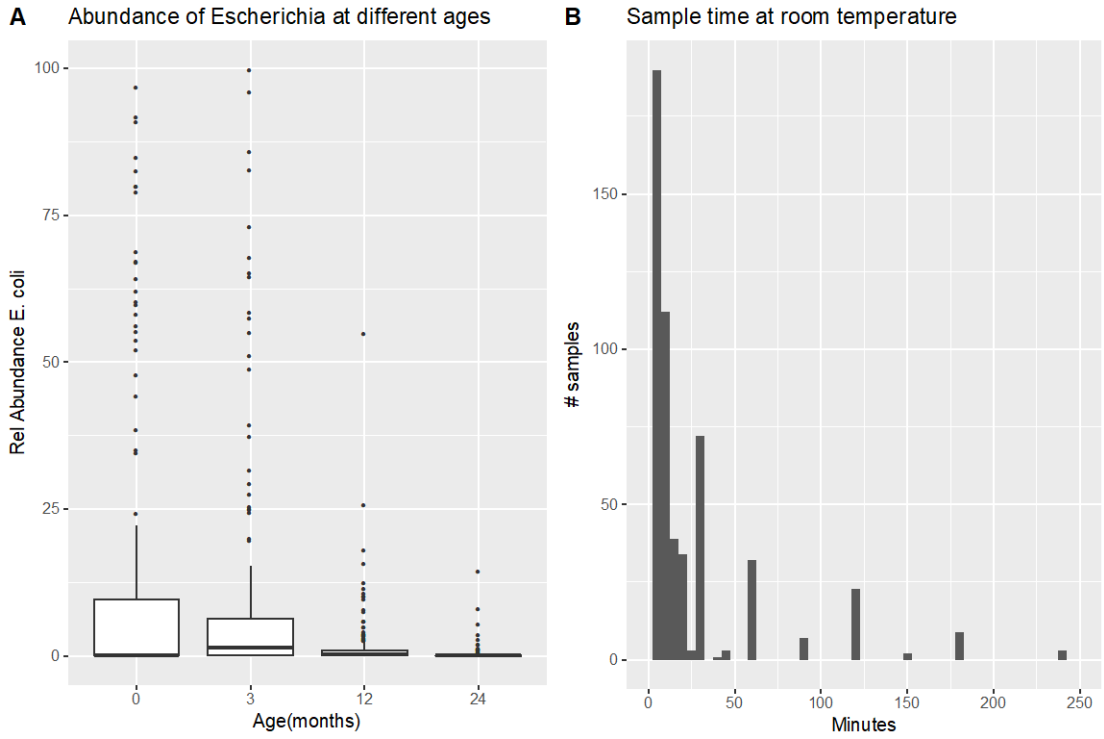

Figure S1. (A) The distribution of *Escherichia coli* relative abundance in all samples in the dataset, stratified by age. (B) Reported duration of storage of study samples at room temperature prior to freezing.

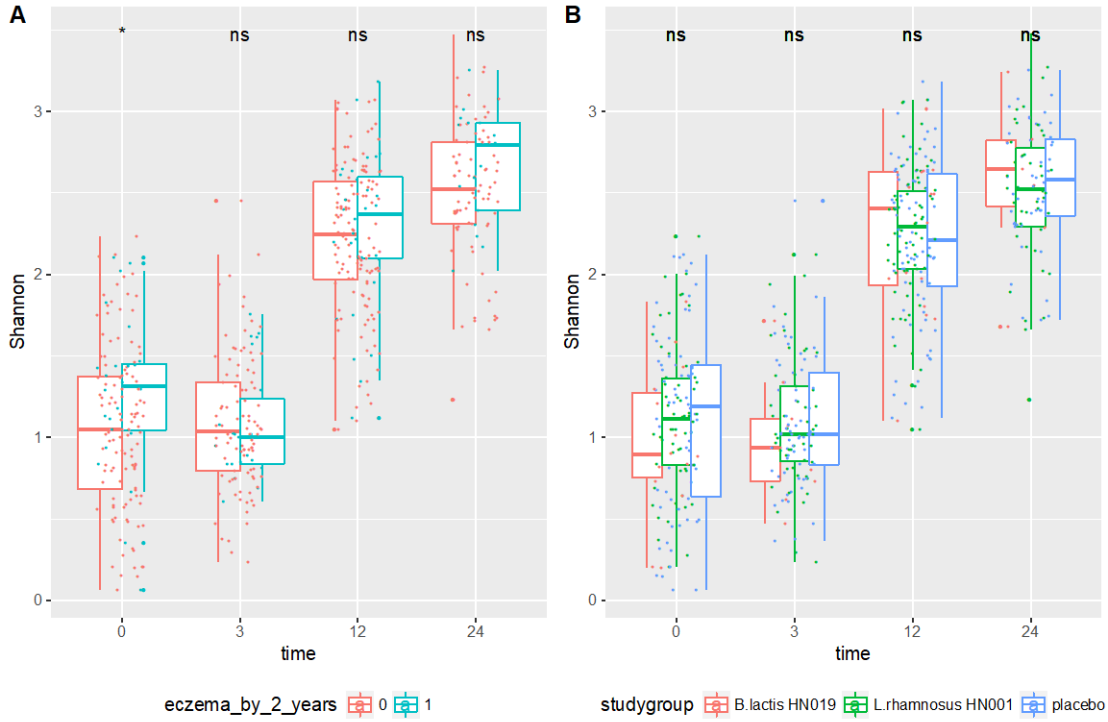

Figure S2. Filtered data excluding samples with *Escherichia coli* outliers showing (A) Shannon diversity, stratified by age, for groups exposed to probiotics or placebo; (B) Shannon diversity, stratified by age and eczema outcome. NS = *P*>0.05; *=*P*<0.05, **=*P*<0.01, *=*P*<0.001, ****=*P*<0.0001, Kruskal-Wallis test. See Supplementary Figure S6 to see this figure with all (unfiltered) data.**

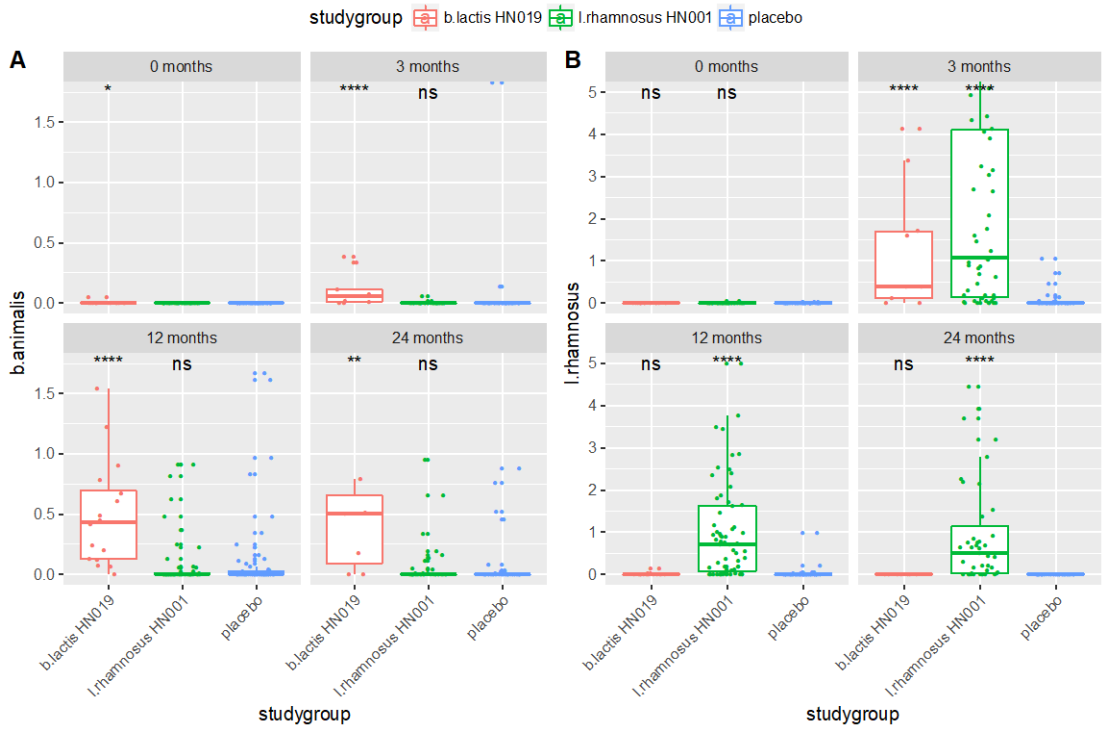

Figure S3. Filtered data excluding samples with *Escherichia coli* outliers showing relative abundances of (A) *Bifidobacterium animalis* subsp. *lactis* and (B) *Lactobacillus rhamnosus* detected within treatment groups. Kruskal-Wallis tests were used to test the null hypothesis that at each time point, there was no difference between the median relative abundance of *B. animalis* or *L. rhamnosus* between study groups and placebo groups (NS=*P*>0.05, *=*P*<0.05, **=*P*<0.01, *=*P*<0.001, ****=*P*<0.0001, Kruskal-Wallis test). See Supplementary Figure S7 to see this figure with all (unfiltered) data.**

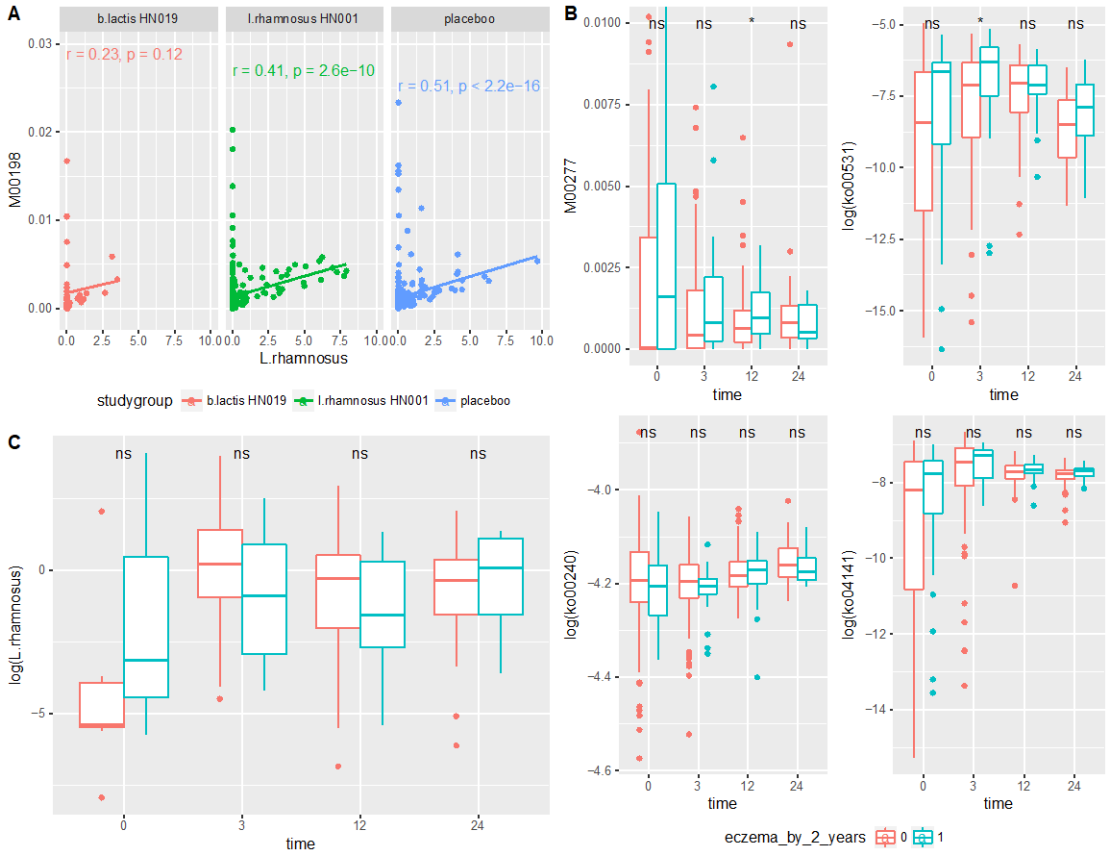

Figure S4. (A) The correlation between *Lactobacillus rhamnosus* relative abundance in a sample and the abundance of M00198. (B) Relative abundance of KEGG modules and pathways significantly associated with eczema, stratified by age of sampling and eczema outcome. NS=*P*>0.05; *=*P*<0.05, **=*P*<0.01; *=*P*<0.0001; Wilcoxon signed-rank test. (C) Mean abundance of *L rhamnosus* in faecal samples by eczema outcome stratified by age of sampling (NS = *P*>0.05 for all eczema/non-eczema pairs; Wilcoxon signed-rank test). See Supplementary Figure S7 to see this figure with all (unfiltered) data.**

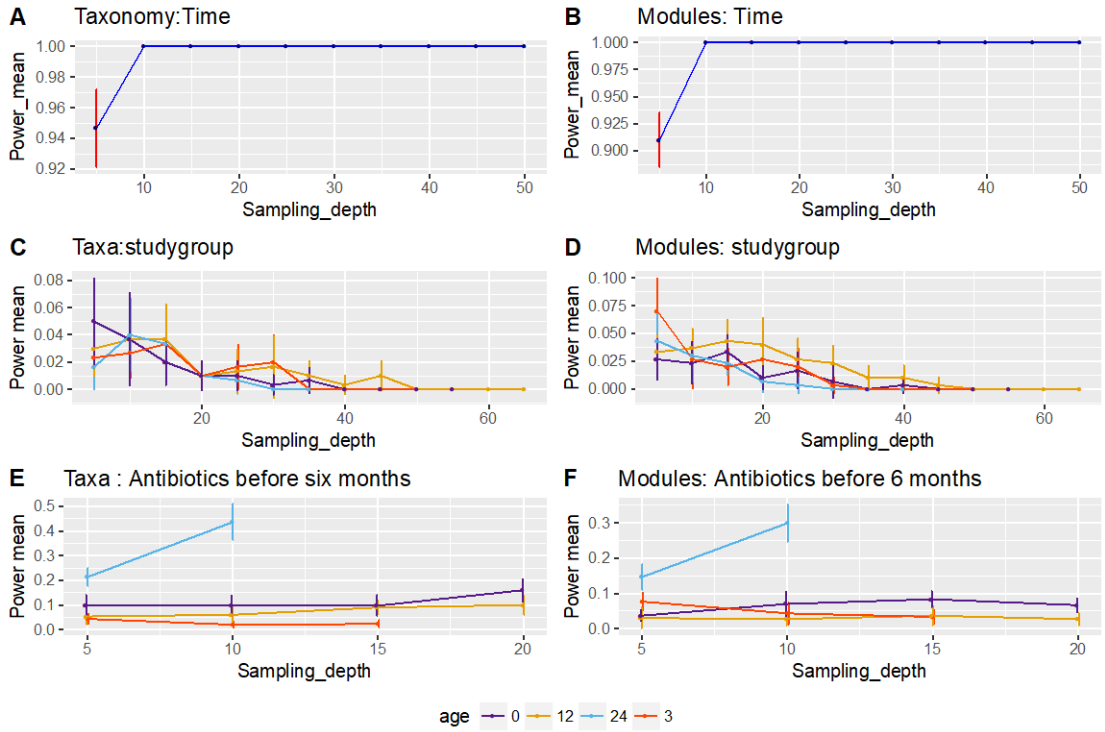

Figure S5. Filtered data excluding samples with *Escherichia coli* outliers showing the power to detect differences in beta diversity based on taxonomy (A, C, E) and functional data (B, D, F), shown by (A,B) age of sampling, (C,D) probiotic exposure, and (E,F) antibiotic use in the first 6 months of age.

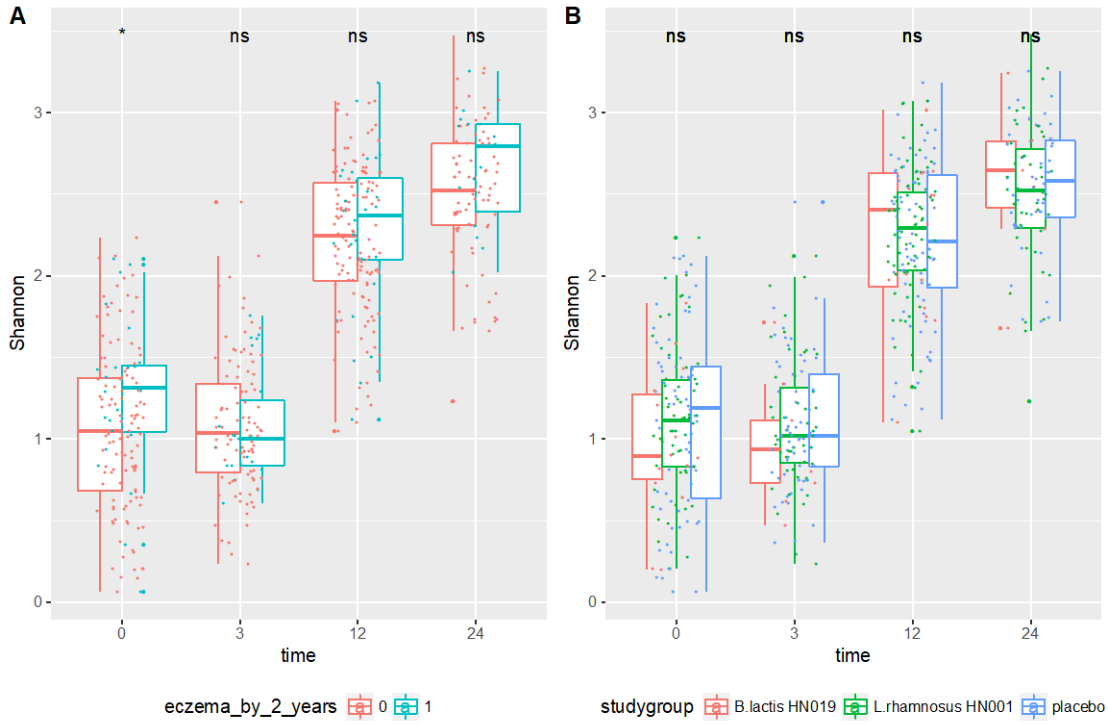

Figure S6. Unfiltered (all) data showing (A) Shannon diversity, stratified by age, for groups exposed to probiotics or placebo; (B) Shannon diversity, stratified by age and eczema outcome. *=*P*<0.05, NS=*P*>0.05, Kruskal-Wallis test.

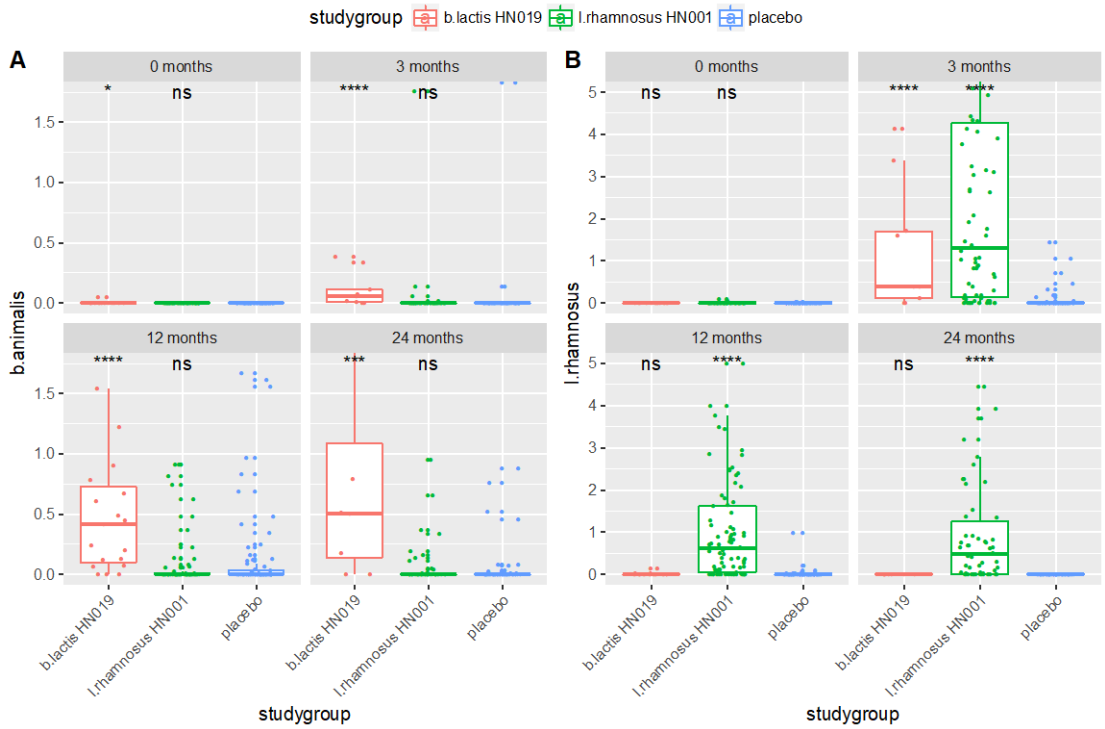

Figure S7. Unfiltered (all) data showing relative abundances of (A) *Bifidobacterium animalis* subsp. *lactis* and (B) *Lactobacillus rhamnosus* detected within treatment groups. Kruskal-Wallis tests were used to test the null hypothesis that at each time point, there was no difference between the median relative abundance of *B. animalis* or *L. rhamnosus* between study groups and placebo groups.

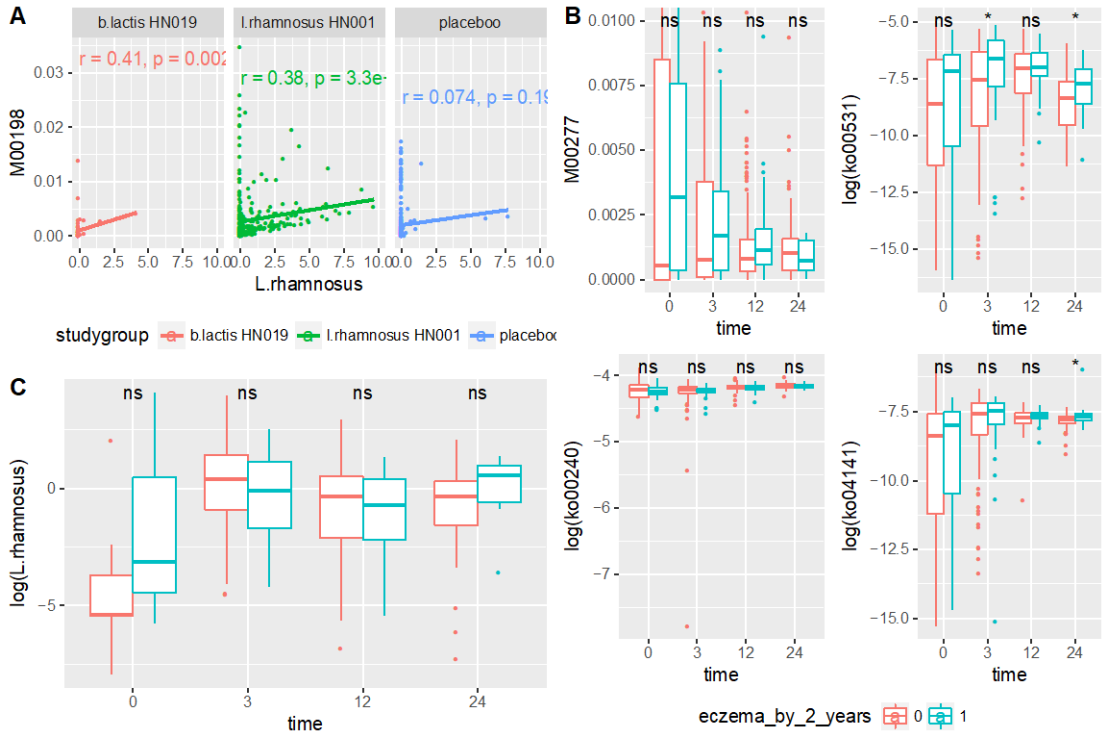

Figure S8. Unfiltered (all) data showing (A) The correlation between the relative abundances of *Lactobacillus rhamnosus* and M00198 in a sample. (B) Relative abundance of KEGG modules and pathways significantly associated with eczema, stratified by age of sampling and eczema outcome. (C) Mean abundance of *L. rhamnosus* in faecal samples, stratified by age of sampling and eczema outcome. (NS=*P*<0.05, *=*P*<0.05, **=*P*<0.01, Wilcoxon signed-rank test for all eczema / non-eczema pairs).

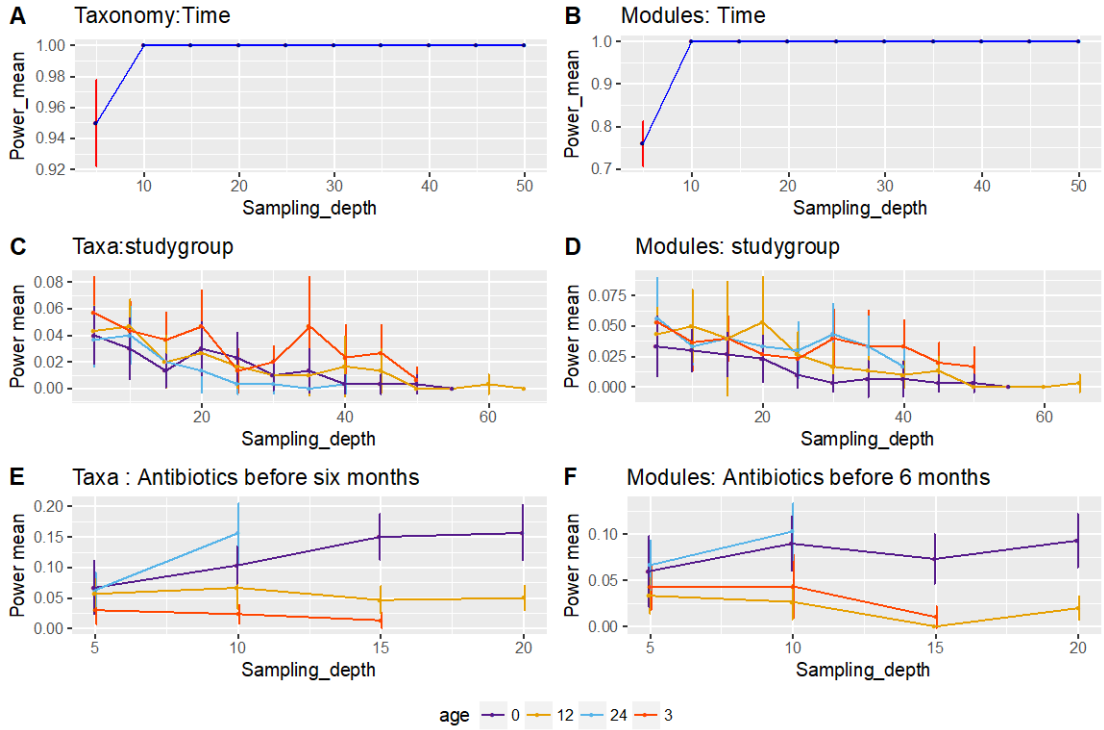

Figure S9. Unfiltered (all) data showing the power to detect differences in beta diversity based on taxonomy (A, C, E) and functional data (B, D, F), shown by (A,B) age of sampling, (C,D) probiotic exposure, and (E,F) antibiotic use in the first 6 months of age. This figure is identical to Supplementary Figure S2, except it was made without filtering data for *Escherichia coli* outliers.

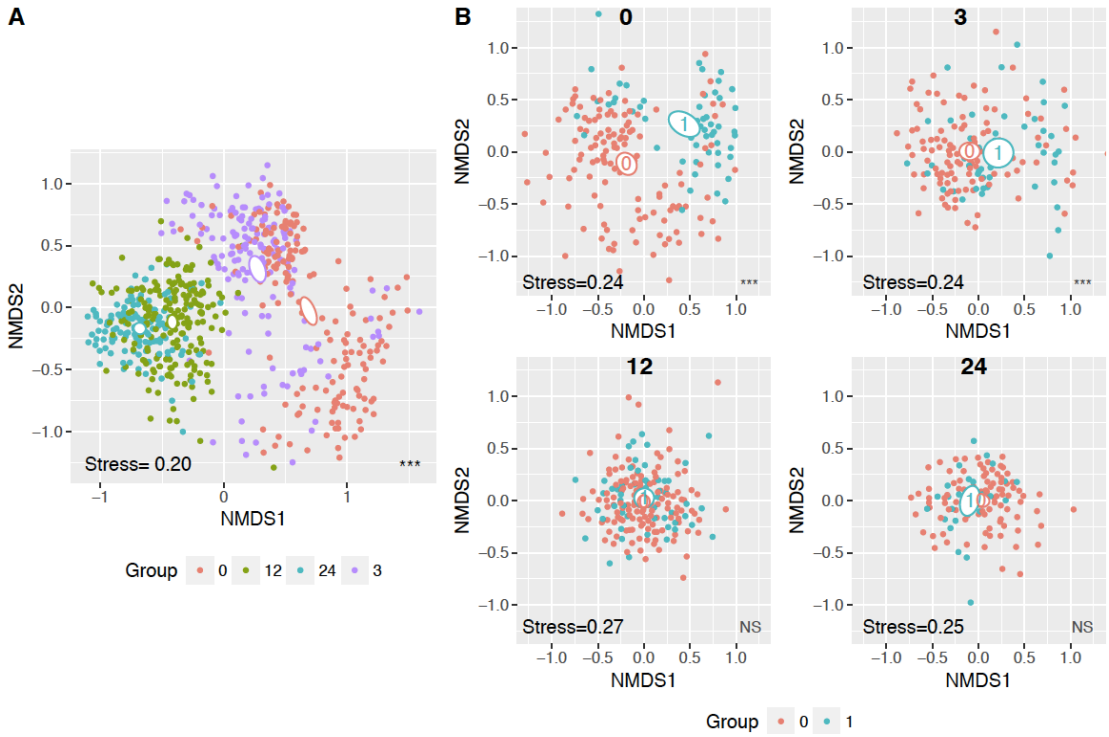

Figure S10. Unfiltered (all) data showing the effects of time and C-section on beta diversity. Non-metric multidimensional scaling of taxonomy by Bray-Curtis distance shows the effects of time and C-section delivery on microbial communities. Adonis was used to assess significance. Colours correspond to time of sample collection (panel A) and C-section delivery (panel B). All samples were used for panel A. In panel B, samples are stratified by the time point in which they were collected. Ovals correspond to the standard error (95% confidence interval) of the weighted average of scores in each group. P-values were calculated using permanova analysis (NS = *P*>0.05, *=*P*<0.05, **=*P*<0.01, *=*P*<0.001, ****=*P*<0.0001). Stress indicates how well the ordination represents the data (ordination stress >0.3 indicates poor fit; 0.2-0.3 indicates medium fit; <0.2 indicates good fit).**

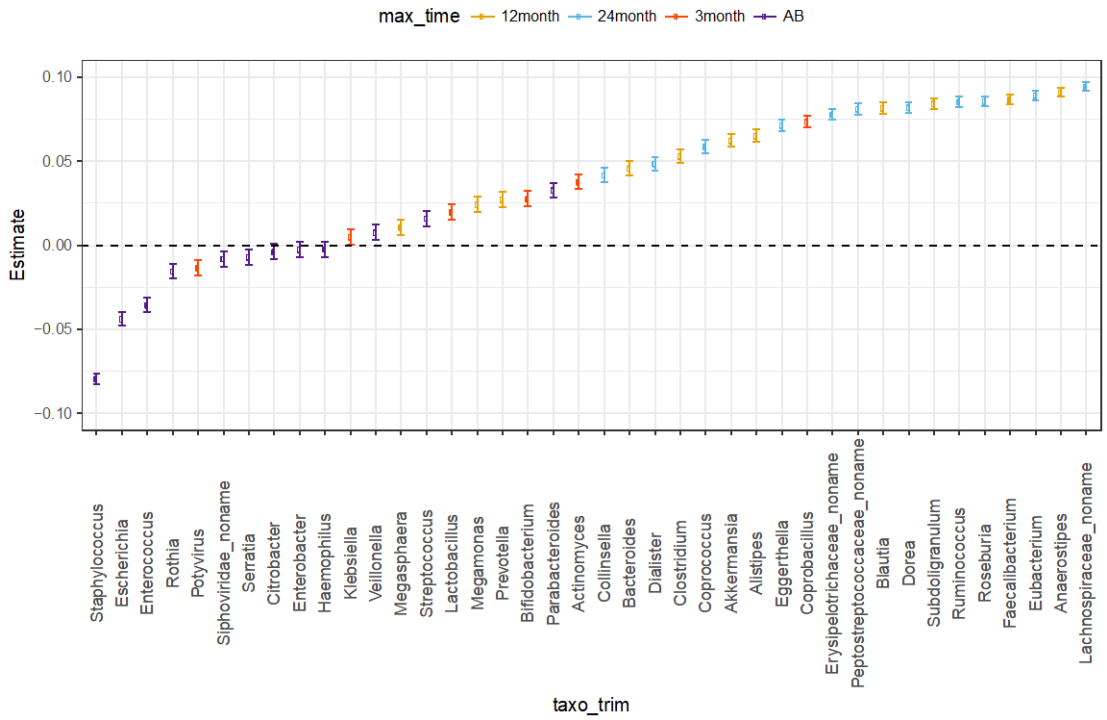

Figure S11. Unfiltered (all) data showing the relationship between age and taxa. To calculate the coefficient estimate, we fitted a linear model between time and the log-10 normalized relative abundance of the 40 most abundant genera. Negative coefficient estimates indicate a negative relationship with time (e.g. more abundant at earlier time points, less abundant at later time points). The complete list of associations is contained in Supplementary Table S3. Each taxa is coloured according to the time point at which its mean abundance is highest. This figure is comparable to Figure 2A except it was made without filtering the data for E. coli outliers and has not been assembled in Inkscape.

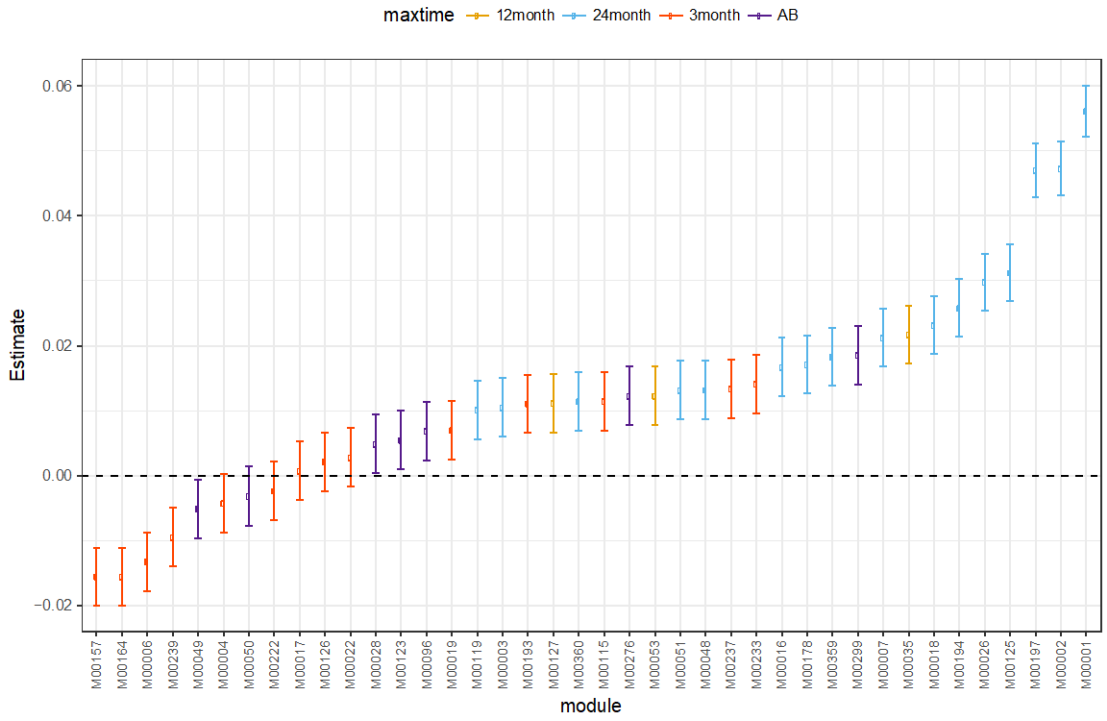

Figure S12. Unfiltered (all) data showing the relationship between age and KEGG modules. To calculate the coefficient estimate, we fitted a linear model between time and the log-10 normalized relative abundance of KEGG modules. Negative coefficient estimates indicate a negative relationship with time (e.g. more abundant at earlier time points, less abundant at later time points). The 40 most abundant KEGG modules in the dataset are shown. The complete list of associations is contained in Table S4. Each taxa and module is coloured according to the time point at which its mean abundance is highest. This figure is comparable to Figure 2B except it was made without filtering the data for *Escherichia coli* outliers and has not been assembled in Inkscape.

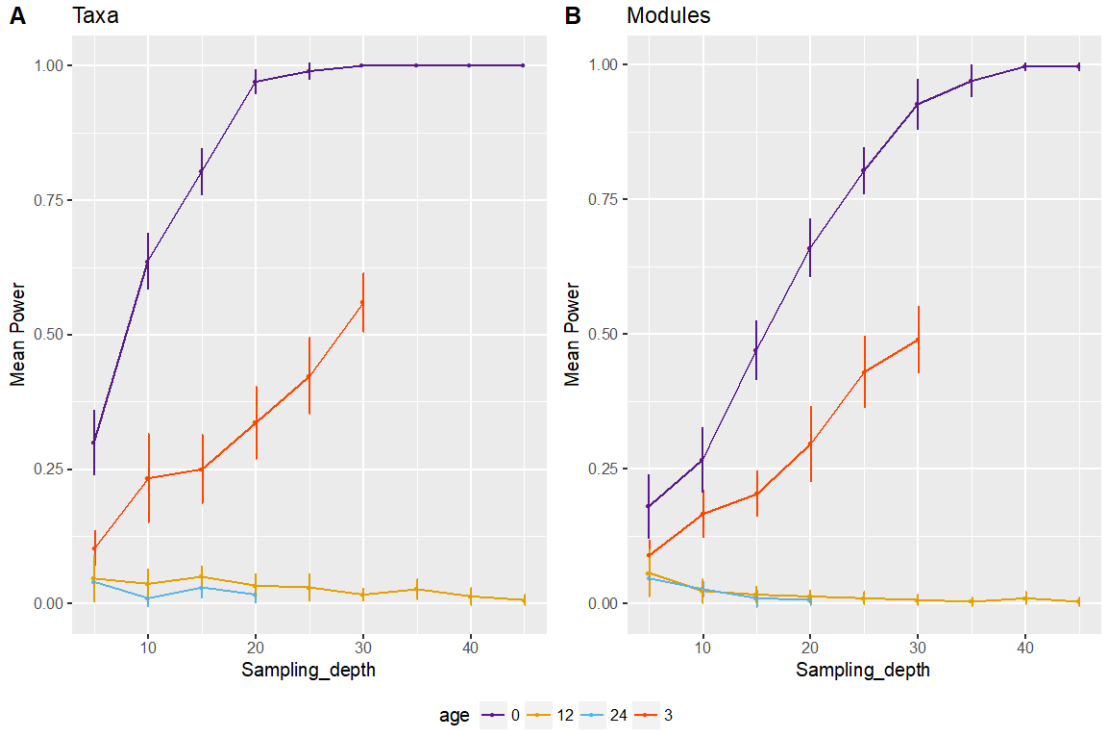

Figure S13. Unfiltered (all) data showing power to distinguish mode of delivery (C-section vs vaginal). (A) Power to distinguish between modes of delivery in (A) taxonomic data and (B) functional data. The X axis corresponds to the number of samples per group (without replacement). This figure is comparable to Figure 4, except it was made without filtering the data for *Escherichia coli* outliers.

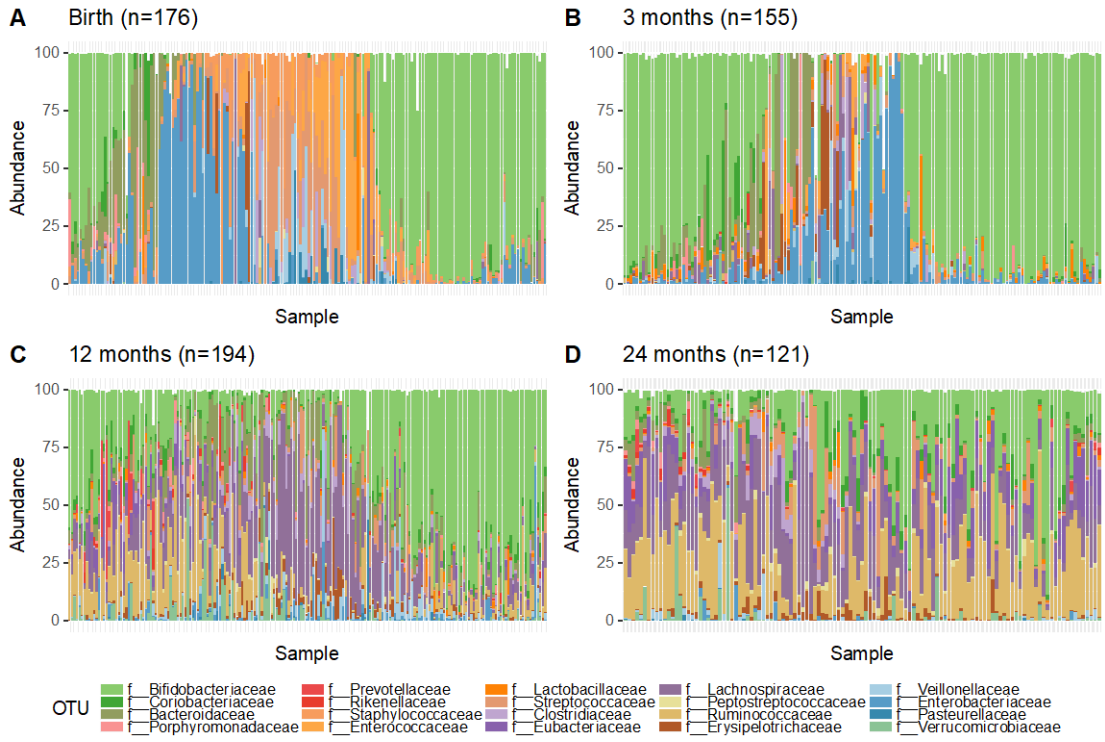

Figure S14. Distribution of taxa across samples, stratified by time and shown at family level.