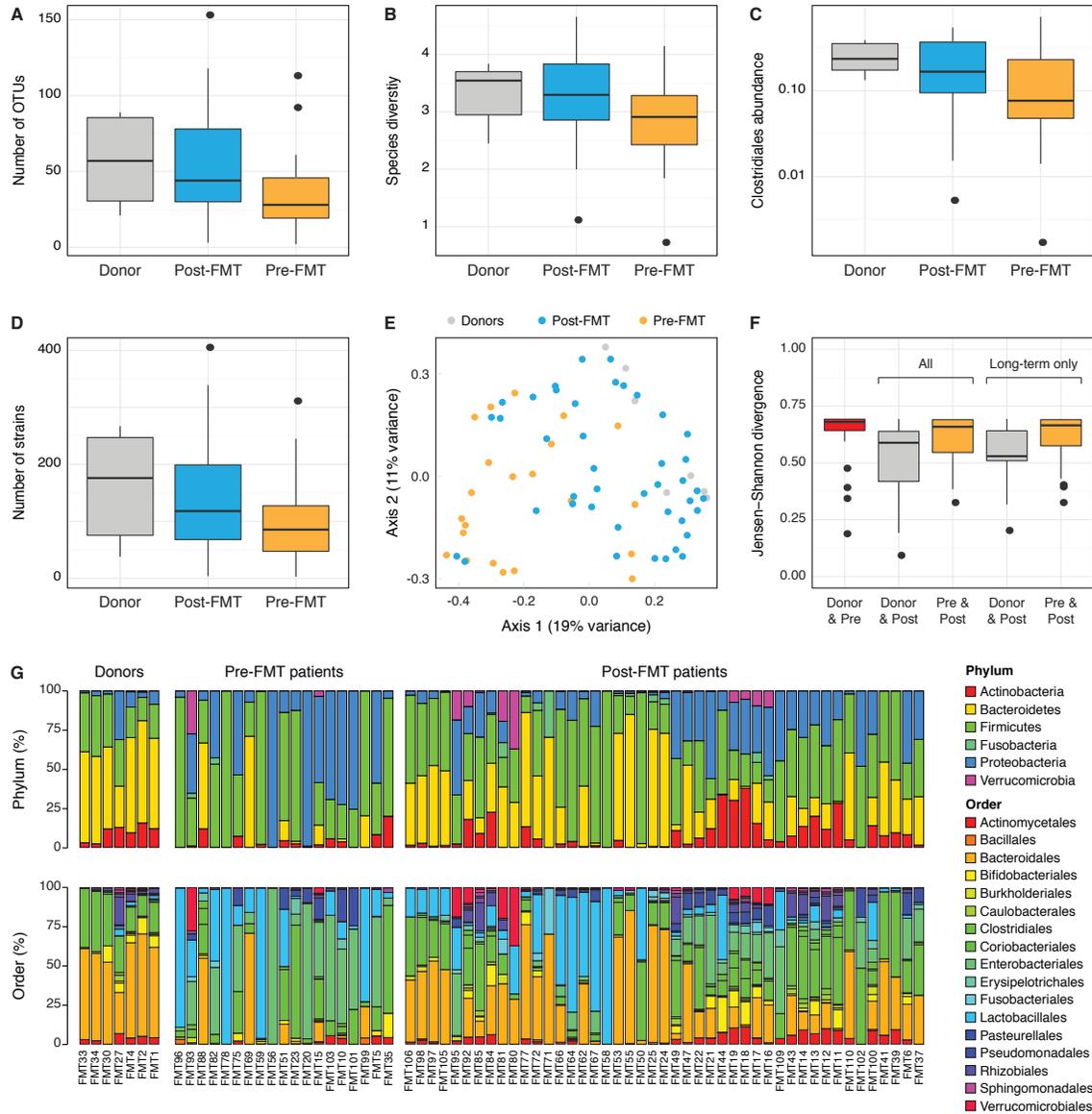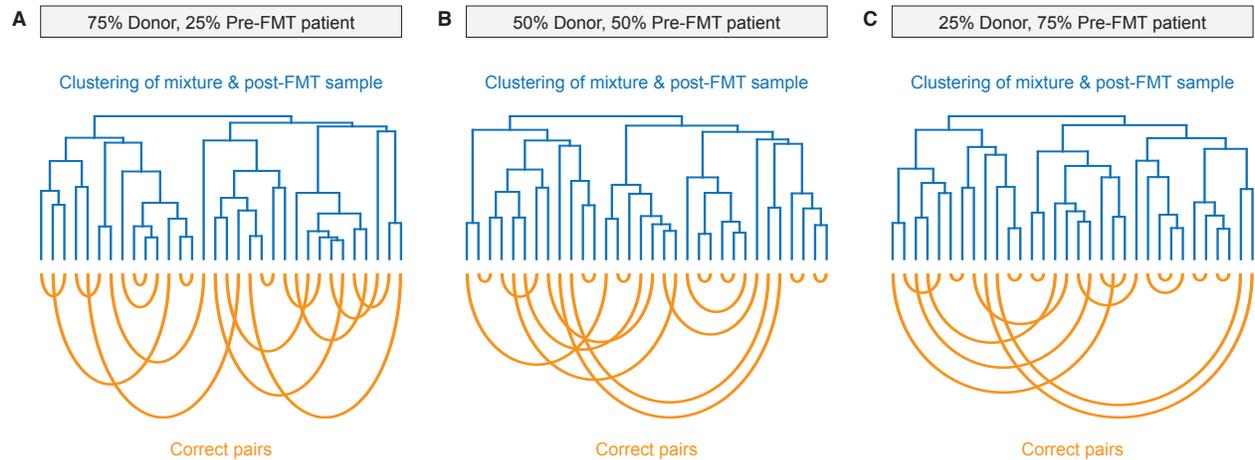# Supplemental Figures

Figure S1, Related to Figure 2: Taxonomic overview of metagenomics samples. Boxplots showing the (A) mg-OTU richness, (B) species diversity, (C) abundances of Clostridiales, and (D) strain richness for samples from the donor, pre-FMT patient, and post-FMT patient. (E) Principal coordinates analysis of the weighted UniFrac distances among all samples. Axes are labeled with the percent of variance they explain. (F) Jensen-Shannon divergence among all samples from the donor, the pre-FMT patient, and the post-FMT patient. After FMT, the patient is more similar to the donor than to itself before the FMT (Wilcoxon test; p-value = 2e-4). (G) Relative abundances of bacterial phyla and orders across all samples.

| A | 75% Donor, 25% Pre-FMT patient | B | 50% Donor, 50% Pre-FMT patient | C | 25% Donor, 75% Pre-FMT patient |

Clustering of mixture & post-FMT sample (A)

Clustering of mixture & post-FMT sample (B)

Clustering of mixture & post-FMT sample (C)

Correct pairs

Correct pairs

Correct pairs

Figure S2, Related to Figure 2: Simple mixtures of the donor samples and the pre-FMT patient samples do not predict the post-FMT gut microbiota. Weighted averages of the mg-OTU abundances from the donor and the pre-FMT patient were constructed for a range of mixing coefficients, including **(A)** 75% of the donor samples and 25% of the pre-FMT patient samples, **(B)** 50% of the donor samples and 50% of the pre-FMT patient samples, and **(C)** 75% of the donor samples and 25% of the pre-FMT patient samples. These mixtures do not cluster with the corresponding samples from the post-FMT patients.
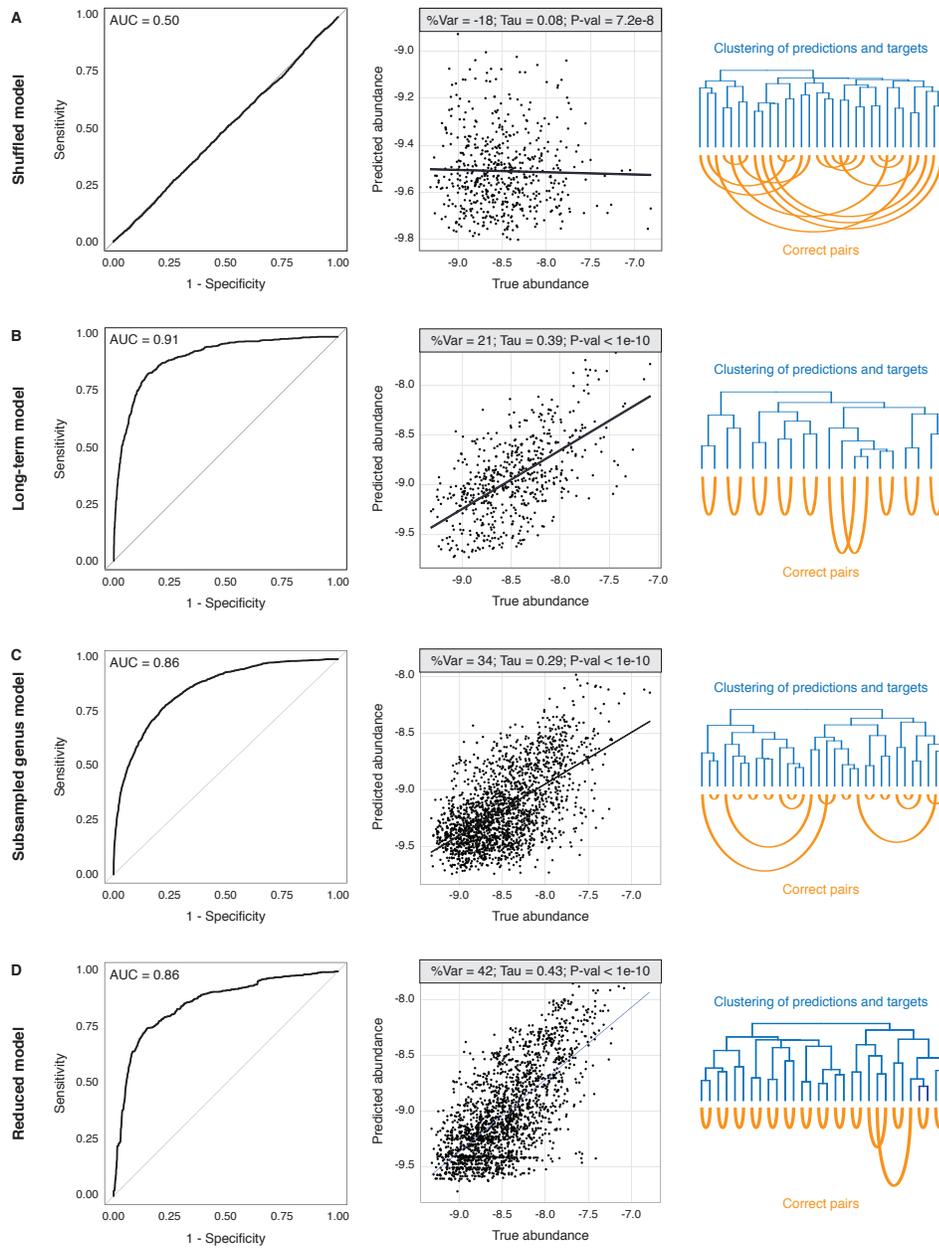
Figure S3, Related to Figure 2: Validation of the model of bacterial engraftment. The model of bacterial engraftment was retrained on (A) shuffled data, (B) long-term follow-up samples only (i.e. samples collected over 14 days after the FMT), (C) a subsampled dataset containing only one bacterial species per genus, and (D) a reduced feature set including only the mg-OTU abundances, the mg-OTU taxonomy, elapsed time, and sequencing depth. For all validations, the ROC curve, a plot of predicted abundances against true abundances, and clustering dendrograms of the predicted microbiota composition and the true microbiota composition are provided.
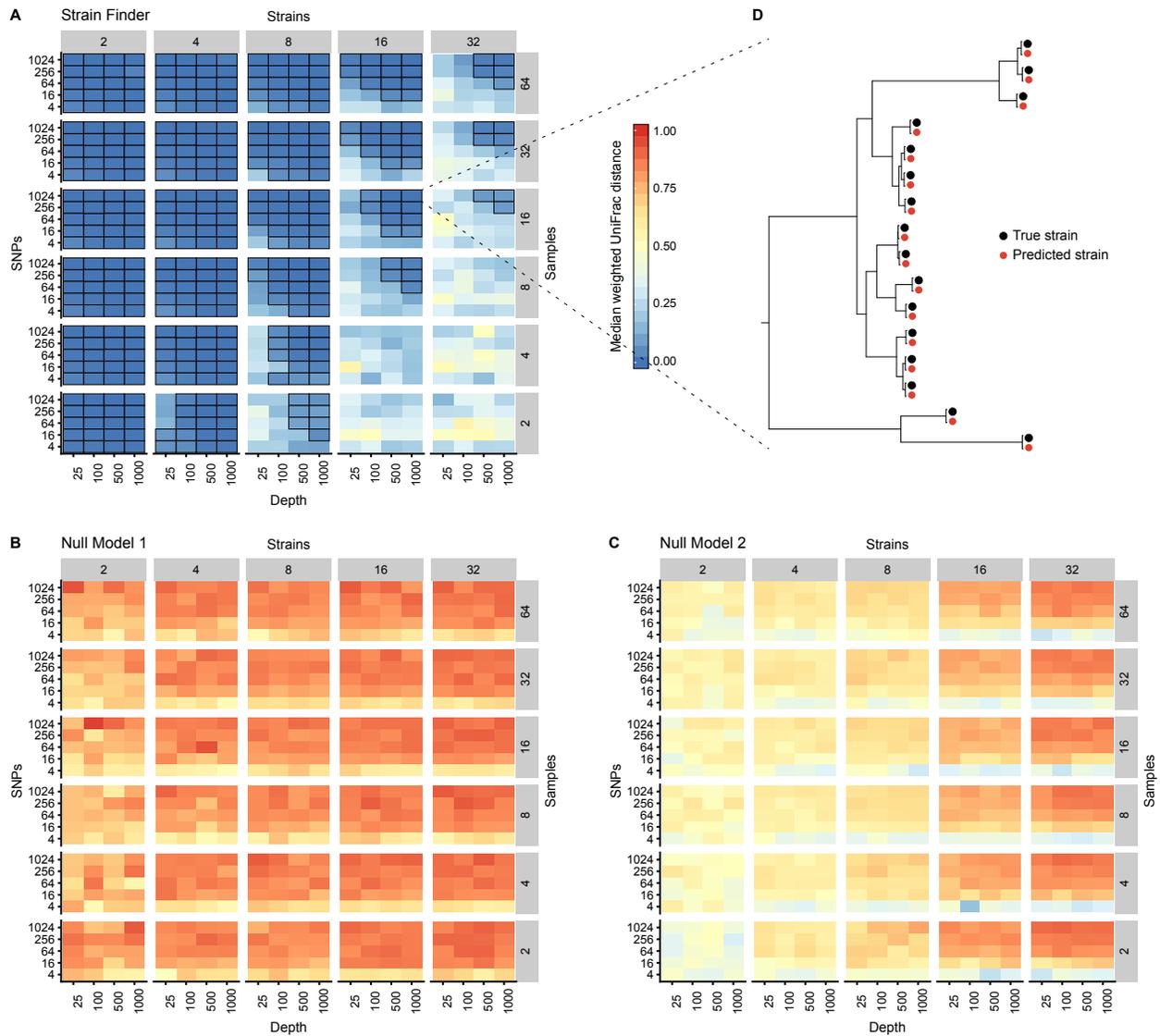
Figure S4, Related to Figure 4: Strain Finder outperforms null models with randomly selected strain genotypes and frequencies. Median weighted UniFrac distances between the inferred strain profiles and the true strain profiles for **(A)** Strain Finder, **(B)** Null Model 1, and **(C)** Null Model 2. Both null models use randomly selected strain frequencies, with the strain genotypes sampled from a discrete uniform distribution (Null Model 1) or from the alignment data itself (Null Model 2). Cases for which the median weighted UniFrac distance is less than 0.1 are outlined in black. **(D)** Strain phylogeny showing the relationships among the true strains and those predicted by Strain Finder, for an example with 16 strains, 16 samples, 1024 SNPs, and 1,000X depth of coverage. The Strain Finder predictions cluster perfectly with their targets.
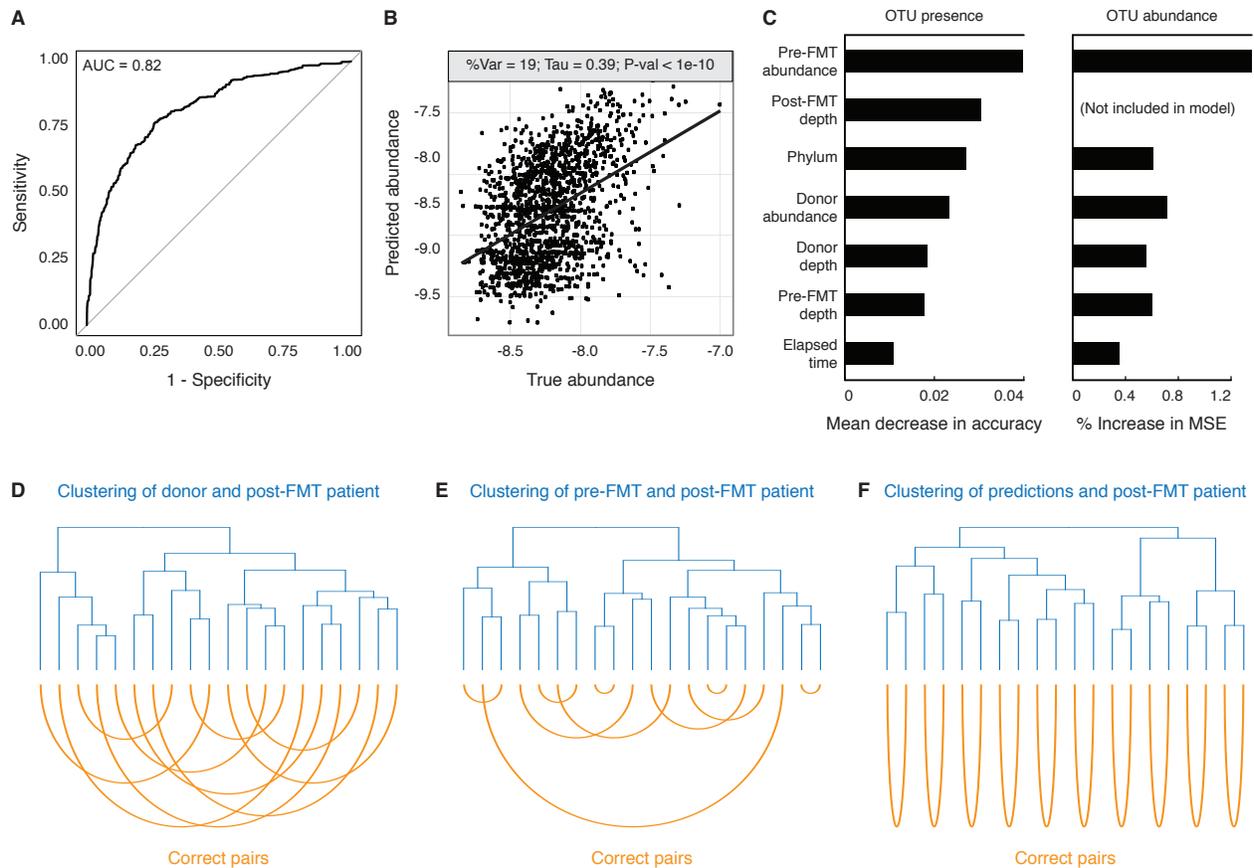
Figure S5, Related to Figure 7: The model of engraftment extends to the treatment of metabolic syndrome. The reduced model of engraftment was retrained on a metagenomics dataset for the treatment of metabolic syndrome with FMT. **(A)** The ROC curve for our predictions of bacterial presence in the post-FMT patient. **(B)** The predicted log10 mg-OTU abundances in the post-FMT patient are significantly correlated to their true values. **(C)** The relative importance of each feature in the models of bacterial presence and abundance. **(D)** Donor samples do not cluster with the correct post-FMT samples. **(E)** Pre-FMT patient samples do not cluster with the correct post-FMT patient samples. **(F)** Predictions cluster perfectly with the correct post-FMT samples.
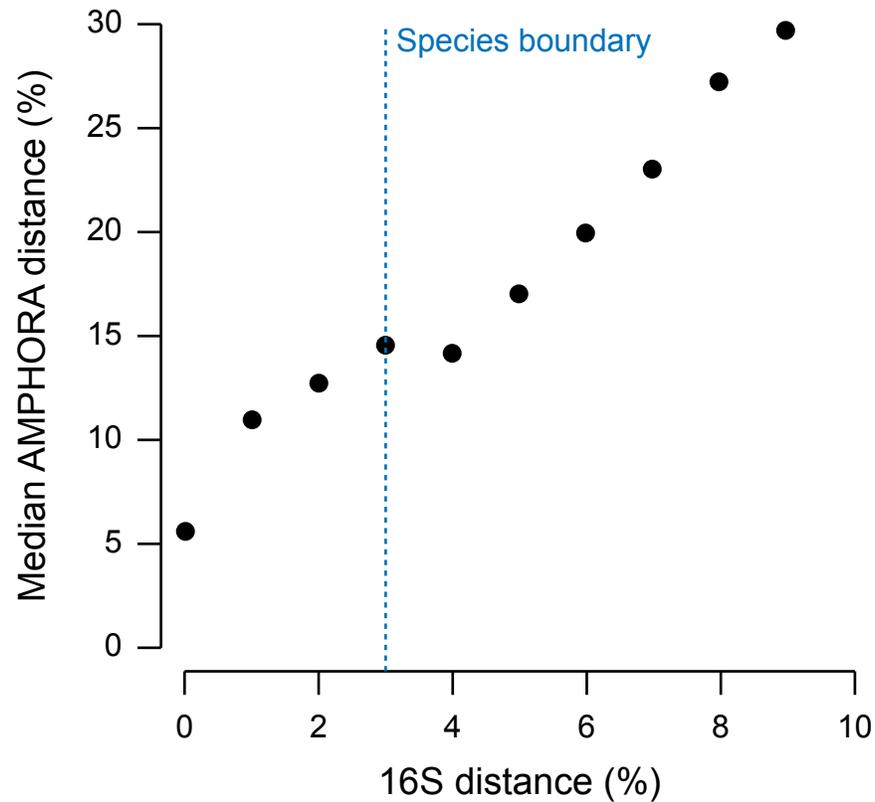
Figure S6, Related to STAR Methods: Comparison of genetic distances at the 16S rRNA gene and the AMPHORA genes. We estimated the genetic distances among all pairs of reference genomes at the 16S rRNA and AMPHORA genes. To ensure that our alignments correspond to bacterial strains, rather than species, we conservatively selected an alignment cutoff of 10% distance at the AMPHORA genes, well below the species boundary of 3% divergence at the 16S rRNA gene.
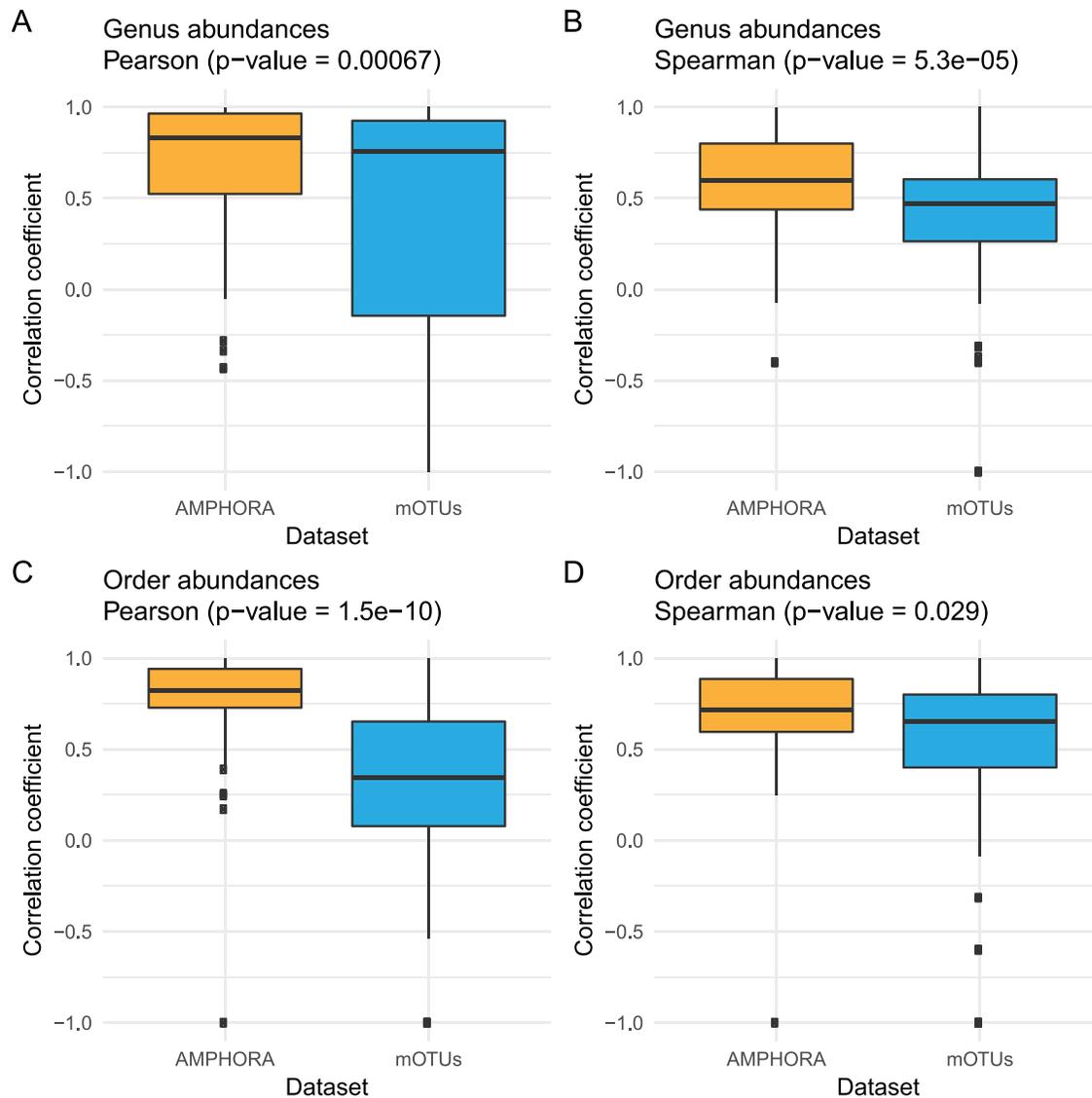
Figure S7, Related to STAR Methods: Comparison of abundances estimated by the 16S rRNA data with the AMPHORA and mOTUs pipelines. We compared the AMPHORA and mOTUs abundances estimates with the abundances estimated from the 16S rRNA data, at the genus and order levels. **(A)** Boxplot of Pearson correlation coefficients at the genus level. **(B)** Boxplot of Spearman correlation coefficients at the genus level. **(C)** Boxplot of Pearson correlation coefficients at the order level. **(D)** Boxplot of Spearman correlation coefficients at the order level. In all cases, the correlations of the 16S rRNA abundances with the AMPHORA abundances were significantly higher than their correlations with the mOTUs abundances (one-sided Wilcoxon signed-rank test, p-values = 7e-4, 5e-5, 2e-10, and 3e-2, respectively).