

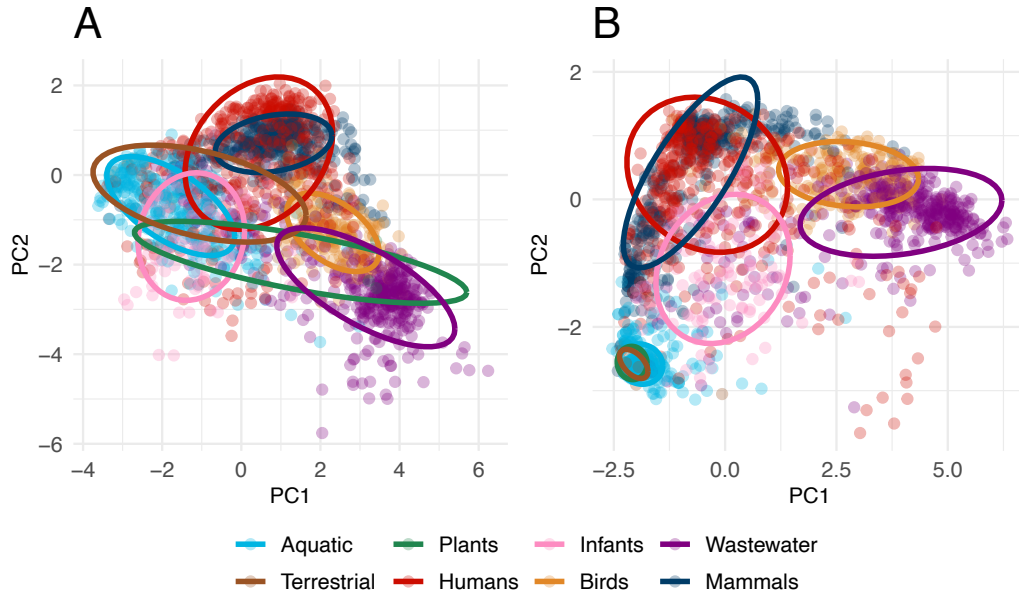


Figure S1: Principal component analysis (PCA) of the α -diversity of A) latent and B) established ARGs. Aquatic includes samples from fresh, lentic, and marine water; Wastewater includes activated sludge, water and sludge, fecal source, and raw wastewater; Birds includes samples from their digestive system; Mammals includes samples from the digestive systems of bovines, mice, and pigs; Human includes samples from the skin, digestive and respiratory systems; and Infants include samples from their digestive system. The PCA was done using rarefied gene counts of all the metagenomic samples, but a maximum of 400 samples per environment are shown. The ellipses are drawn from a multivariate t-distribution at a 75% confidence level. The aquatic, terrestrial, and plant environments are more widely spread for the latent ARGs, while they are gathered together for the established ARGs. In both PCAs, the human/mammal metagenomes form a cluster, while birds and wastewater are separated from the other environments.

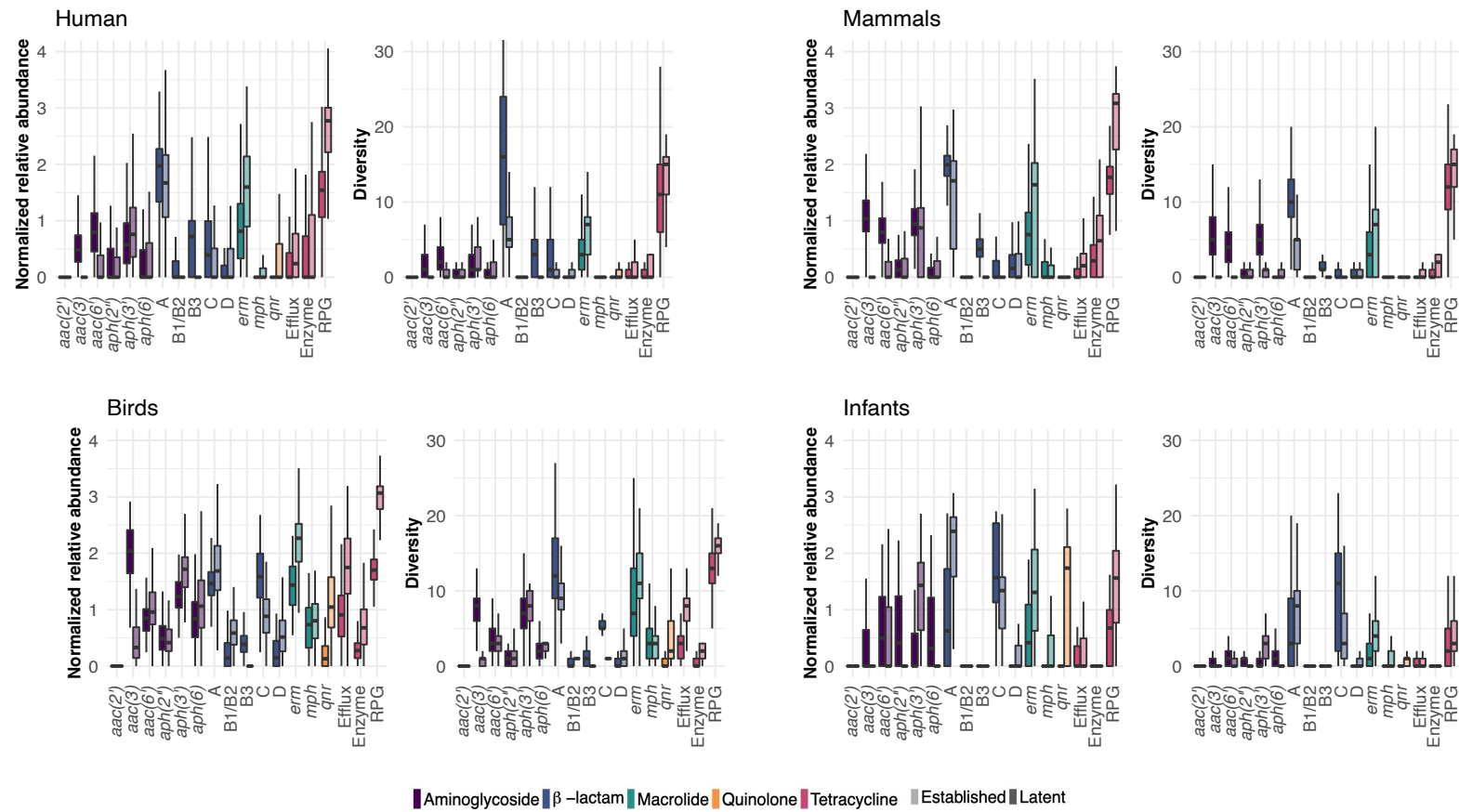


Figure S2: Distribution of the log-transformed abundance and α -diversity for latent and established ARGs for each gene class and environment. RPG is short for ribosomal protection gene.

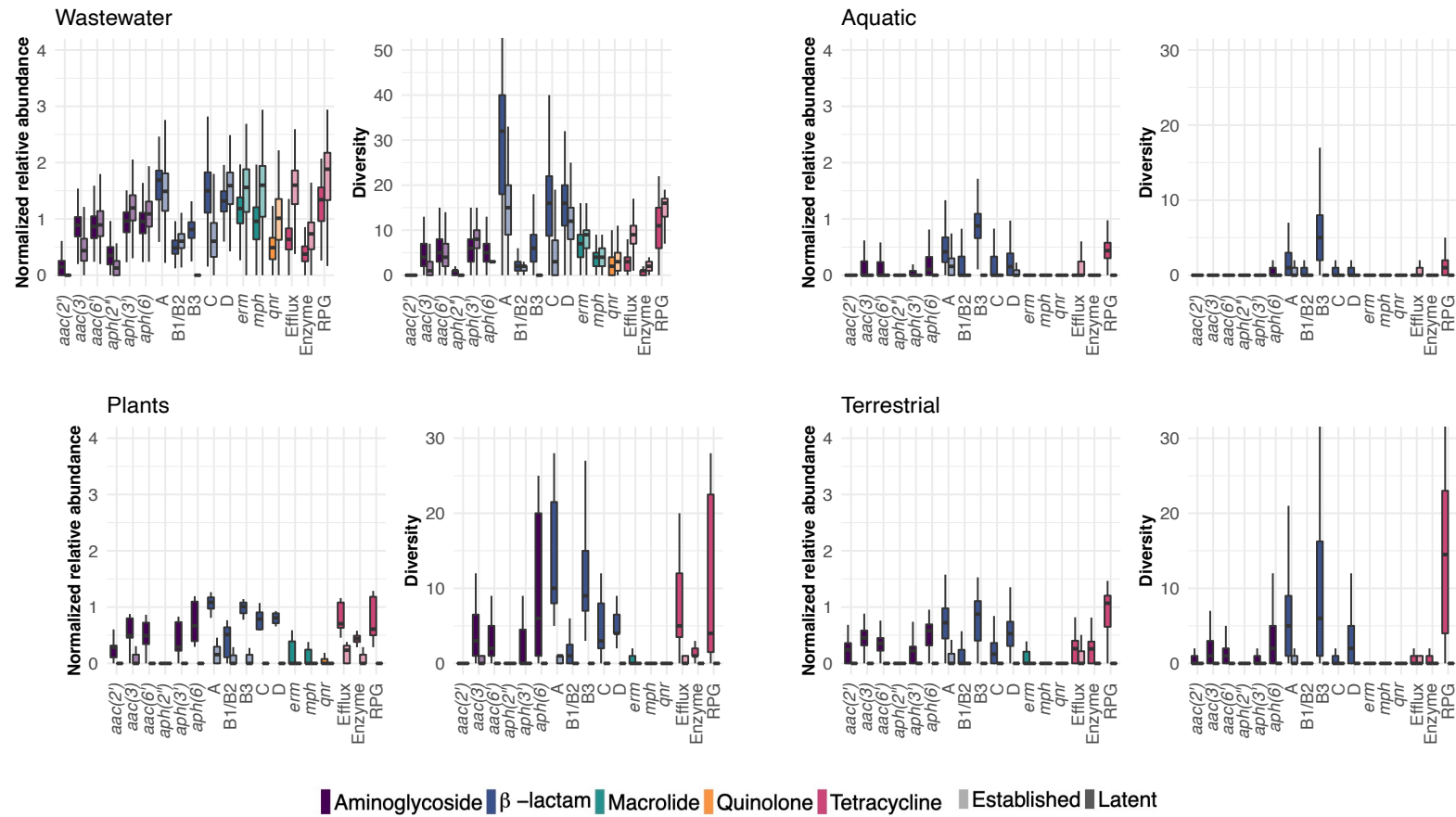


Figure S3: Distribution of the log-transformed abundance and α -diversity for latent and established ARGs for each gene class and environment. RPG is short for ribosomal protection gene.

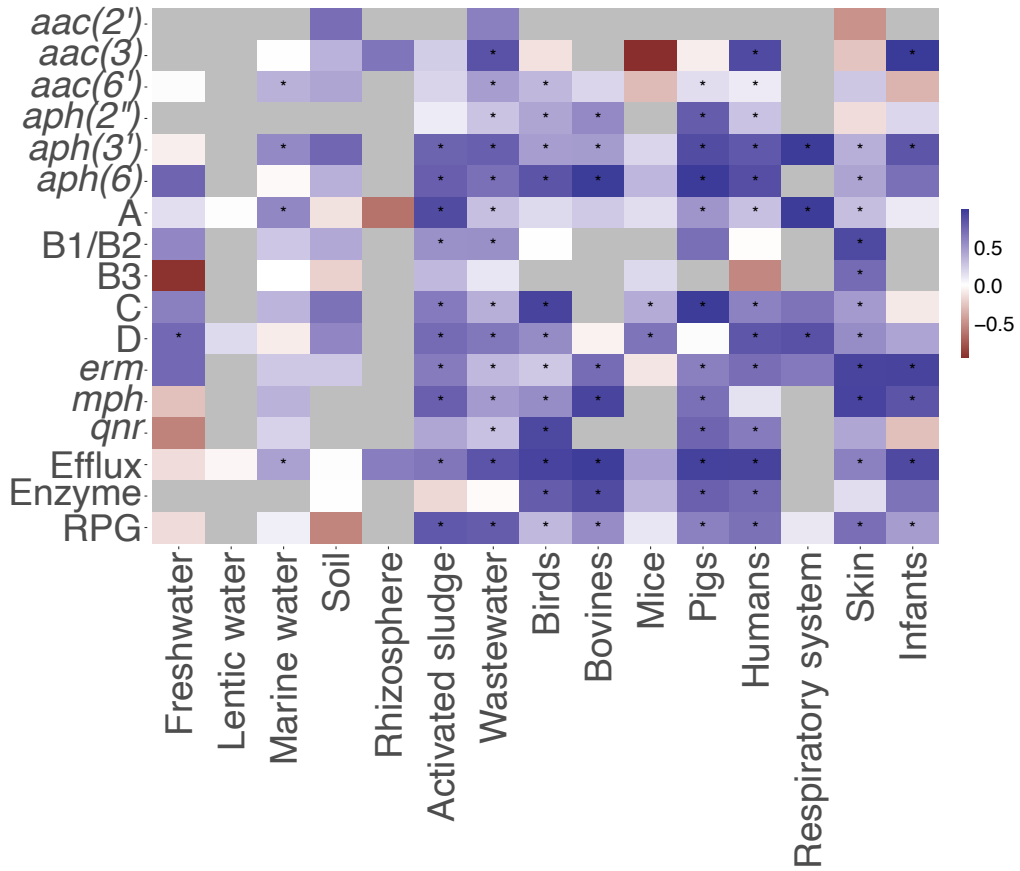


Figure S4: Pearson correlation between the latent and establish ARGs divided into gene class and environment. The labels Birds, Bovines, Mice, Pigs, Humans, and Infants denotes metagenomes from the corresponding digestive system. Respiratory system and skin only includes human samples. The color intensity reflect the size of the estimated correlation coefficient and a star (*) marks significant correlations ($p < 0.001$). Gray squares indicate environment and gene classes with insufficient number of non-zero observations to calculate the correlation coefficient and/or the p-value.