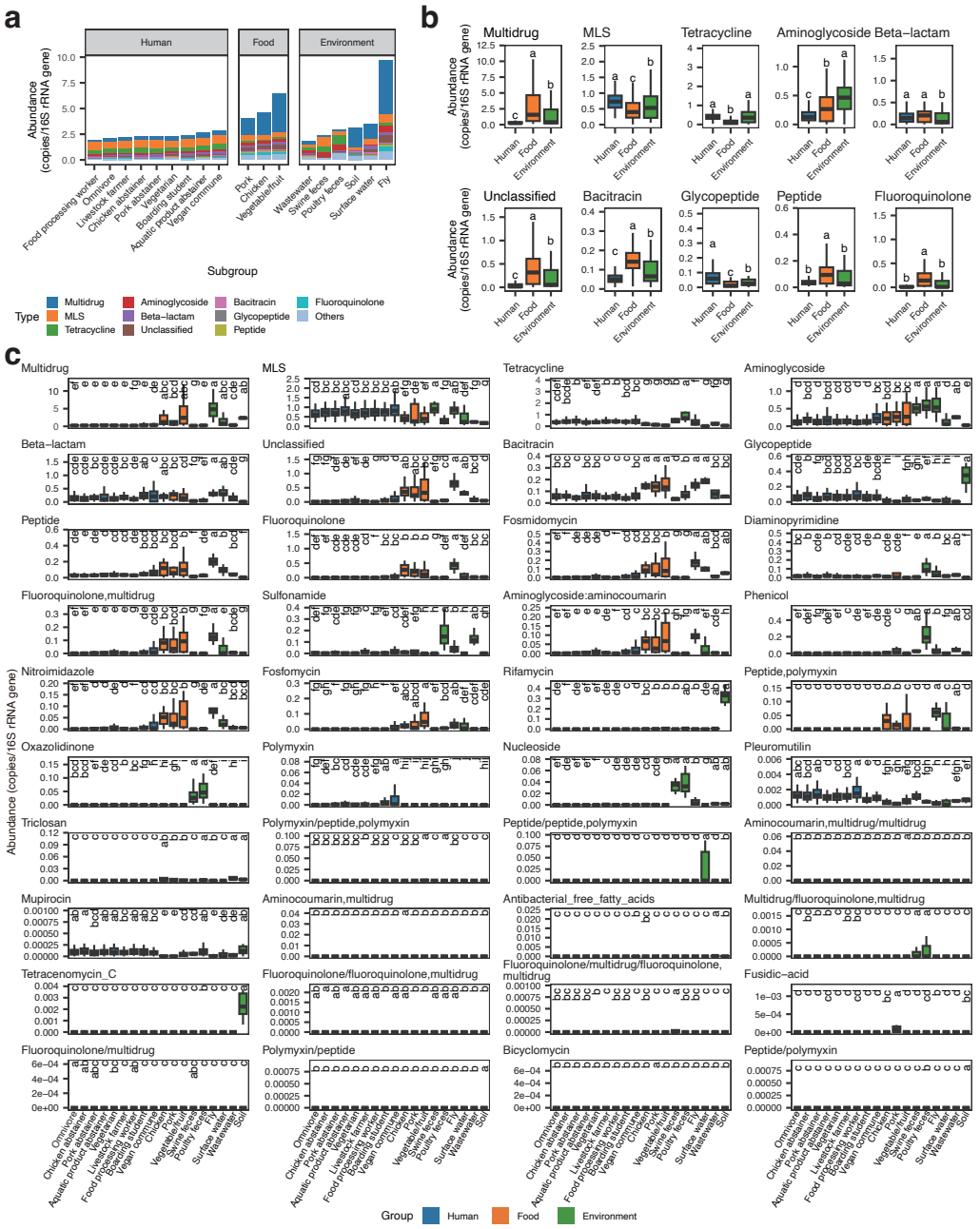

Fig. S1: The abundance of the ARG types. (a) Average abundance of the main ARG types in various subgroups. (b) The abundance of the top 10 ARG types in the three groups. (c) The abundance of the 40 ARG types in all subgroups. Left panels: the ARG types (n = 20) with relative higher abundance; right panels: the ARG types (n = 20) with relative lower abundance.

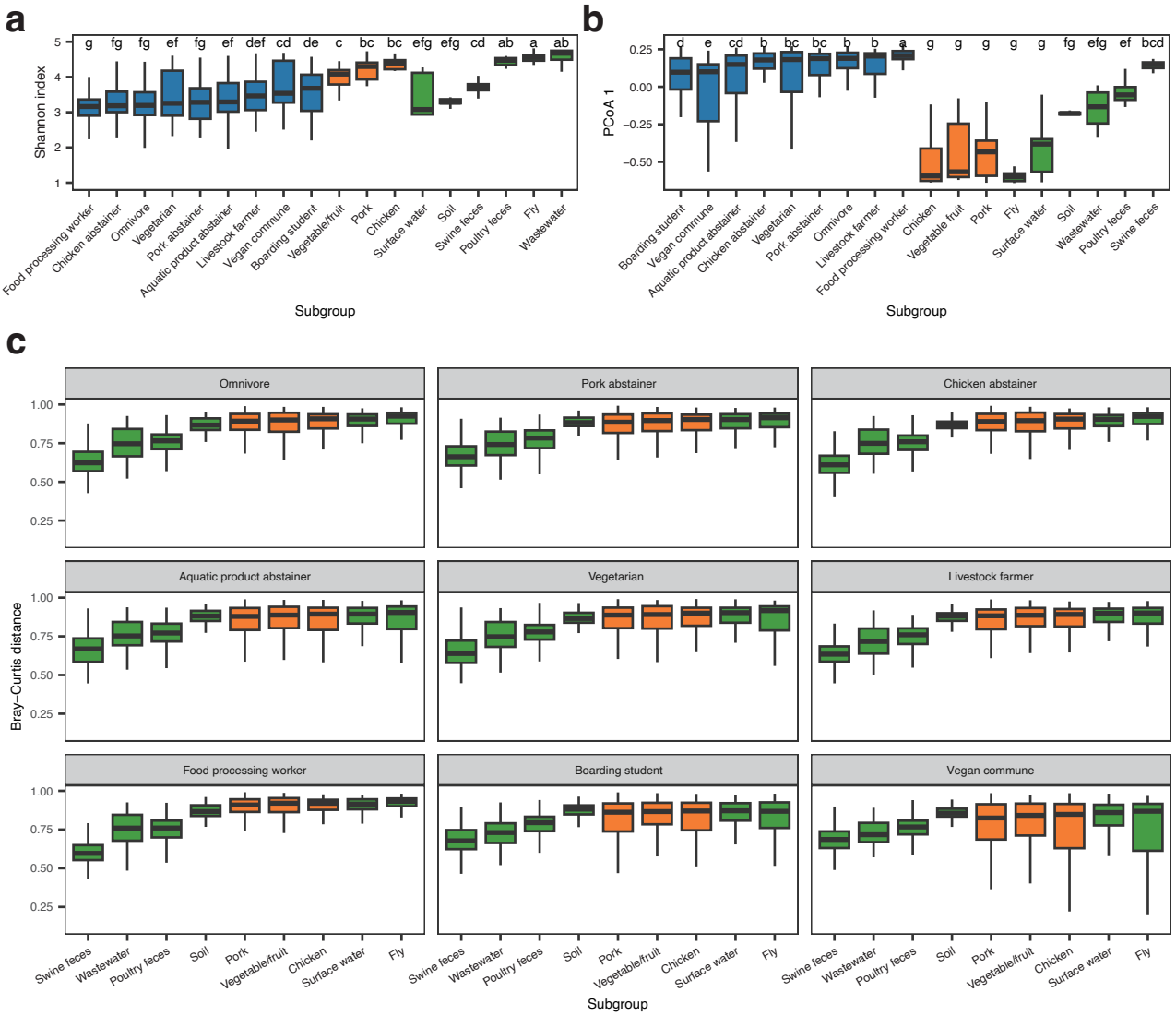

Fig. S2: Diversity analysis of the resistome profiles. (a) The Shannon indices of the resistome profiles. (b) Axis 1 values of PCoA of the resistome profiles from various subgroups. (c) Bray–Curtis distances between the resistome profiles of various human fecal samples and those from food and environmental samples.

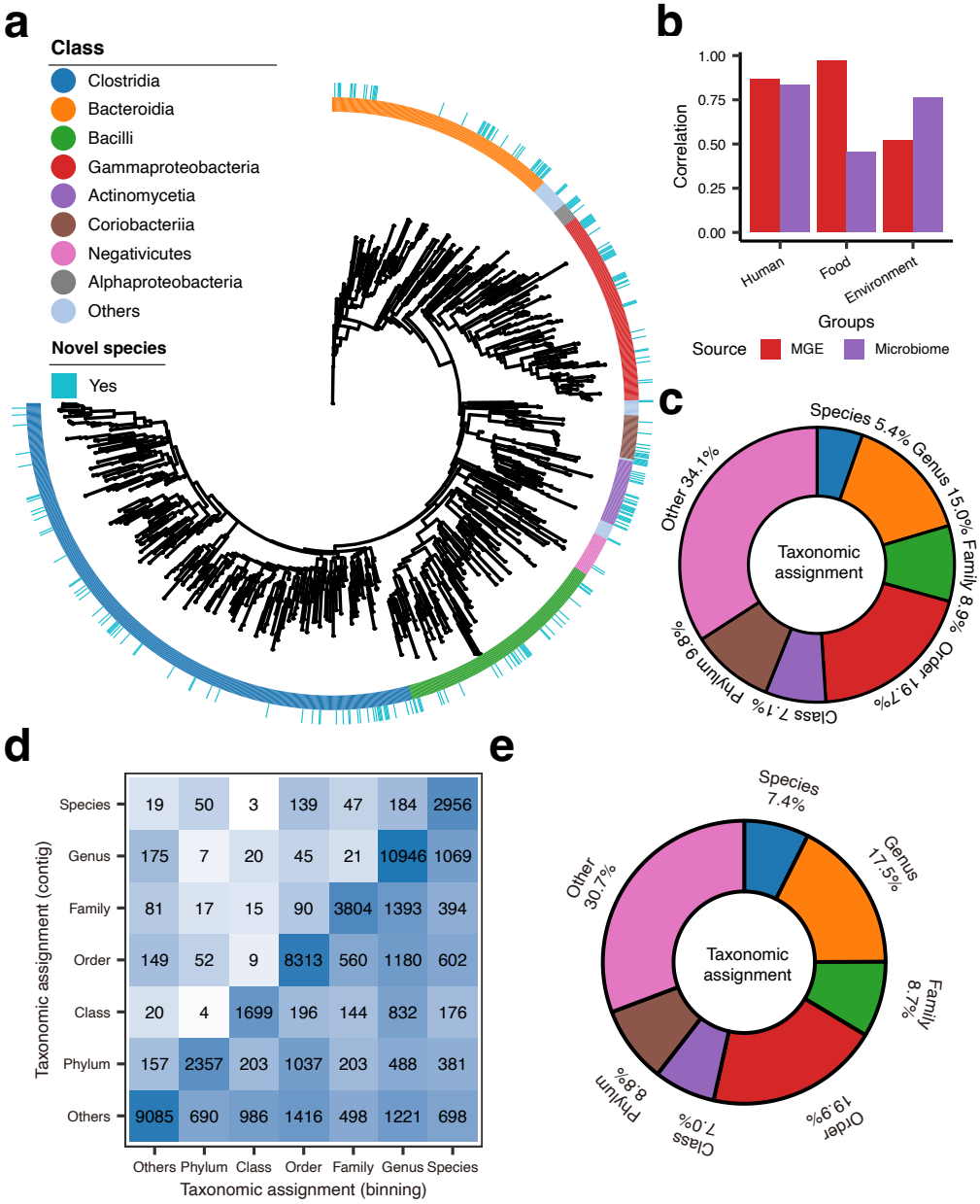

Fig. S3: Metagenome-assembled genomes and ACC taxonomic assignment. (a) Phylogenetic tree of the 1,302 MAGs at the species level. A total of 14,787 MAGs (completeness ≥ 70.0% and contamination ≤ 10.0%) were constructed in this study. The novel species are labeled with light blue bars in the outer ring. The color of the inner ring represents the class level of the MAG. (b) Procrustes analysis of the correlations between the resistome and both MGEs and the microbiome in the three groups. The correlations were calculated by the function "protest" in R package vegan. (c) Proportion of taxonomic assignment of the ACCs by the software CAT. CAT exploits homology searches of individual open reading frames to classify ACCs directly. (d) The taxonomic assignment performance of the two methods based on the 54,831 ACCs, which were binned into MAGs out of 162,001 ACCs. (e) Proportion of the taxonomic assignment of the ACCs with the integration of taxonomic information of MAGs.

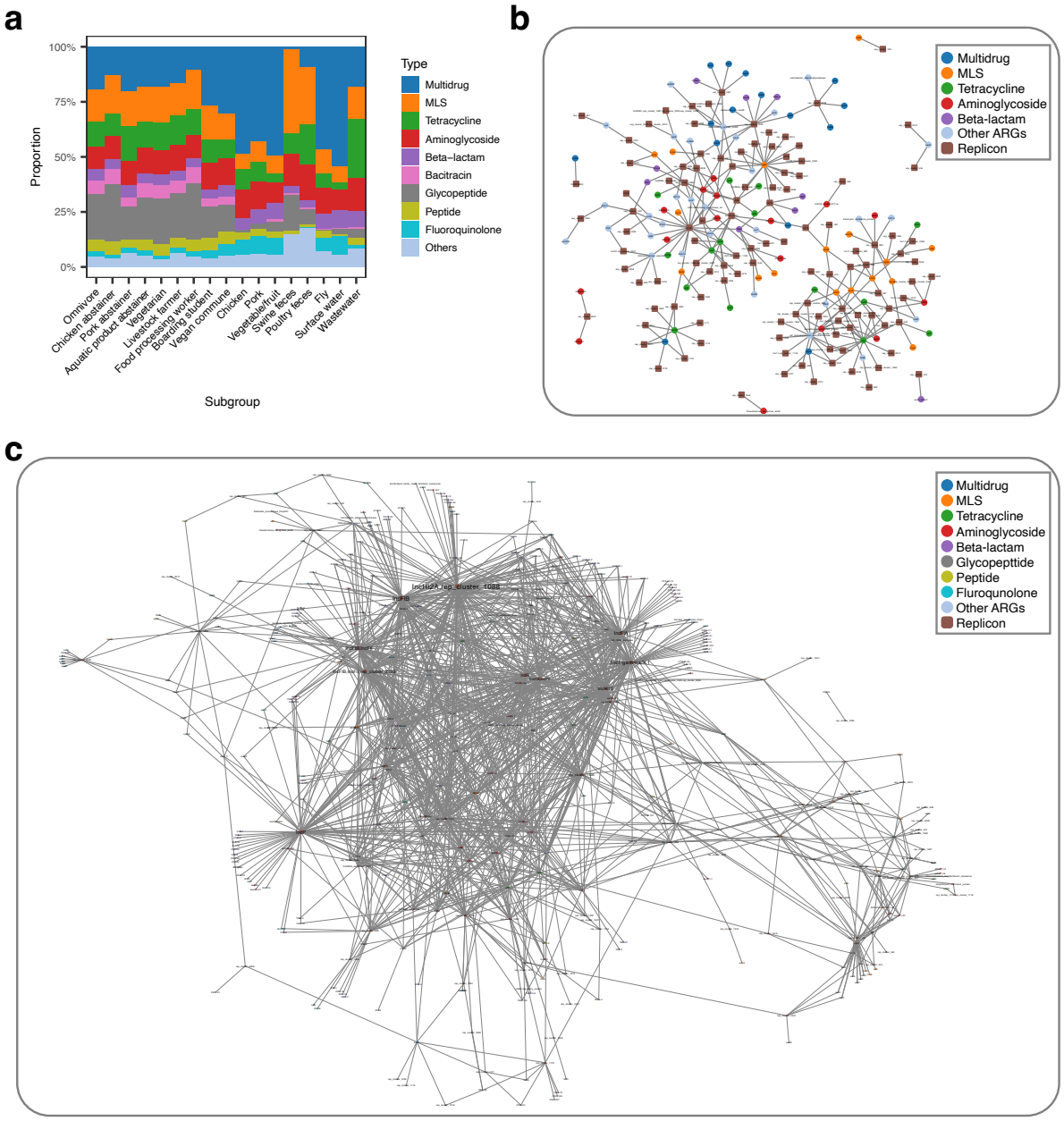

Fig. S4: Plasmid-associated ARGs. (a) The proportion of ARG types carried by plasmids in the different subgroups. (b) Co-occurrence network between plasmid replicon types and the ARGs revealed by the ACCs. The brown rounded rectangles represent the plasmid replicons, the colored circle nodes represent the ARG types, and the gray lines represent the associations between the ARGs and the plasmid replicons. (c) An expanded association between the ARGs identified in (b) and the plasmid types identified by integrating the results from the PLSDB database. The brown squares represent the plasmid replicons; the colored circles represent the ARG types; and the gray lines represent the associations between the ARGs and the plasmid replicons.

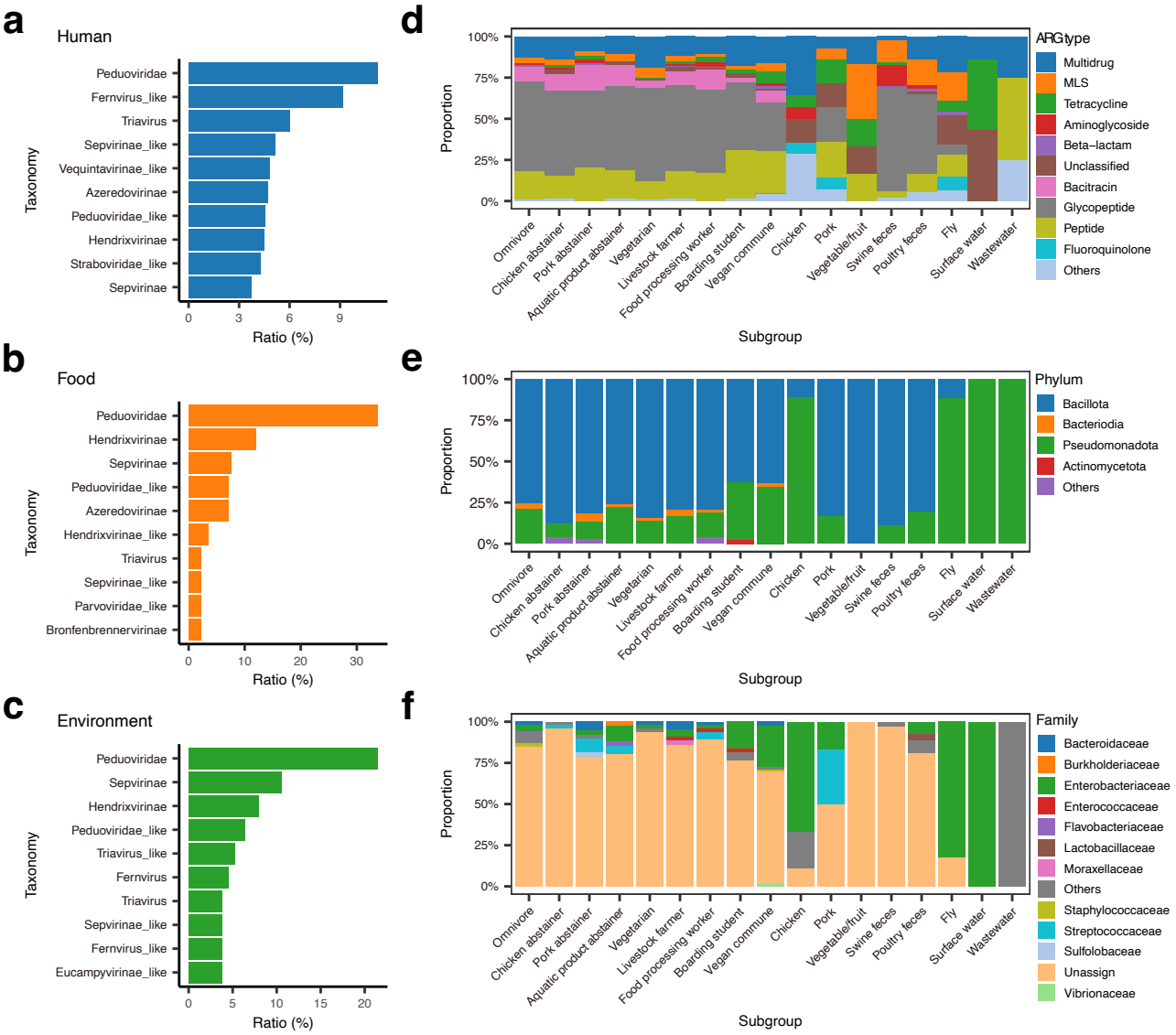

Fig. S5: Analysis of the prophages/phages identified from ACCs. The ratio of different taxa of prophage/phage identified in samples from human (a), food (b), and environmental (c) sources. (d) Proportion of ARGs carried by phage/prophage fragments. Host prediction of prophage fragments at the phylum (e) and family (f) levels.

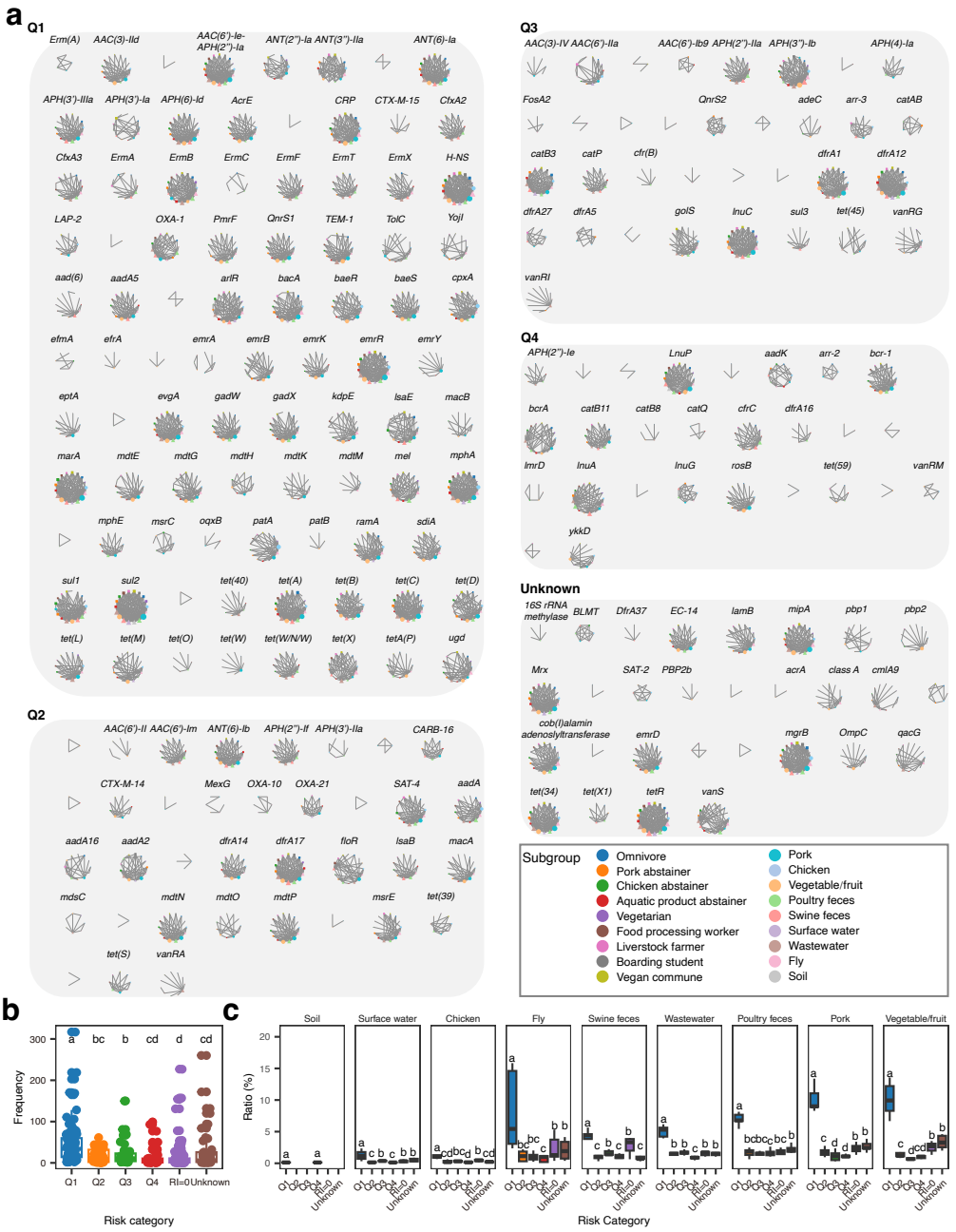

Fig. S6: Analysis of the ARG-sharing events. (a) ARG-sharing network. A cluster of shapes represents an ARG subtype shared among certain subgroups. The color of each node indicates the subgroup, while the different shapes represent human, food, and environmental samples. The size of the circle corresponds to the frequency of shared events. (b) Frequency of ARG-sharing events in the different risk ranks. Unknown represents the ARGs without a classification. (c) The ratio of ARGs found in food and environmental subgroups shared with human fecal groups.

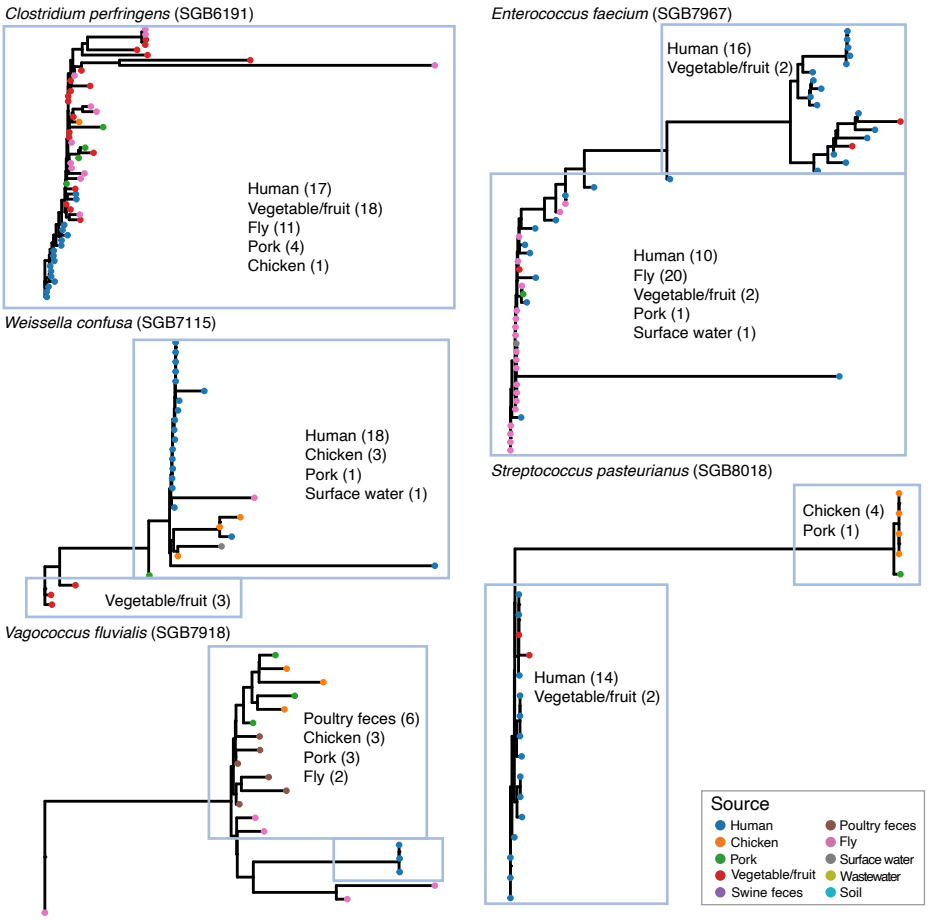

Fig. S7: Strain-level phylogenies of species present in all 592 samples. Each node represents the main strain of a sample. Different colors represent various subgroups. Light blue rectangles indicate phylogenetic trees of the same strain (nGD < 0.1). The annotations are the number of samples harboring the same strain.

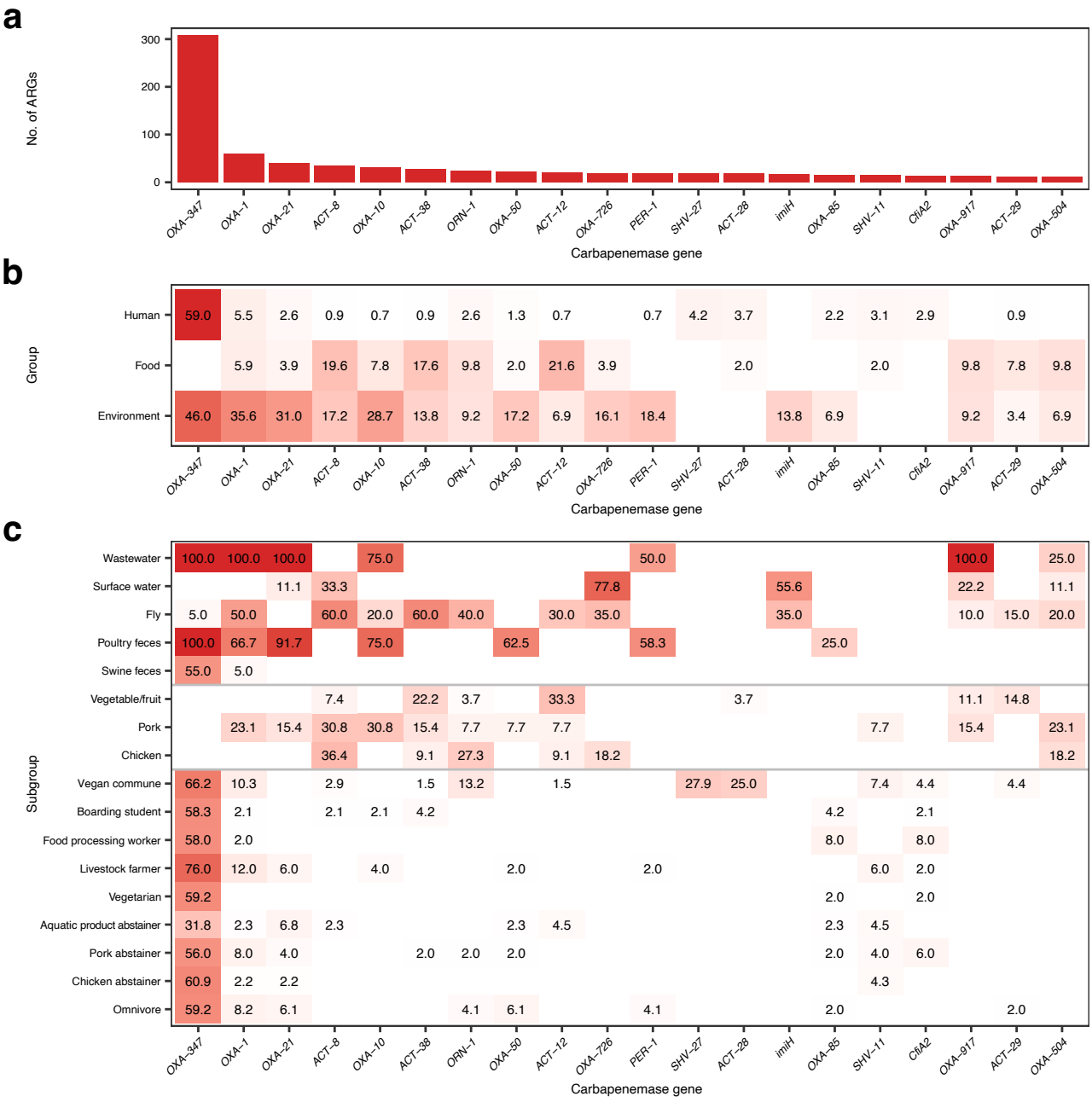

Fig. S8: Distribution of carbapenemase genes found in the metagenomic sequencing data. (a) The prevalence rates of the top 20 carbapenemase genes in all 592 samples. The ratio of positive samples for carbapenemase genes in the three groups (b) and the 17 subgroups (c). The numbers labeled on cells represent the ratio of positive samples for each carbapenemase gene.

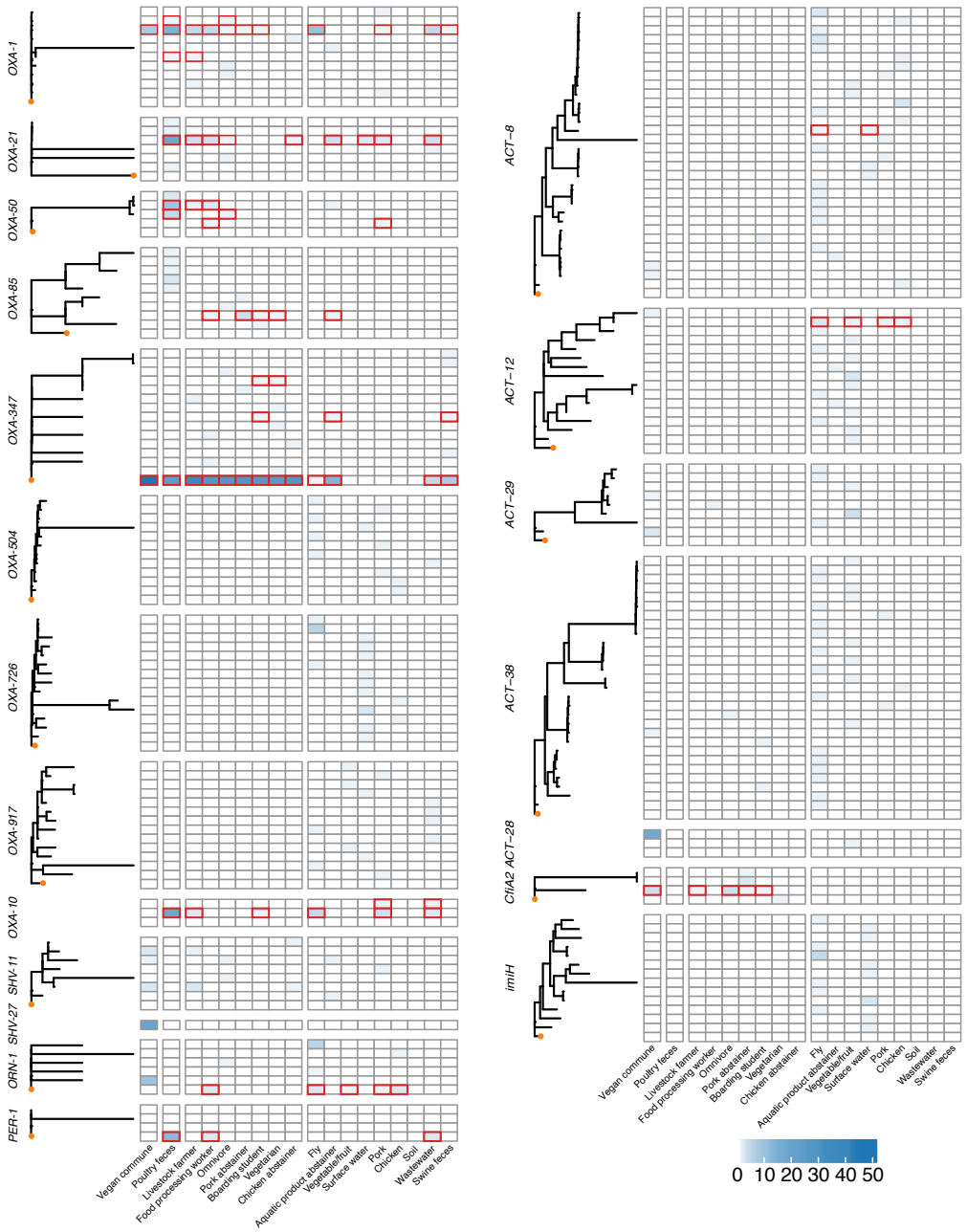

Fig. S9: Phylogenetic trees of the top 20 carbapenemase genes. In the phylogenetic tree, each branch represents a sequence type of the carbapenemase gene. Orange circle nodes in the phylogenetic trees represent the reference carbapenemase gene sequences from CARD. Heatmap illustrates the prevalence of each sequence type in the different subgroups, and the color of the cells reflects the number of positive samples. The red border of each cell represents the sharing events of the carbapenemase gene across the different subgroups.

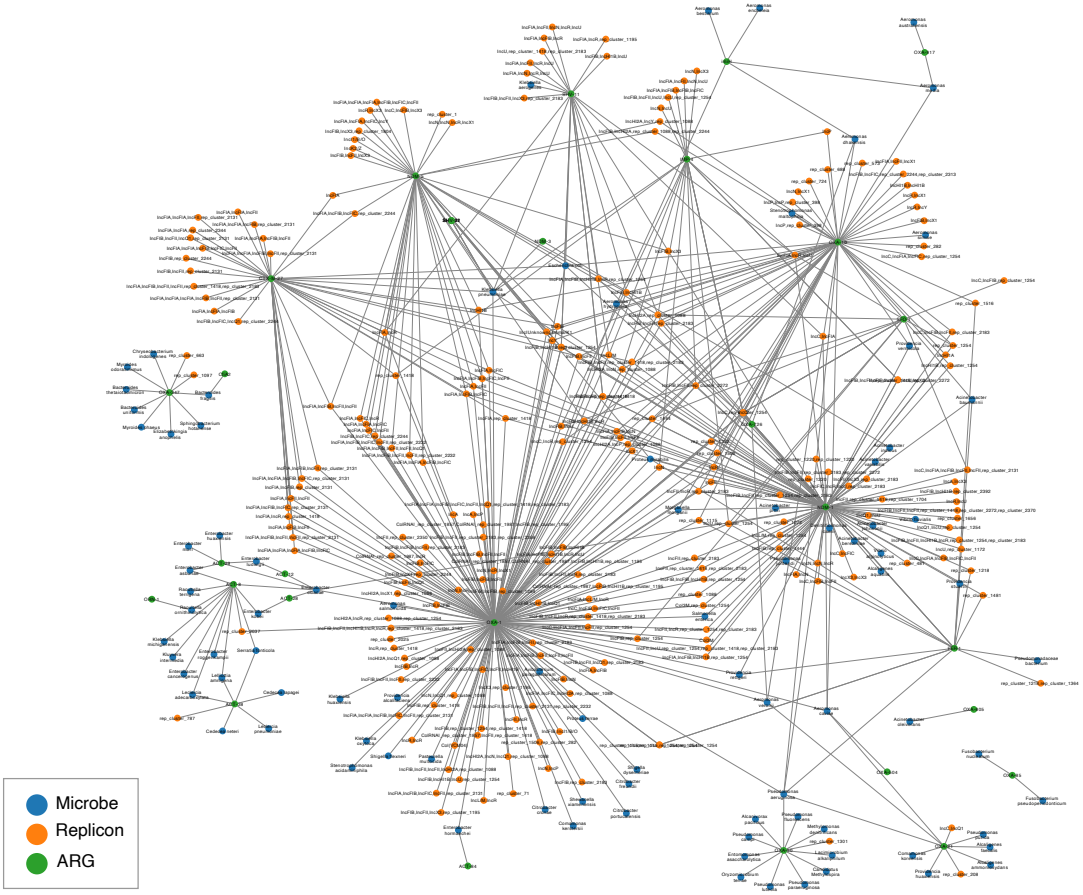

Fig. S10: Co-occurrence network of carbapenemase genes and their associated hosts identified in complete genomes. The carbapenemase genes includes those from both metagenomic sequencing data and the strain genomic data in our study. The blue nodes represent microbes; the orange nodes represent plasmid replicons, the green nodes represent the ARGs, and the gray lines represent the locations of the ARGs.