

Description of Additional Supplementary Files:

Supplementary Data 1: List of genomes considered for the analyses (prokaryotes and eukaryotes) and associated metadata obtained from GTDB (for prokaryotes).

Supplementary Data 2: KEGG pathways enriched per subcategory for the core OGs and the ASPs

Supplementary Data 3: Events of tandem duplications

Supplementary Data 4: KEGG pathways enriched in the tandem duplications

Supplementary Data 5: Main output statistics for the HGT pipeline for the 14 seed genomes