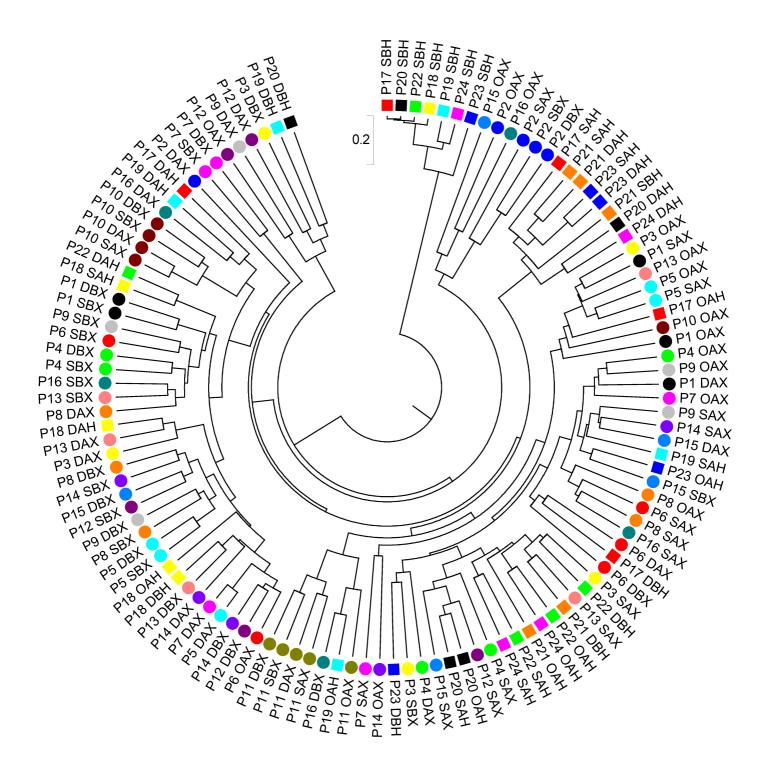
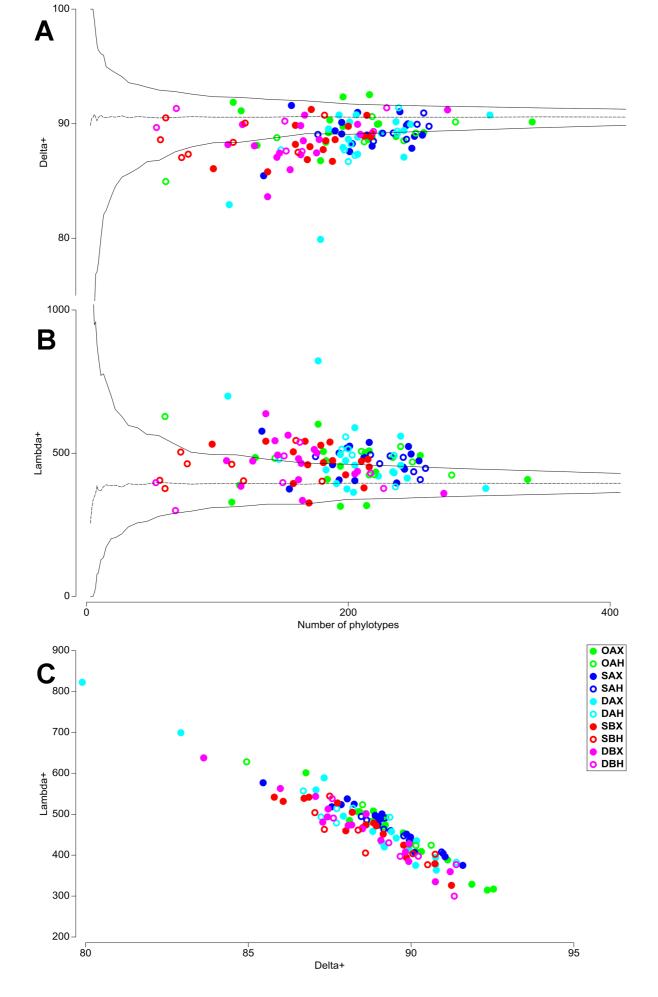


Supplementary Figure 1: Rarefaction curves portraying the number of resolved phylotypes against sequencing depth for each sample. Samples from individuals infected with *H. pylori* are shown in red, those of non-infected individuals in black.



Supplementary Figure 2: Group-average agglomerative hierarchical clustering analysis of bacterial communities along the upper gastrointestinal tract at the genus level. The Bray-Curtis algorithm was used to assess the similarity between samples based on the relative abundance of genera. The sample regions are saliva (O), stomach (S), duodenum (D) and sample types are aspirates (A), biopsies (B), and *H. pylori* negative (X) or *H. pylori* positive (H). The 24 individuals are indicated by a unique symbol.



Supplementary Figure 3: Taxonomic distinctness and evenness of microbial communities along the gastrointestinal tract. Funnel plots chart the (A) average taxonomic distinctness (delta+) and (B) variation in taxonomic distinctness (lambda+) against the number of phylotypes within each sample. The funnel indicates the limits within 95% of the simulated TD values and the middle line represents the mean expected TD. (C) Charts average taxonomic distinctness (lambda+) against variation in taxonomic distinctness (delta+)..