Figure S2   Distribution of valid sequences for BPHD-f3 (outer ring) and BPHD-r1 (inner ring). Valid sequences and 467 reference sequences were clustered into 56 and 20 clusters based on 0.6 distance cutoff for BPHD-f3 sequences and BPHD-r1 sequences, respectively. The clusters are composed of (i) only valid sequences (blue background), (ii) both valid sequences and reference sequences (yellow background), and (iii) only reference sequences (pink background). The bar height indicates the number of valid sequences in each cluster. The clusters with less than 15 valid sequences are enlarged in white sections. The frequency distribution of unique sequences within the largest three clusters are indicated as histograms.