

Figure S1: Histogram of bootstrap support values for interpatient patristic (tip-to-tip) distances. In total, 77,164 interpatient patristic distances (0.25% of all possible pairwise comparisons) fell below the cutoff of 0.02 expected nucleotide substitutions per site in one or more replicate phylogenies which were generated from nonparametric bootstrap samples of the sequence alignment. The histogram was rescaled to elucidate differences among bins; the leftmost bin contained 39,009 distances with bootstrap supports below 5%. We used a cutoff of 50% to determine which patristic distances would be used to generate phylogenetic clusters. All statistical associations between phylogenetic clustering and clinical, demographic, or behavioural factors (Table 1) were robust to varying this cutoff.

