

Fig. S1. Classification of the isolates into Clades I-VIII highlighted on the minimum spanning tree (MST) based on the multi-locus sequence typing (MLST) allelic profile. This was generated using the goeBurst full MST algorithm and was visualised with Phyloviz 2. Each node is a sequence type (ST). The isolates classified in clades I-III, V and VI appear to be correlated also in the MST. Moreover, clades I-III and V-VI, which are more closely related to each other according to the phylogenetic relationship based on the cgSNPs and the DAPC analysis of accessory genome composition, are also related on the MST. However, isolates of clades IV, VII and VIII are dispersed over the entire MST.

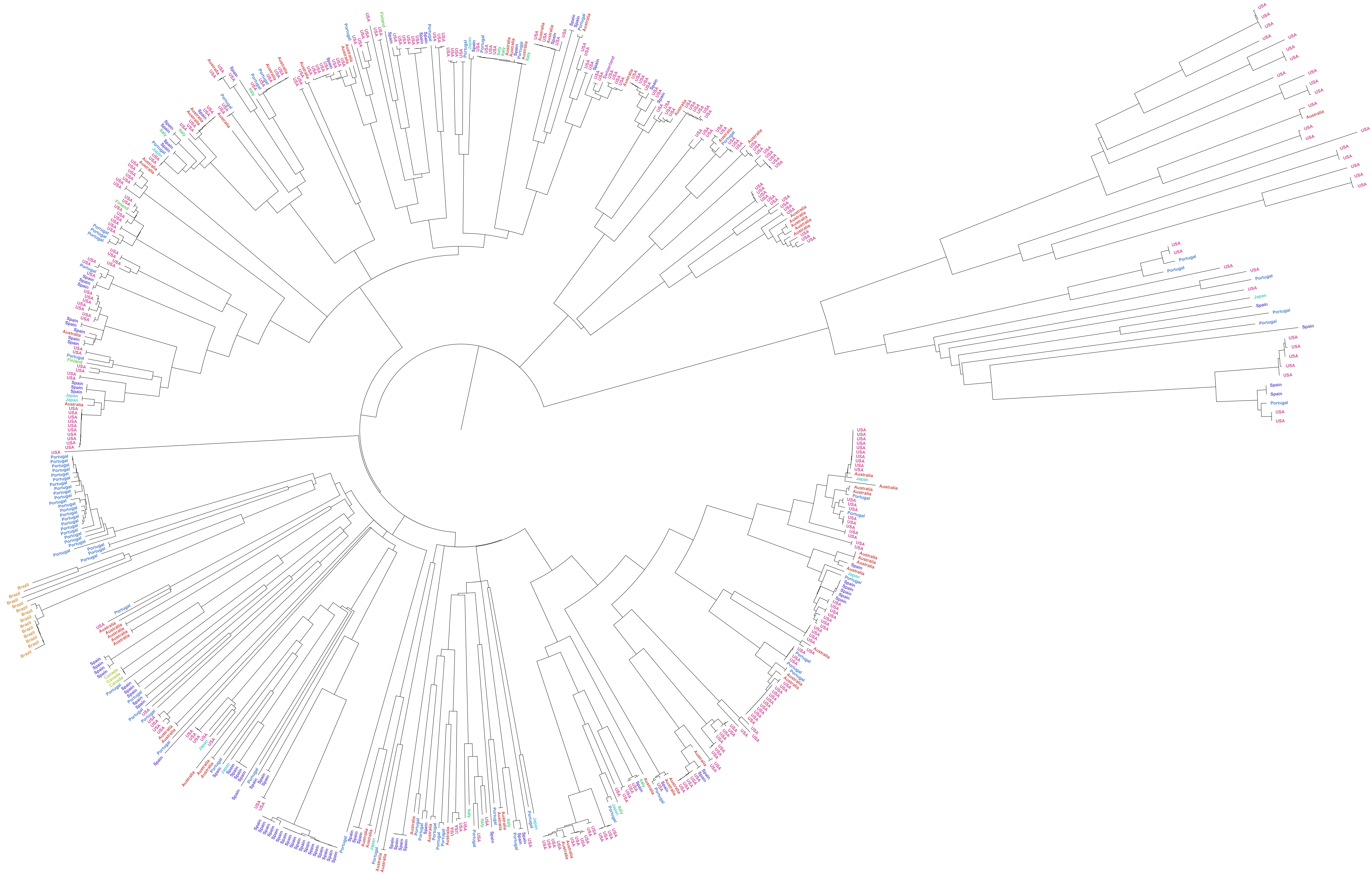


Fig. S2. Distribution of NTHi isolates on the core-genome SNPs-based Neighbour-joining phylogenetic tree, which is annotated with the geographic locus of isolation of the samples. There was no absolute separation of the strains according to geography.