



Figure S1. The evolutionary history was inferred by using the Maximum Likelihood method and Kimura 2-parameter (K2+G+I) model. The tree with the highest log likelihood (-15530.64) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.1338)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 43.59% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 95 nucleotide sequences. All positions with less than 90% site coverage were eliminated, i.e., fewer than 10% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). There were a total of 1209 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [30]. The Polish EHV-5 sequences obtained in the current study are labelled POL_EHV5_stud number_horse ID number. Samples from the same stud are labelled with the rectangle of the same color. Accession numbers for sequences from GenBank are included in the description of each sequence.