

Figure S1 Definition of sequence subsets of different evolutionary divergence in Puumala hantavirus following Duchêne et al. 2015. Sequences in a ML topology connecting to basal branches (black) were removed to create a sequence subset of reduced evolutionary age (1) relative to the full data set. Subsets of further reduced evolutionary levels (2 and 3) were created by removing additional basal sequences where level 3 represents the least diverged subset.

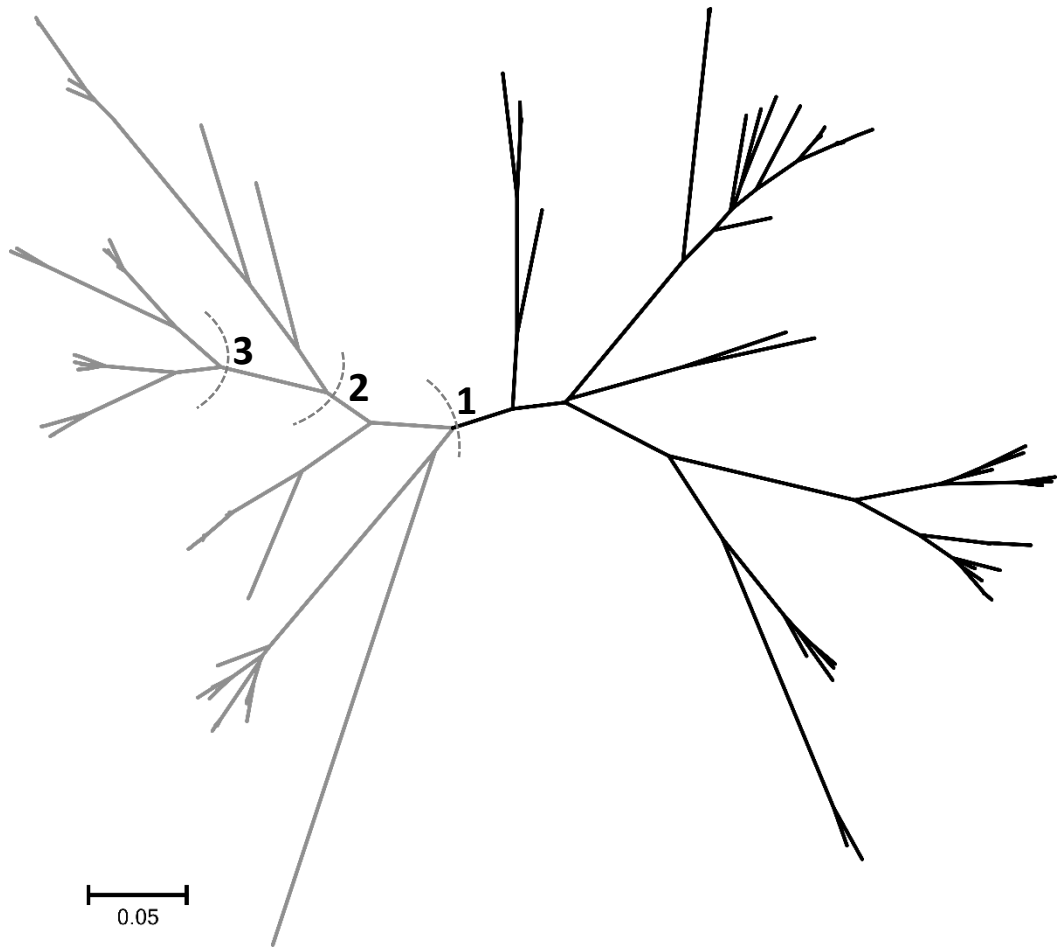


Figure S2 Geographic distance vs. genetic divergence based on whole genome sequences of Puumala hantavirus (PUUV). Genetic distance is calculated as p-distance (A), under the TrN+I+G model of sequence evolution (B) and as sum of branch lengths in a phylogenetic tree (C). Lines indicate fitted saturation growth models with residual standard errors indicated.

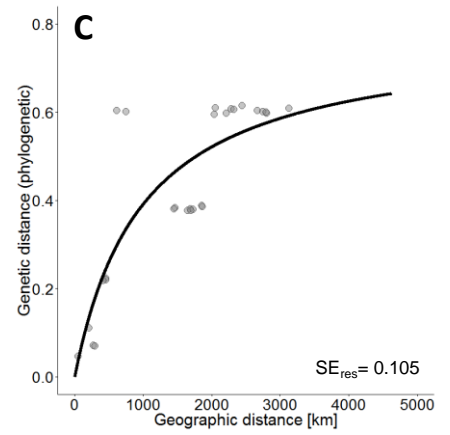
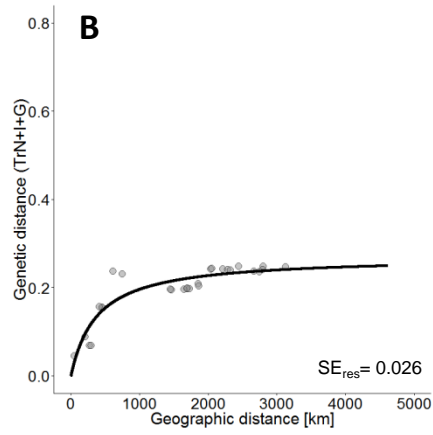
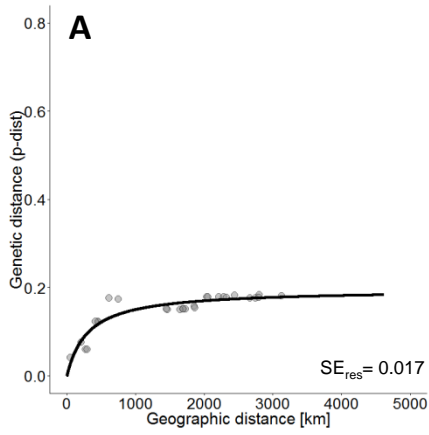


Figure S3 Relationship between geographic distance and genetic divergence based on deduced partial amino acid sequences of the Puumala (PUUV) and Tula (TULV) hantavirus nucleocapsid proteins. p-distances for TULV (A) and PUUV (C) and sum of branch lengths from a phylogenetic tree estimated on the amino acid sequences are shown for TULV (B) and PUUV (D). Lines indicate fitted Saturation Growth Rate models with goodness of fit indicated as residual standard errors.

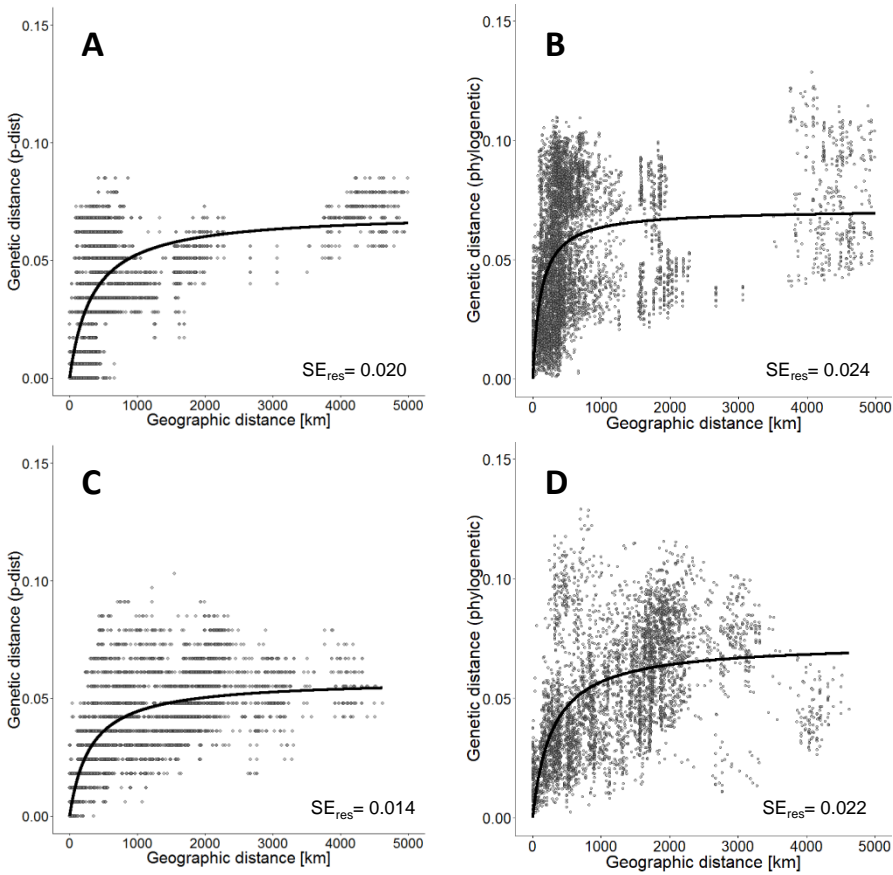


Figure S4 Pairwise geographic distance plotted against genetic divergence estimated for 1st + 2nd (A and C) and 3rd codon position separately (B and D). Divergence was modelled using the TrN+I+G model of nucleotide substitution for PUUV (A and B) and TULV (C and D) partial S-segment sequence.

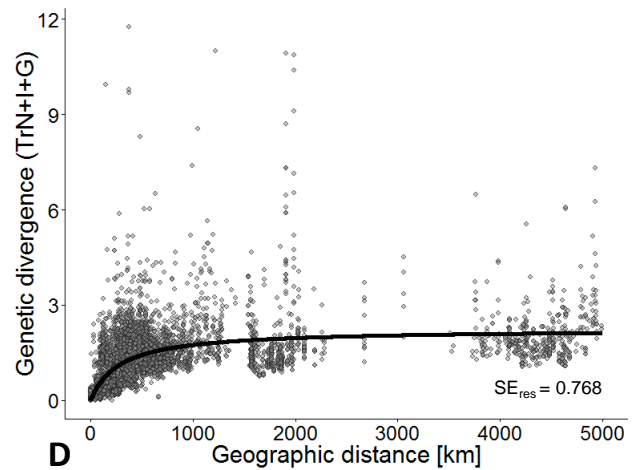
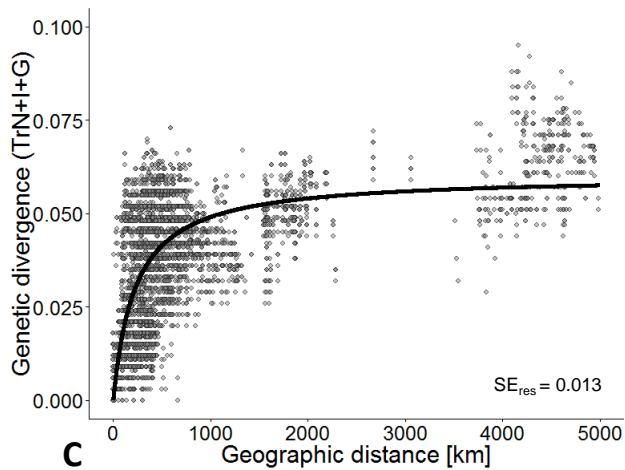
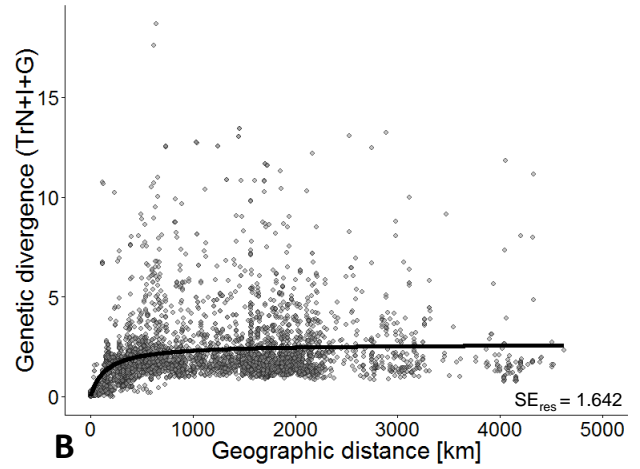
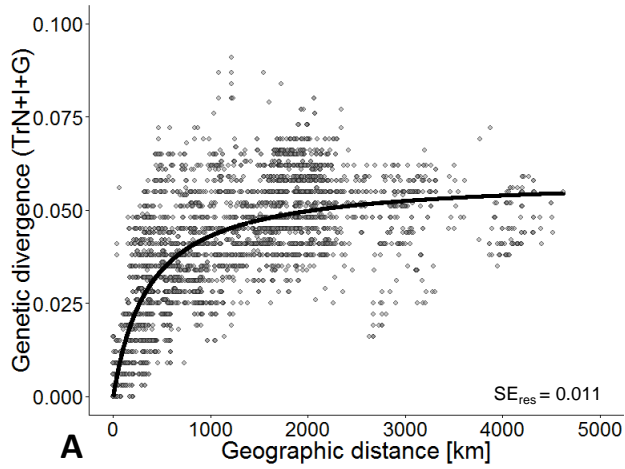
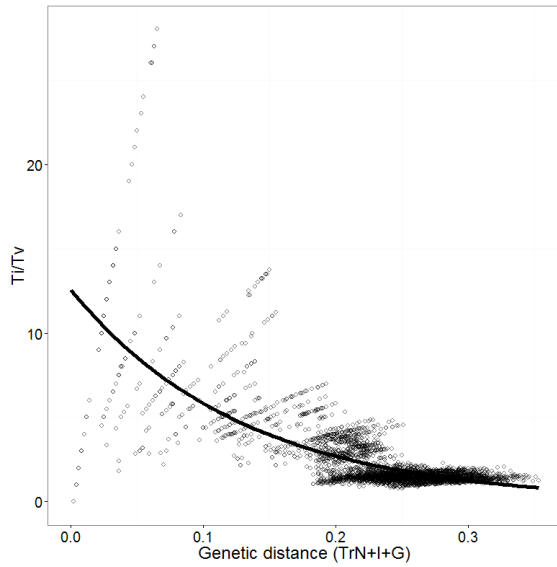
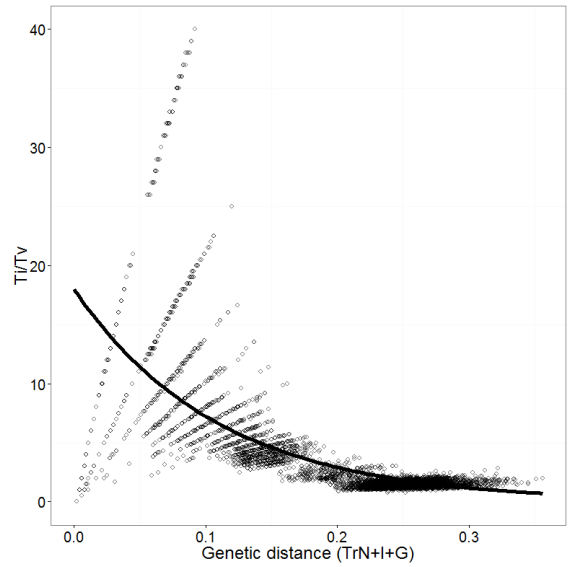


Figure S5 Correlation between pairwise genetic divergence inferred with the TrN+I+G model of sequence evolution and transition to transversion ratios (Ti/Tv) for PUUV (A) and TULV (B) partial S-segment sequences. An exponential model was fitted to data points indicated by a solid line.



A



B

Figure S6 TMRCA estimated using BEAST for PUUV (A) and TULV (B) based on partial S-segment sequences. Node ages are indicated in years for major nodes. Bars indicate 95% highest posterior density (HPD) intervals.

A

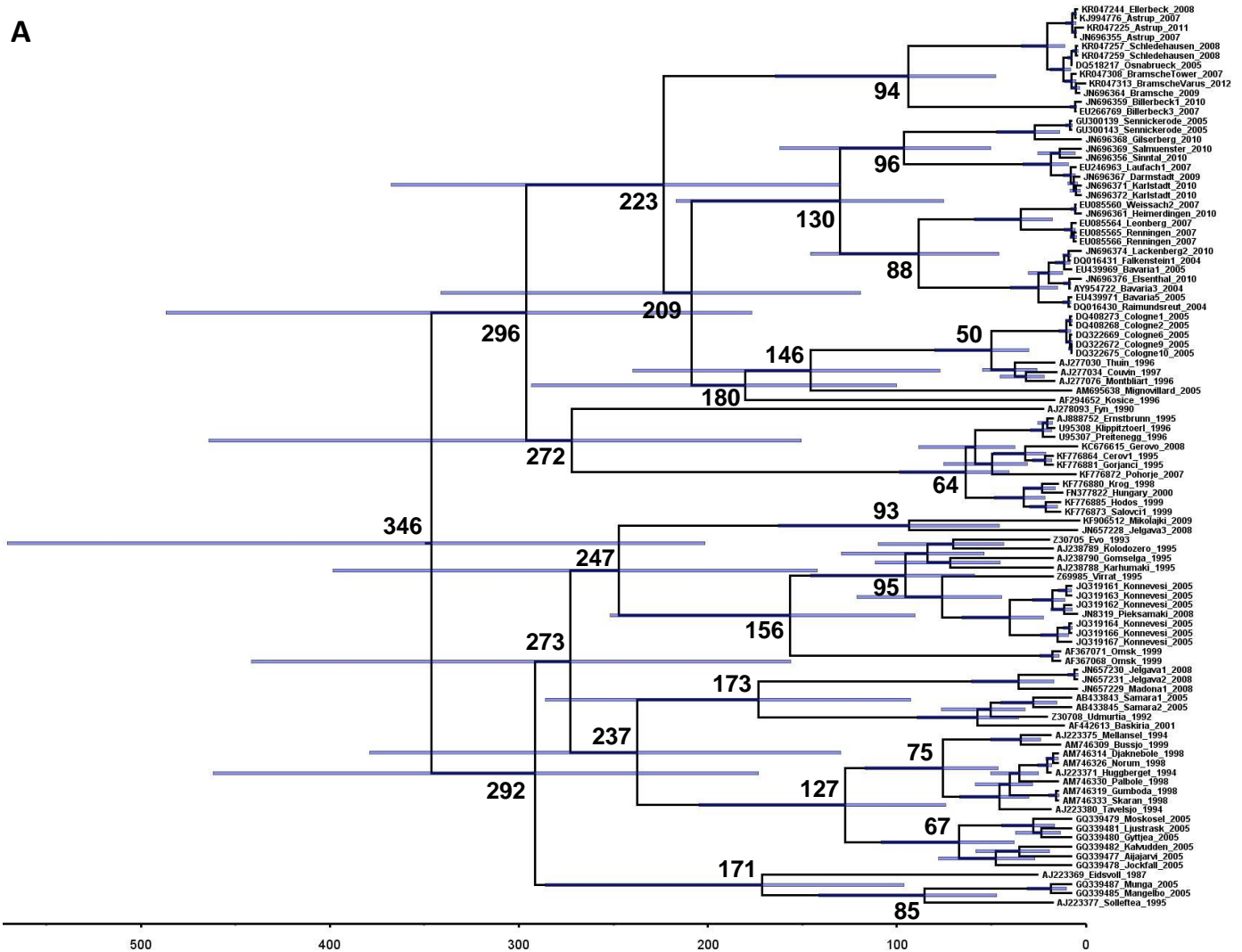


Figure S6

B

