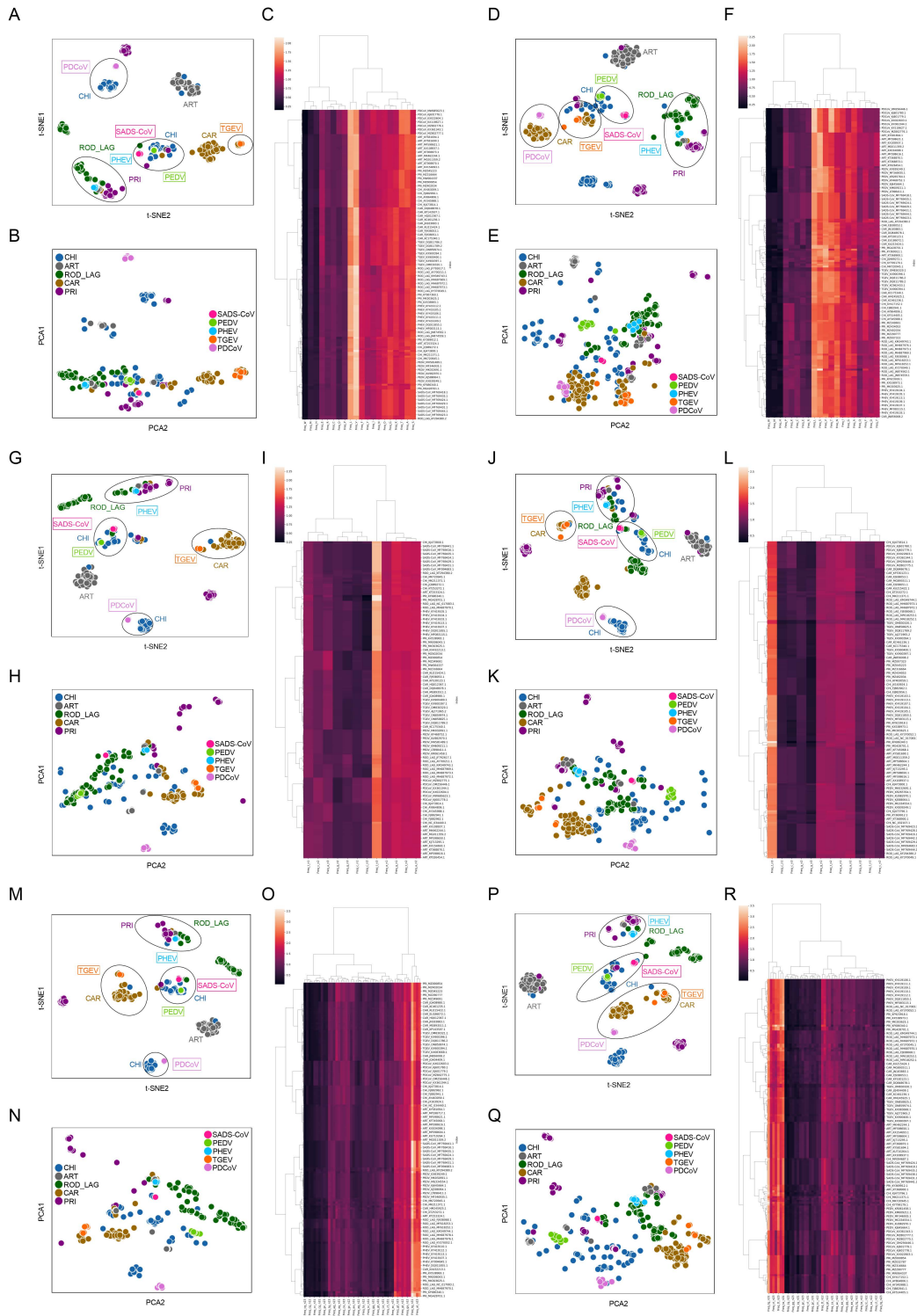

Supplementary data

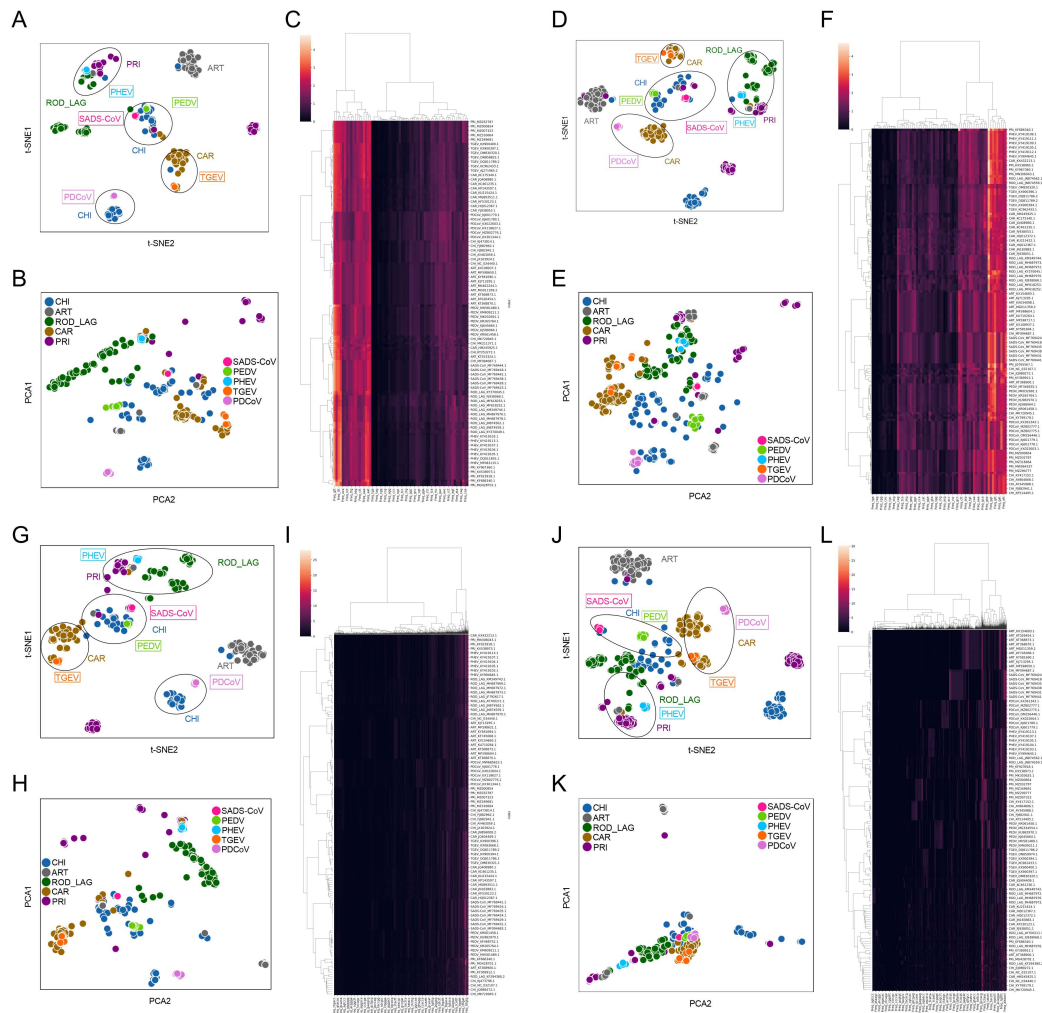
Supplementary Figure S1. Reduction, visualization and clustering of each type of SUI CoVs and other CoVs based on AAs, NTs and DNTs of ORF1ab and Spike sequences.

Visualization of AAs reduced with t-SNE (**A**) and PCA (**B**), and hierarchical clustering of AAs of ORF1ab of each CoV sample (**C**); Visualization of AAs reduced with t-SNE (**D**) and PCA (**E**), and hierarchical clustering of AAs of Spike of each CoV sample (**F**); Similar reduction, visualization and hierarchical clustering based on NTs of ORF1ab (**G–I**) and Spike (**J–L**) sequences; Similar reduction, visualization and hierarchical clustering based on DNTs of ORF1ab (**M–O**) and Spike (**P–R**) sequences.



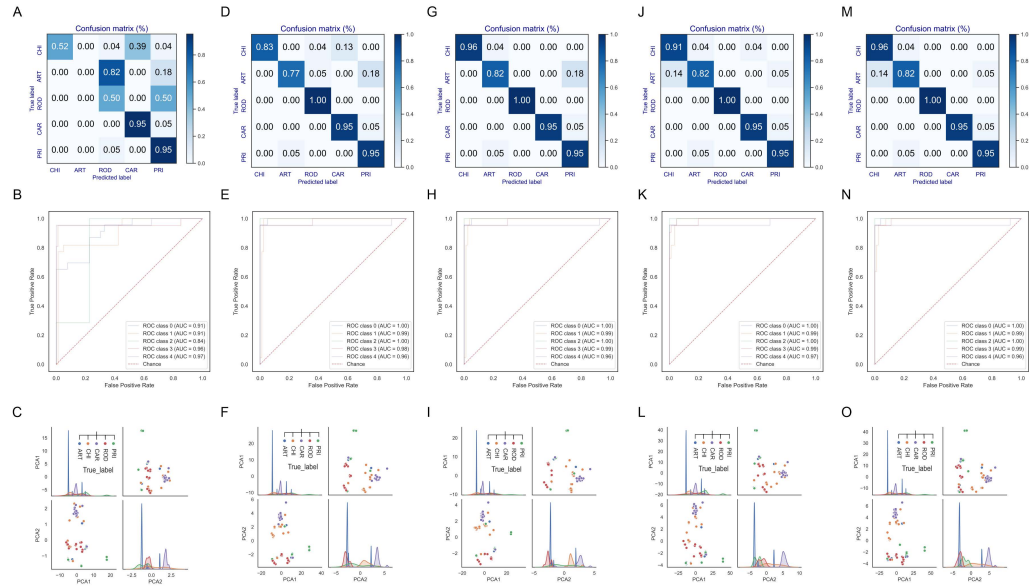
Supplementary Figure S2. Reduction, visualization and clustering of each type of SUI CoVs and other CoVs based on codons and codonpairs of ORF1ab and Spike sequences.

Visualization of codons reduced with t-SNE (A) and PCA (B), and hierarchical clustering of codons of ORF1ab of each CoV sample (C); Visualization of codons reduced with t-SNE (D) and PCA (E), and hierarchical clustering of codons of Spike of each CoV sample (F); Similar reduction, visualization and hierarchical clustering based on codonpairs of ORF1ab (G–I) and Spike (J–L) sequences.



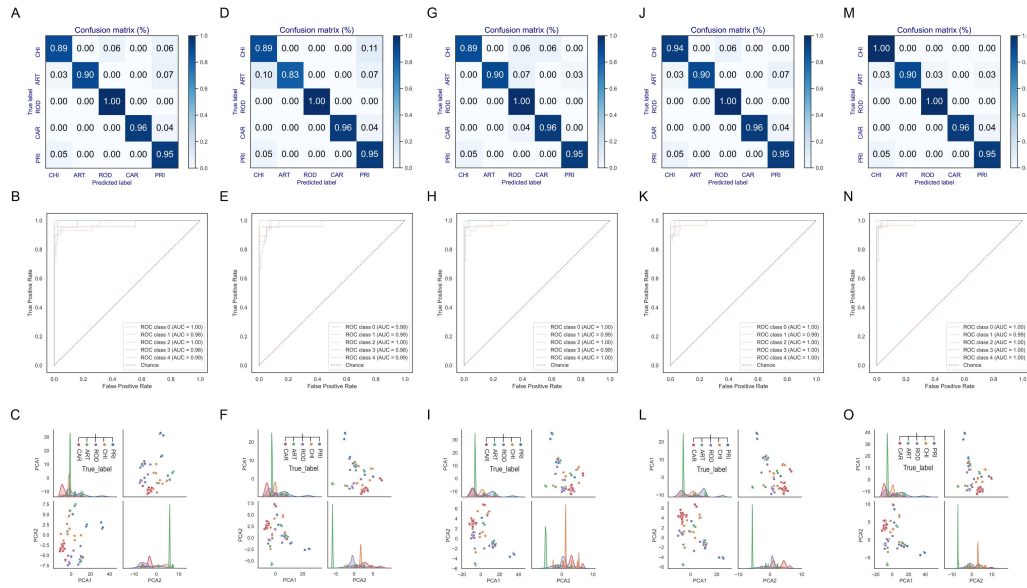
Supplementary Figure S3. Performance of the DCR-based CNN models of ORF1ab gene.

The confusion matrices (A), ROC with AUC (B), and the pair-plotting of PCA1 and PCA2 of the FC data (C) of the ORF1ab model with training epoch of 10, 20 (D–F), 30 (G–I), 40 (J–L), 50 (M–O). ROD means ROD_LAG.



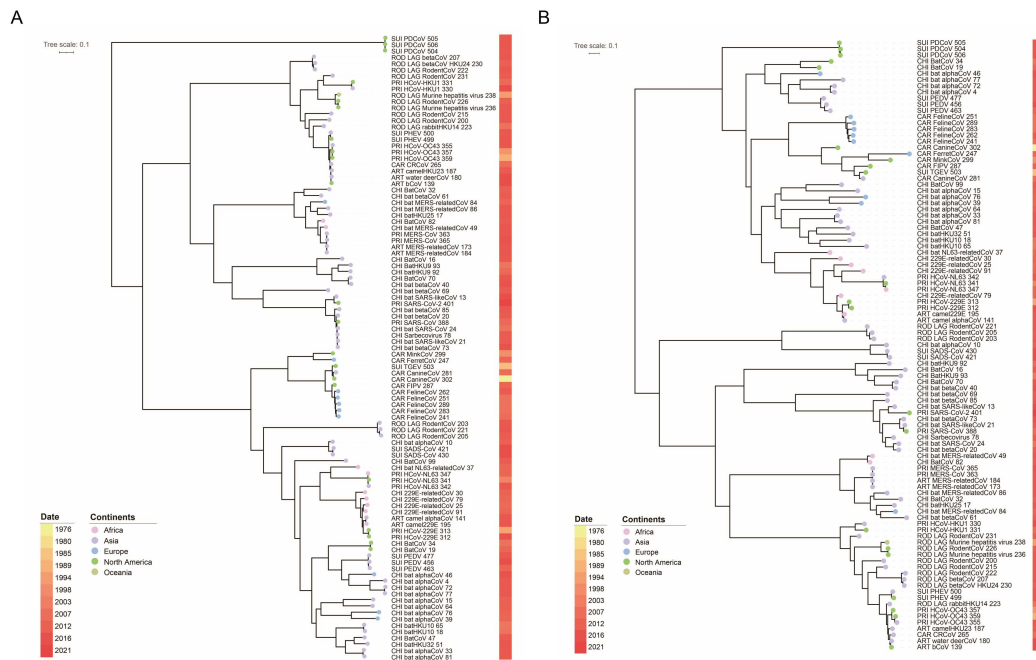
Supplementary Figure S4. Performance of the DCR-based CNN models of Spike gene.

The confusion matrices (A), ROC with AUC (B), and the pair-plotting of PCA1 and PCA2 of the FC data (C) of the Spike model with training epoch of 10, 20 (D–F), 30 (G–I), 40 (J–L), 50 (M–O). ROD means ROD_LAG.



Supplementary Figure S5. Phylogenetic tree of CoVs from different adaptive hosts for ORF1ab and Spike gene.

The phylogenetic tree was constructed using iqtree with 100 bootstrap replicates for randomly sampled CoVs from CHI, ART, ROD_LAG, CAR, PRI and SUI for *ORF1ab* (A) and *Spike* (B) genes. Different colored circles in the tree represent different sample location (pink: Africa, purple: Asia, blue: Europe, green: North America, grey: Oceania). Sampling date is indicated by the progressive color bar at right.



Supplementary Figure S6. Receptor binding verification of PHEV with different hosts.

Biolayer Interferometry (BLI) assay for the interaction between RBD of the PHEV Spike protein and NCAM of rabbit (**A**), rat (**B**) and human (**C**).

