

Crossing Wallace’s Line: A Gibbon Ape Leukemia Virus like retrovirus identified from the Philippine flying lemur (*Cynocephalus Volans*)

Kyriakos Tsangaras^{1,4}, Jens Mayer ², Alex D. Greenwood³x*

Affiliations:

¹ University of Nicosia, Department of Life and Health Sciences, Nicosia, Cyprus

² Institute of Human Genetics, Medical Faculty, University of Saarland, Homburg, Germany

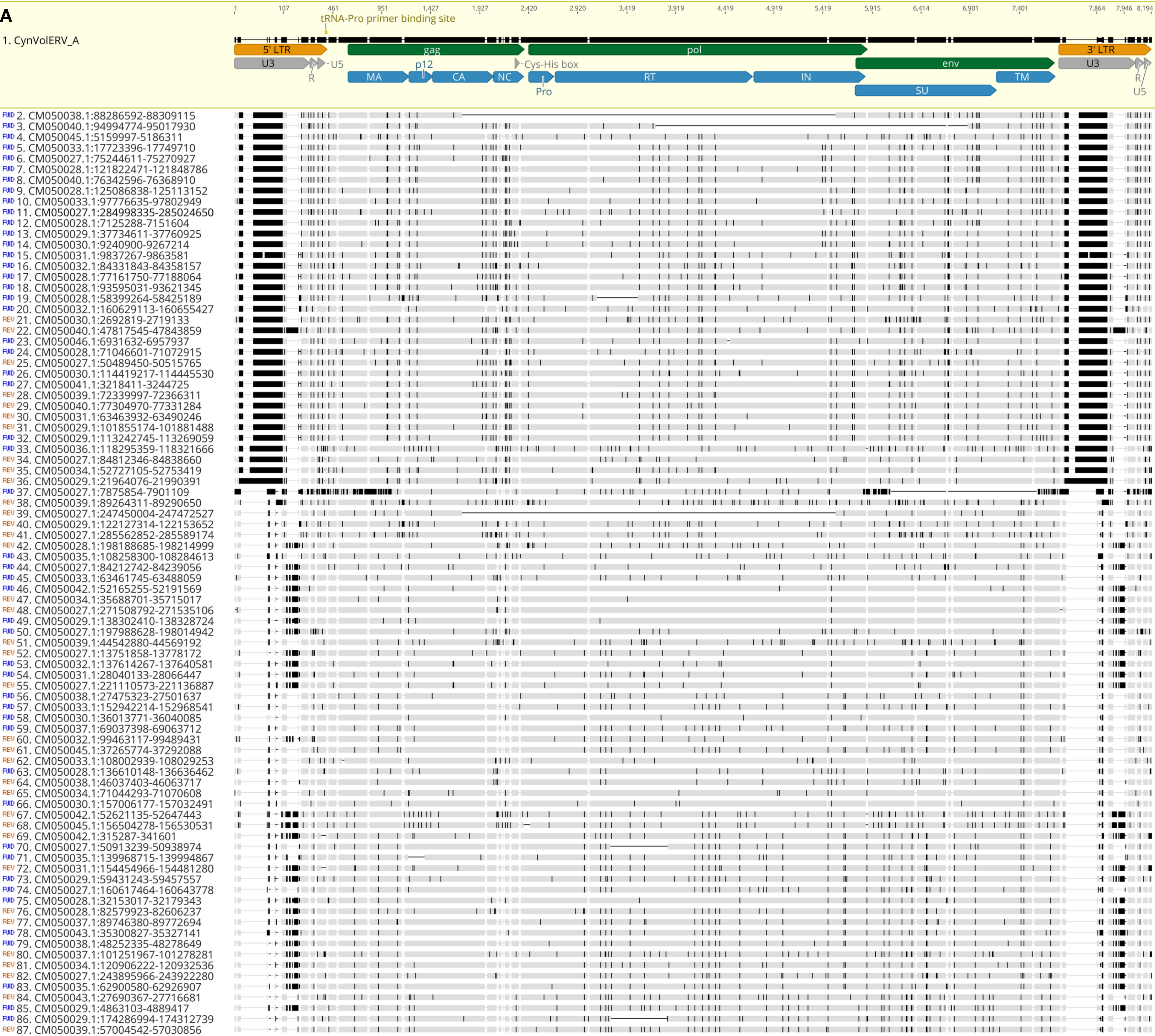
³ Department of Wildlife Diseases, Leibniz Institute for Zoo and Wildlife Research (IZW), Berlin, Germany

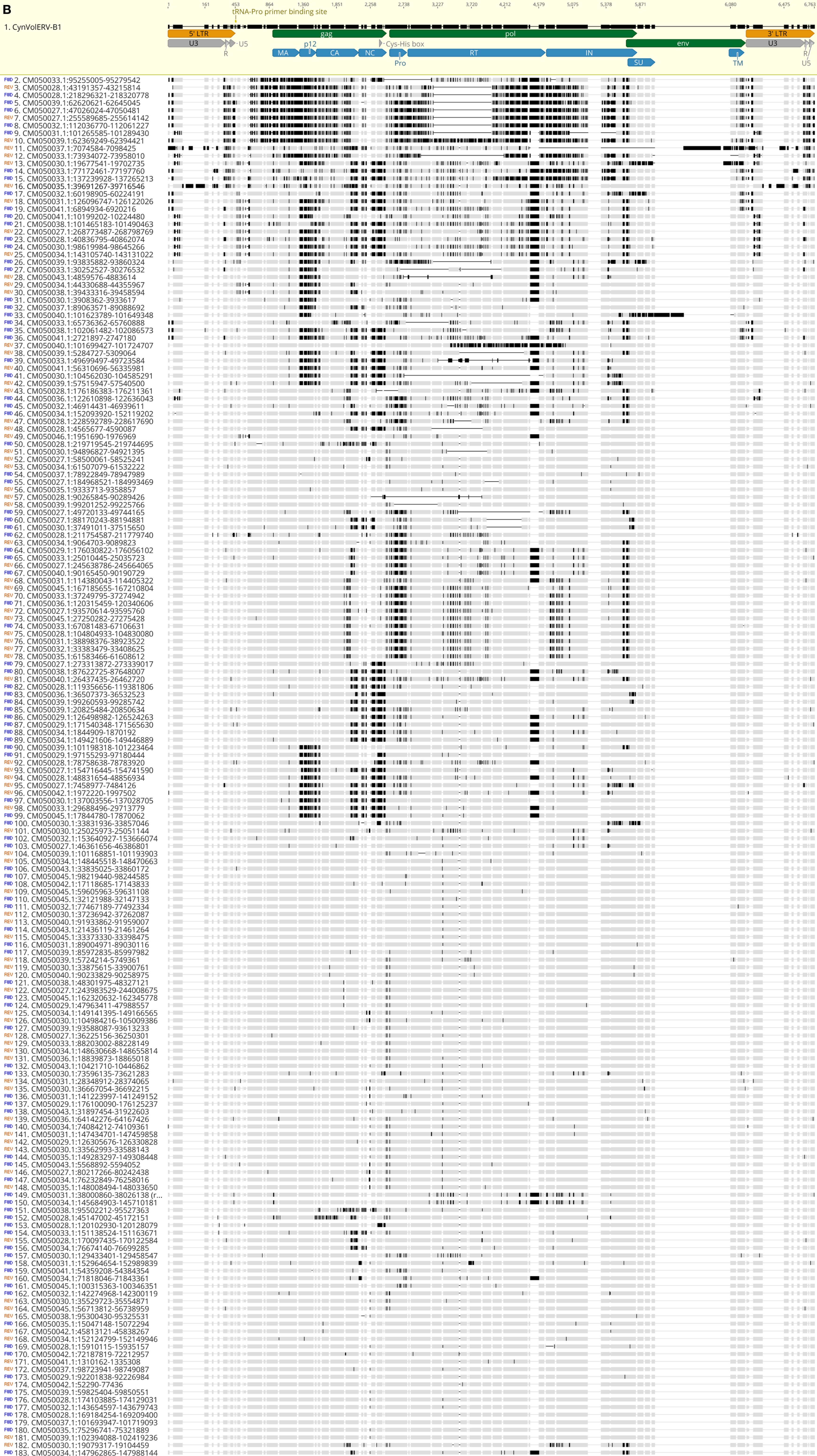
⁴ Department of Cardiovascular Genetics and the Laboratory of Forensic Genetics , Cyprus Institute of Neurology and

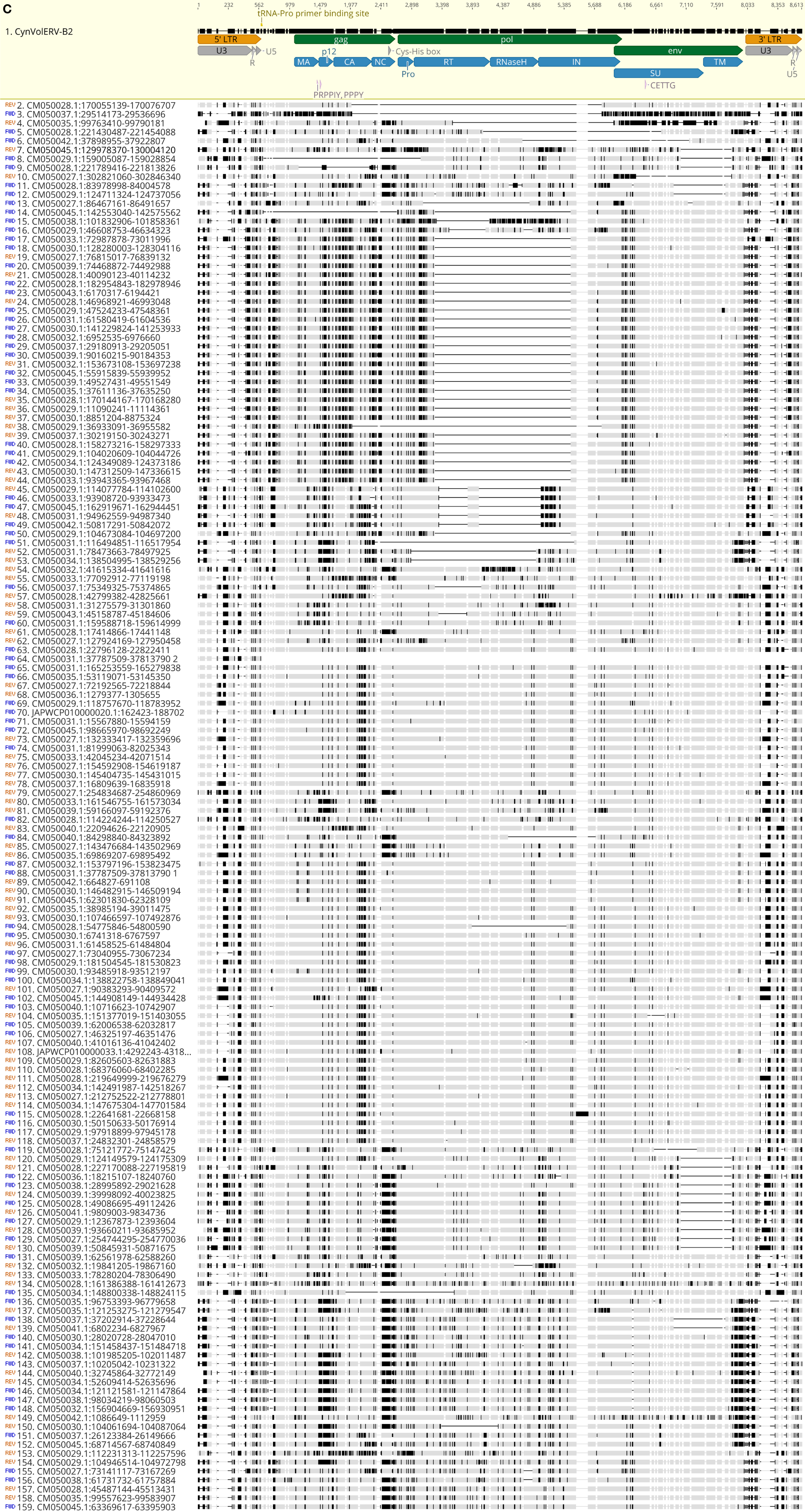
Genetics, Nicosia, Cyprus

x School of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany

* Corresponding author: greenwood@izw-berlin.de







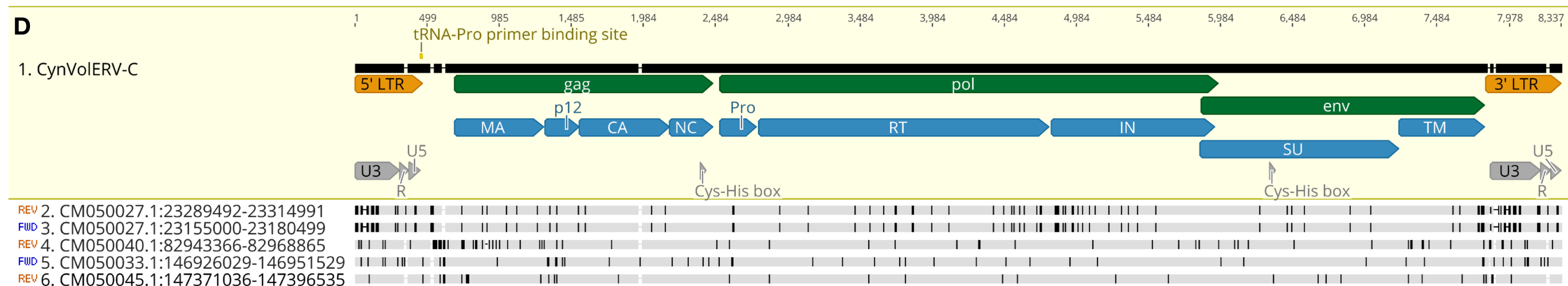


Fig S1: Multiple sequence alignments of CynVolERV-A, CynVolERV -B1, CynVolERV-B2, and CynVolERV-C proviral loci. CynVolERV sequences were identified using megablast in the *Cynocephalus volans* reference genome (GCA_027409185.1). Proviral loci were extracted from the genome and multiply aligned using MAFFT v7. A) Proviral alignment of 86 sequences using the CynVolERV-A consensus sequence as reference. The first 36 sequences belong to the CynVolERV-A LTR2 group, while the remaining sequences belong to the CynVolERV-A LTR1 group. B) CynVolERV-B1 reference sequence aligned to 182 CynVolERV-B1 proviral sequences. C) 152 CynVolERV-B2 proviral sequences aligned to CynVolERV-B2 reference sequence. D) Proviral alignment of 5 sequences against the CynVolERV-C consensus sequence. Black vertical lines indicate sequence differences compared to the reference sequence, while gray color indicates sequence identities. Note here that due to resolution, proviral regions appear more dissimilar in sequence than they actually are.

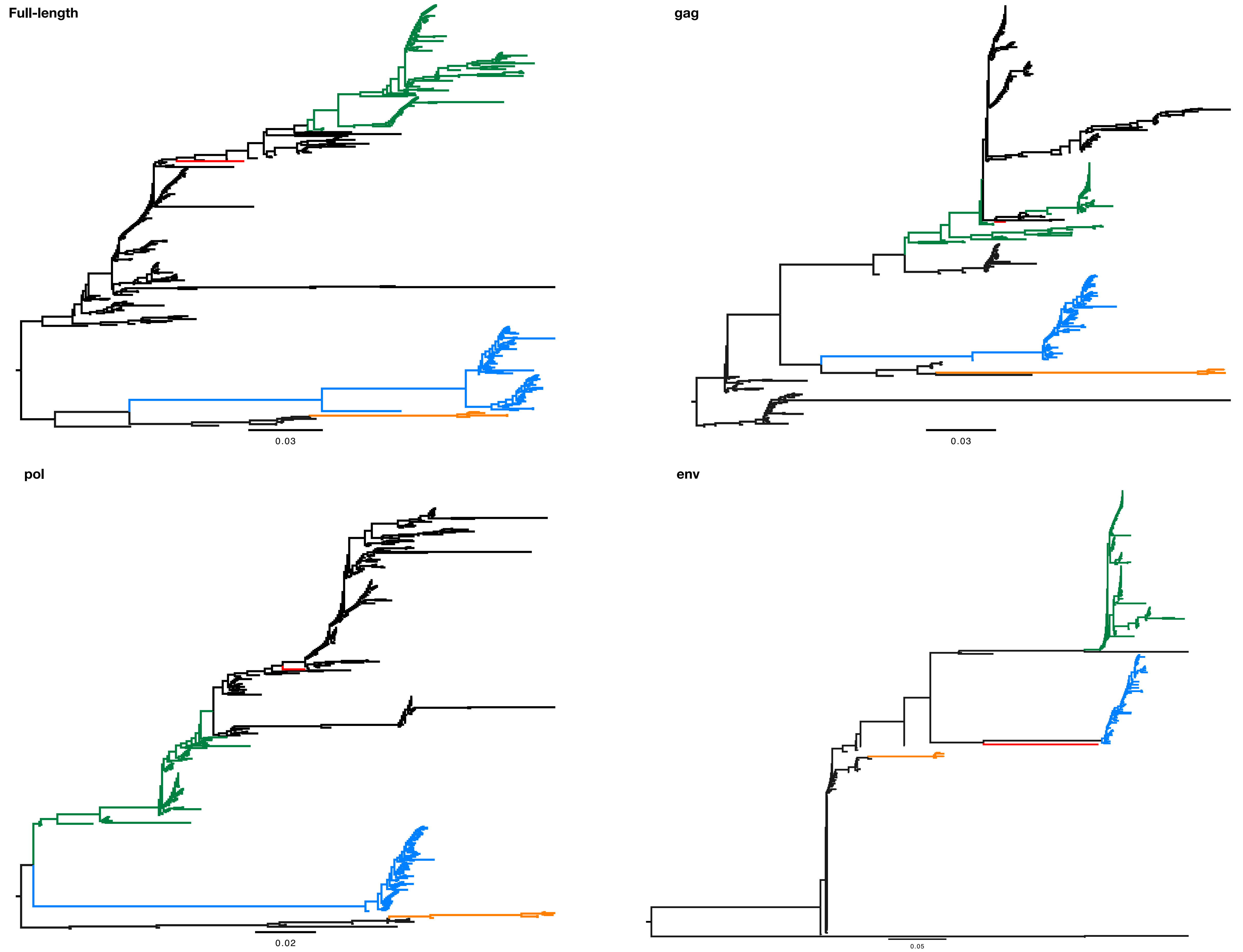


Fig S2: Phylogenetic trees for CynVoIERV full-length proviral sequences and CynVoIERV gag, pol and env gene regions. Trees were generated using FastTree. CynVoIERV-A sequences are highlighted in blue, CynVoIERV-B1 in black, CynVoIERV-B2 in green, CynVoIERV-C in orange. The CynVoIERV-consensus sequence of all sequences included in the tree is highlighted in red each.

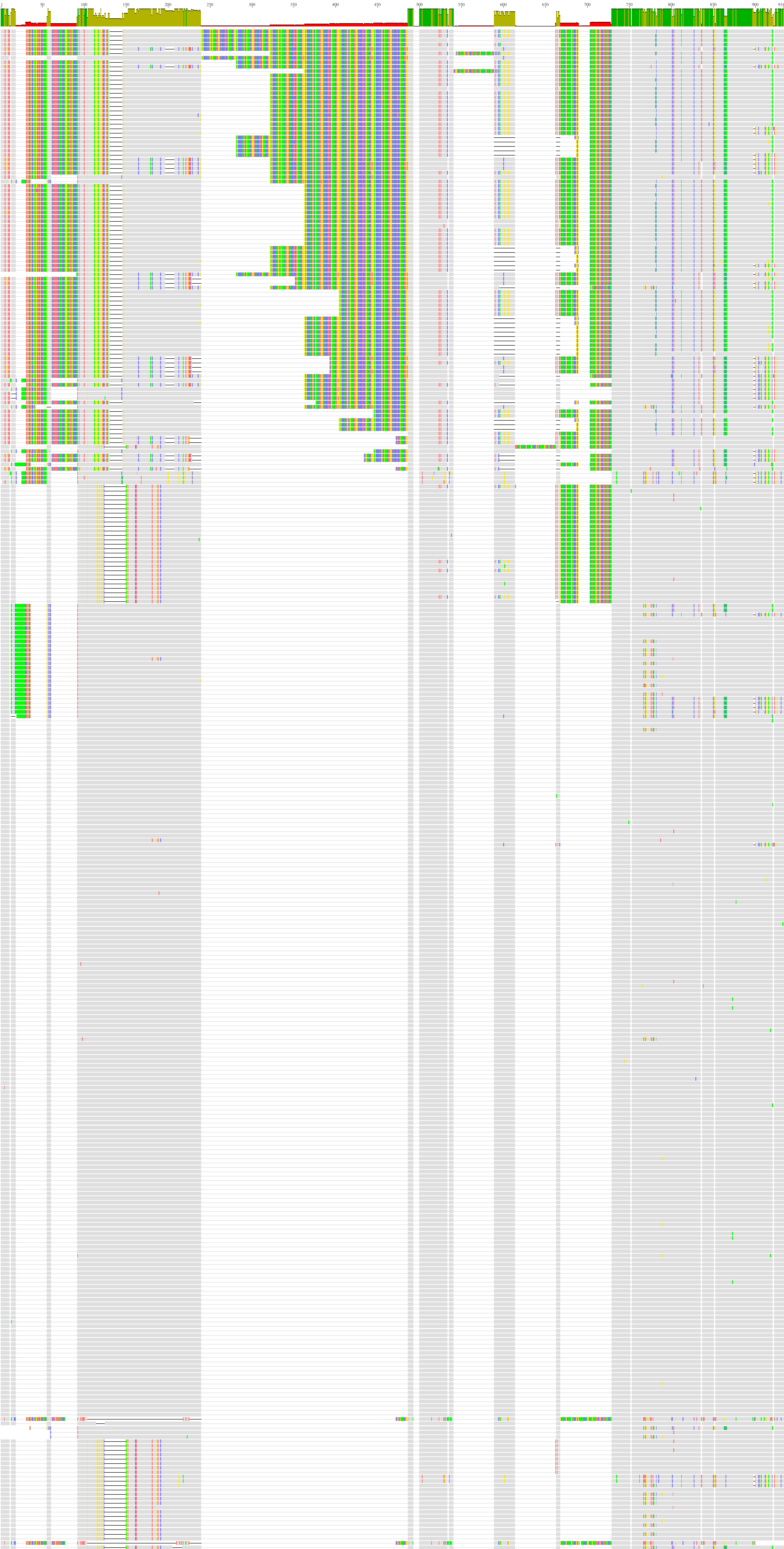


Fig S3: Multiple alignment of CynVolERV-B 5'LTRs demonstrating considerable indel variation between proviral LTRs. LTR sequences in the multiple alignment are ordered by sequence length for better depicting indel variations. Nucleotide positions identical to the consensus are in grey

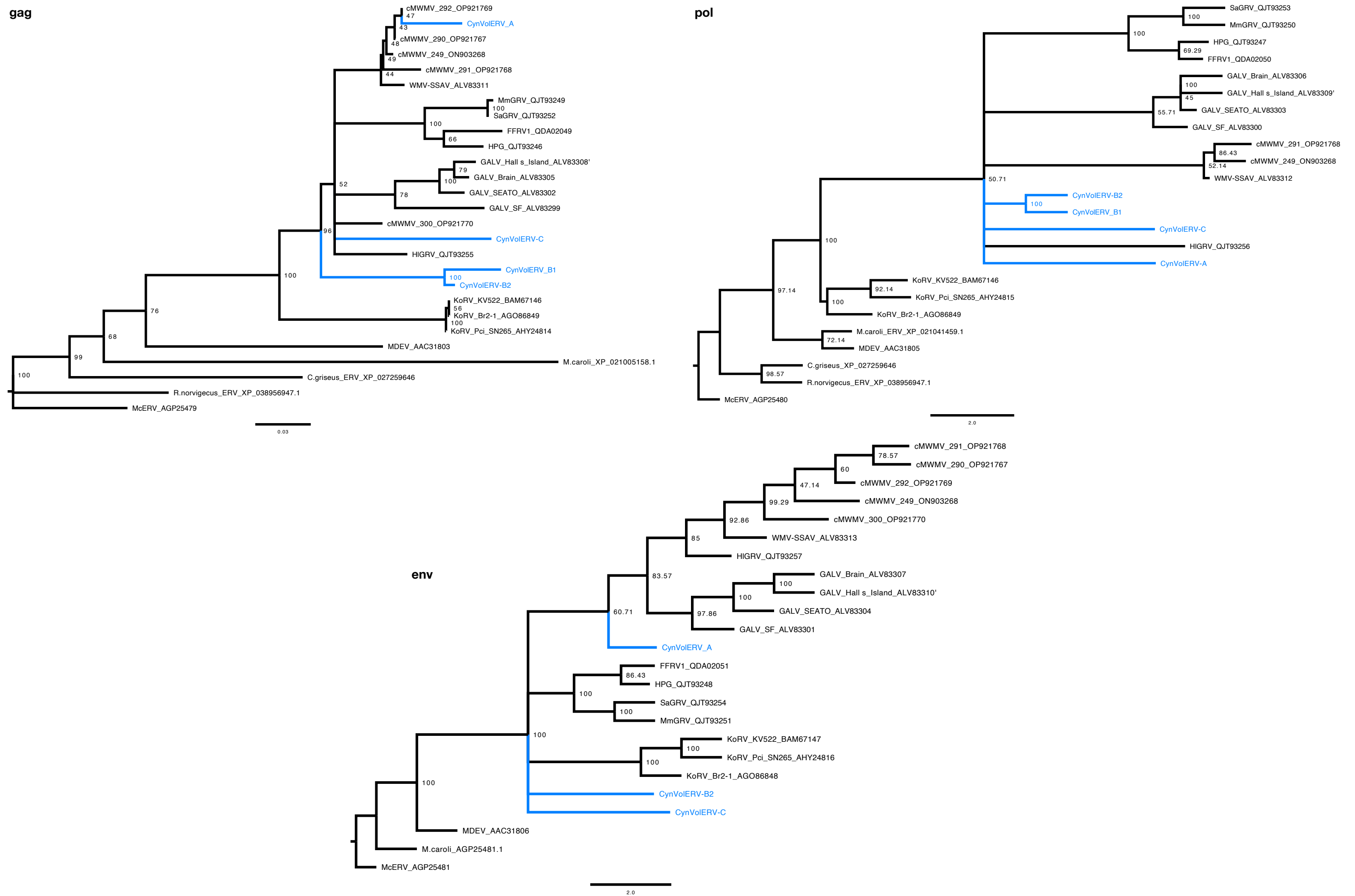


Fig S4: Phylogenetic trees generated with RAxML for proviral proteins, as predicted from each CynVolERV subgroup's consensus sequence by Retrorector. McERV sequence was used as an outgroup.

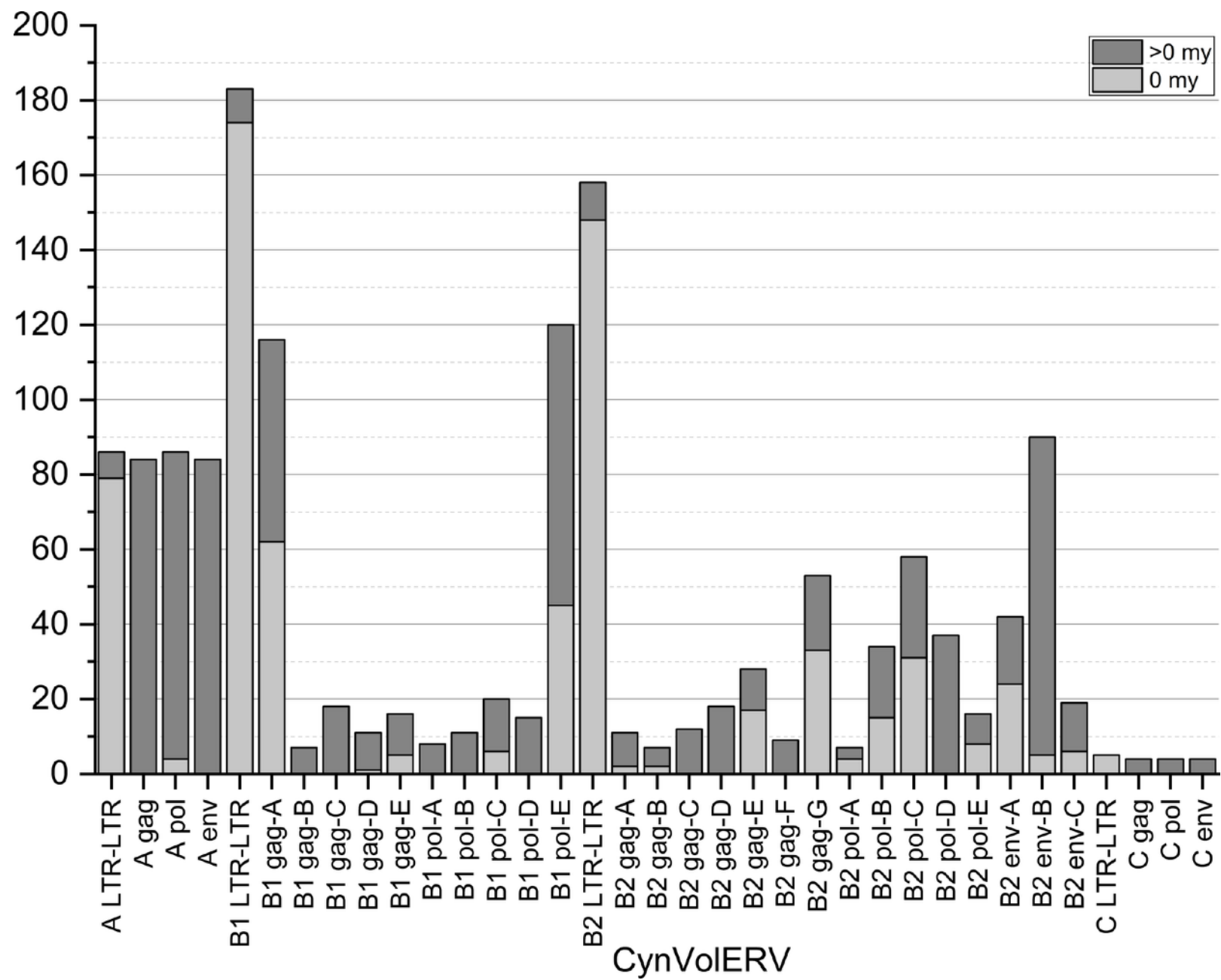


Fig S5: Stacked bar chart that illustrates zero vs. non-zero million years provirus age estimations