



**Fig S4.** Superimpositions between the modern KoRV (Pci-SN265-green) protein structure and its historical variants show the overall similarity of the structures of the Pol protein. Amino acid variations (red) between these sequences are mapped on the protein models (arrows). The structural differences predicted are attributed to changes in the polarity, charge, and atom conformations. These differences are largely localized onto flexible loop regions. The models are shown in cartoon representations and the atoms of the variable amino acid residues in line representations to view the side chains. In all comparisons the Pci-SN265 consensus sequence was used as the reference. The domain organization (as depicted at the CDD database) and the location of the modeled structure of the protein are shown at the top of the figure. RP: Retropepsin of the RTVL\_H family of human endogenous retrovirus-like elements; RT\_ZFREV\_like: Reverse transcriptase subfamily found in sequences similar to the intact endogenous retrovirus from zebrafish and Moloney murine leukemia virus; Rnase\_HI\_RT: Bel/Pao family of RNase HI in long-term repeat retroelements ; rve: Integrase core domain.