



**Additional file 4:** Figure S2. Tree showing the inferred relationships among the 2556 ARV genes sequences from the 10 main groups classified by the phylogenetic analysis of the 105 ARV genes sequences from the 10 main groups and the 105 ARV genes sequences from each of the next species in the genus *Baculovirus*. Phylogeny Groups of related taxa are shown and numbers according to the species of the type strain sequence included in the phylogeny. Phylogeny Group names are shown in italics. These ANVs with relative abundance greater than 1% were used to infer the phylogeny of the 2556 ARV genes sequences. The tree was inferred using the Minimum Evolution method [11]. The optimal tree with the sum of branch length = 46.940145 is shown. The percentage of replicates in which the associated taxon clustered together in the bootstrap test (100 replicates, only branches present > 70% support) is shown at the nodes. Bootstrap values less than 70% are not shown. The evolutionary distances were computed using the maximum likelihood method [11]. The tree is rooted at the midpoint of the root node. The evolutionary distances were computed using the same units as those of the evolutionary distances used to infer the phylogeny.

- and are in the units of the number of base substitutions per site. The MT tree was searched using the Close-Neighbour-interchange (CNI) algorithm [4] at a search level of 1. The Neighbour-joining algorithm [5] was used to generate the initial tree. The analysis involved 2626 nucleotide sequences (including 1000 reference sequences) and 100000 randomised trees. There were a total of 38 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [5].

  - Rahman A, and Nee S. (1992). A simple method for estimating and testing maximum entropy models. *Molecular Biology and Evolution* 9:94-967.
  - Felsenstein J. (1993). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 37:783-791.
  - Tanabe S, Kuroki N, and Kumar S. (2004). Prospects for inferring very large phylogenetic trees. In: *Advances in bioinformatics and computational methods*. Proceedings of the National Academy of Sciences USA 101:10352-10357.