Supplement 1

Immunohistochemistry (IHC) protocol

IHC was employed in 3 µm sections of organs with tumor and the immunophenotype was determined using polyclonal rabbit anti-CD3 antibodies (code IS503; ready-to-use; Dako) for T-lymphocytes, and monoclonal mouse anti-CD79α (code M067A; Biocare Medical) and polyclonal rabbit anti-CD20 antibodies (code PA516701; Thermo Fisher Scientific) for B-lymphocytes. For the B-cell markers, initially, endogenous peroxidase was blocked by submerging the sections in 10% hydrogen peroxide for 15 min. Antigen retrieval for CD79α and CD20 was performed by boiling the sections in Tris-EDTA pH 9.0 in a digital pressure cooker (Dakocitomaker) for 20 min at 96°C for CD79α and in two 5-minute cycles in the microwave for CD20. To block nonspecific binding, the sections were submerged in 5% dried milk solution for 30 min. The sections were then incubated with the primary antibodies, with dilutions of 1:10 (CD79α) and 1:250 (CD20). Amplification signal was obtained with MACH4 universal HRP polymer (Biocare Medical) and the reactions were revealed with 3-amino-9-ethylcarbazole chromogen kit (AEC, Biocare Medical) and counterstained with hematoxylin. For CD3, IHC was performed using Autostainer link 48 (Agilent-Dako®). The equipment steps are: antigen retrieval with PT Link equipment (Agilent-Dako®) with low pH buffer (EnVision FLEX Target Retrieval Solution, 50x, concentrate) for 20 min at 97°C and high pH buffer (Target Retrieval Solution, 50x, concentrated) for 30 min at 97°C. The dilution of the solutions was performed as recommended (1:50); employment of EnVision FLEX Wash Buffer (1:20), concentrate buffer in the sections for 5 min; blockage of endogenous enzymes with FLEX peroxidase block reagent for 5 min; incubation with primary antibody for 30 min; detection of the reactions with EnVision FLEX/HRP polymer (Dako®); staining with FLEX DAB + Sub-chromogenic system (Dako®) with two cycles of 5 min; and counterstaining with hematoxylin for 5 min.

Supplement 2

Cladogram methodology

A set of 11 sequences including the sequence obtained from the sample were aligned using Clustal Omega and the tree was inferred and edited using W-IQ-TREE web server and iTOL web server. The method used was the Maximum Likelihood based on the Kimura-2 model with gamma distributed sites and the internal branch test was made with 1000 bootstrap replicates.

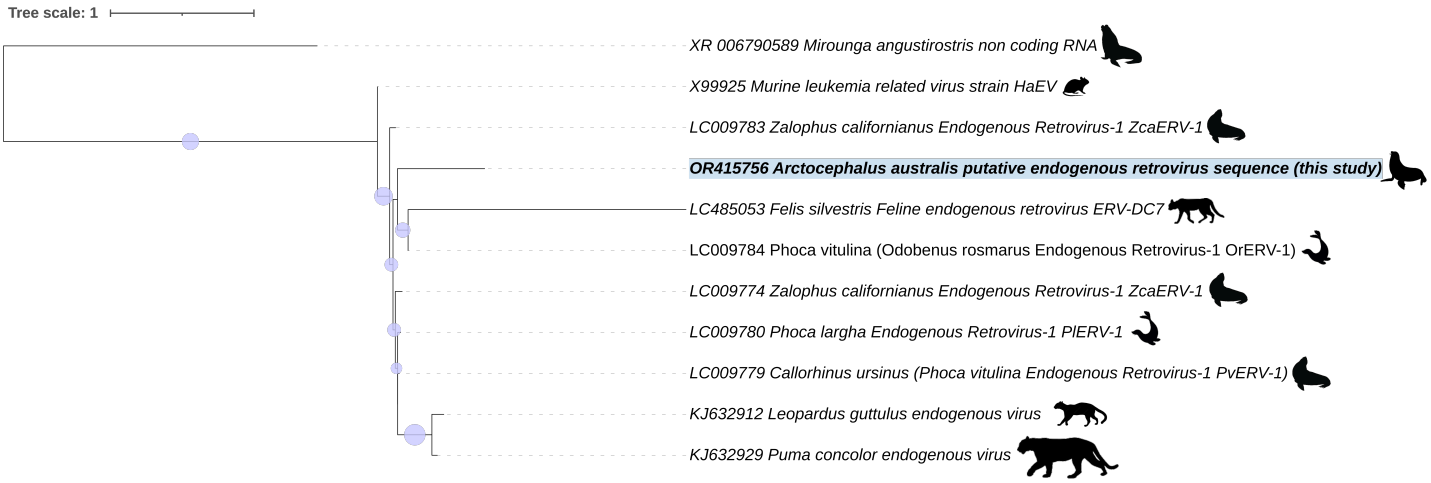

Figure S1: Cladogram exhibiting the phylogenetic relationships with sequences from exogenous and endogenous retrovirus detected in other species. The silhouette belongs to the host species from which the sequence was obtained. In parenthesis is the original host of the sequence detected. The dots on the branches represent posterior probabilities above 60%