

s22

s22

**From:** Heidi Drumme  
**Sent:** Tuesday, 3 September 2019 4:11 PM  
**To:** Joshua Hayward  
**Cc:** mary.tachedjian; Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; philip-daniel.solymos; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos; garycramer; Andy Poumbourios; Glenn.Marsh; Michelle Baker; Edward Holmes; Wang Linfa; <Ina.Smith>; Gilda Tachedjian  
**Subject:** Re: HPG Bat Retrovirus Manuscript  
**Attachments:** HPG paper Manuscript Final v2 HD.docx; HPG paper - SI Appendix Final v2 HD.docx

Attachments Removed

Hi Joshua and all authors,  
Great job from everyone to pull all this work together. Here are a few more comments on the latest version.  
Cheers and good luck  
Heidi

Duplicate Email - Removed

s22

**From:** Joshua Hayward [s22]  
**Sent:** Tuesday, 3 September 2019 7:11 PM  
**To:** Heidi Drummer  
**Cc:** mary.tachedjian; Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; philip-daniel.solymosi [s22] Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos [s22] garycramer [s22] Andy Pombourios; Glenn.Marsh; Michelle Baker; Edward Holmes; Wang Linfa; <Ina.Smith [s22]> Gilda Tachedjian  
**Subject:** Re: HPG Bat Retrovirus Manuscript

Hi Heidi,  
 Thank you for that!  
 Cheers,  
 Josh

**Joshua Hayward PhD**  
**Research Officer**  
**Retroviral Biology and Antivirals Research Laboratory**

**Burnet Institute**  
 Disease Elimination and Maternal & Child Health Programs

s22

GPO Box 2284 Melbourne Victoria Australia 3001

s22

The Macfarlane Burnet Institute for Medical Research and Public Health Ltd ABN 49 007 349 984

Duplicate Email - Removed

s22

**From:** Joshua Hayward [s22]  
**Sent:** Thursday, 5 September 2019 12:46 PM  
**To:** mary.tachedjian; Kohl, Claudia; Adam Johnson; Deamley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; Solymosi, Philip; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos [s22]; garycrameri [s22]; [s22]; Heidi Drummer; Andy Poubourios; Glenn.Marsh; Michelle Baker; Edward Holmes; Wang Linfa; <Ina.Smith [s22]>; Gilda Tachedjian  
**Subject:** HPG Paper - Fwd: PNAS MS#2019-15400 Submitted  
**Attachments:** HPG paper - Manuscript and SI v1.pdf

Dear all,

I am happy to report that the HPG manuscript has been submitted! The submitted single-document manuscript/SI is attached.

Best regards,

Josh

**Joshua Hayward PhD**

**Research Officer  
 Retroviral Biology and Antivirals Research Laboratory**

**Burnet Institute  
 Disease Elimination and Maternal & Child Health Programs**

s22

GPO Box 2284, Melbourne, Victoria, Australia 3001

s22

The Macfarlane Burnet Institute for Medical Research and Public Health Ltd ABN 49 007 349 984

----- Forwarded message -----

**From:** Gilda Tachedjian [s22]  
**Date:** Thu, 5 Sep 2019 at 12:30  
**Subject:** Fwd: PNAS MS#2019-15400 Submitted  
**To:** Joshua Hayward [s22]

Can you please forward this email to all authors along with the submitted manuscript attached for their records

Begin forwarded message:

**From:** [journalstaff@pnascentral.org](mailto:journalstaff@pnascentral.org)  
**Subject:** PNAS MS#2019-15400 Submitted  
**Date:** 5 September 2019 at 12:25:45 pm AEST  
**To:** gildat [s22]  
**Cc:** gtachedjian [s22]  
**Reply-To:** [pnas@nas.edu](mailto:pnas@nas.edu)

September 4, 2019

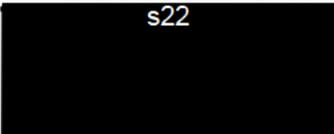
Dear Dr. Tachedjian,

Thank you for your submission to PNAS. Your Direct Submission was received and will be processed within 24–48 hours. You will receive an email once the staff have confirmed everything is in order or letting you know if any changes are necessary.

Thank you for submitting to PNAS.

Sincerely yours,

s22

A large black rectangular redaction box covers the signature area of the letter.

## SI Appendix

## SI Appendix

### Contents

Supplementary Methods.....	23
Ethics .....	23
Sample collection along east coast of Australia and Daintree Rainforest .....	23
Metagenomic analysis of viral nucleic acid in bat samples .....	23
Isolation and assembly of the HPG genome sequence .....	24
Identification and assembly of KoRV-related viruses in publicly available databases .....	25
Annotation of Retroviral Genomes .....	26
HPG-specific analysis of Pteropid genomes .....	26
Phylogenetic analysis.....	27
Generation of HPG proviral sequence and synthesis of HPG proviral expression construct	27
Cell cultures .....	28
Transfection of 293T cells for generation of HPG & M-MLV viral particles .....	28
Electron Microscopy and viral particle morphology .....	28
Reverse transcriptase divalent cation preference .....	29
Replication kinetics assay .....	29
Generation of gammaretroviral <i>env</i> gene expression constructs for pseudotyping .....	30
Generation of pseudotyped retroviruses and assay of host cell tropism .....	31
RT-qPCR for detecting the presence of gammaretroviral nucleic acids in bat samples .....	31
Generation of HPG Envelope protein for serological assays.....	32
Generation of anti-HPG Envelope sera .....	32
Serological assay for the presence of anti-HPG antibodies in bats.....	33
Supplementary Figures.....	34
Supplementary Tables .....	44
Supplementary References .....	49

## SI Appendix

### 1 Supplementary Methods

2

#### 3 Ethics

4 Fieldwork in all locations and on all occasions was approved under the following permits:  
5 in Queensland, the (then) Department of Employment, Economic Development and  
6 Innovation Animal Ethics Committee (AEC) Permit SA 2011/12/375 and 1710 the  
7 Environmental Protection Agency/Department of Environment and Resource  
8 Management Scientific Purposes Permits WISP14939514, WISP05810609, and  
9 WISP14100614; in New South Wales, The University of Sydney AEC Permit 04/3  
10 2011/1/5498, the Elizabeth Macarthur Agricultural Institute AEC Permit M11/15, the  
11 Office of Environment and Heritage AEC Permit 120206/02, and the Office of Environment  
12 and Heritage Scientific Licenses SL100086 and SL 100537.

13

#### 14 Sample collection along east coast of Australia and Daintree Rainforest

15 To assess Australian bats for the presence of unidentified viruses, samples including scat,  
16 blood, urine, and oral swabs were collected from multiple species of bats, including 373  
17 bats across Hervey Bay, Boonah, Byron Bay, Alstonville, Redcliffe, and Nambucca Heads,  
18 and 106 bats in the Daintree rainforest, between 2007 and 2014. Bats were  
19 macroscopically identified. To confirm the species of origin of each sample, nucleic acids  
20 were extracted as described below in supplementary methods section "RT-qPCR for  
21 presence of Gammaretroviral nucleic acids in bat samples", and a cytochrome B gene  
22 TaqMan PCR assay for species determination was performed, as described in (1).

23

#### 24 Metagenomic analysis of viral nucleic acid in bat samples

25 For samples collected in the Daintree rainforest, total RNA was extracted with the  
26 QIAamp viral RNA Mini Kit (QIAGEN) and DNA was digested using the TURBO DNA-free kit  
27 (Thermo Fisher Scientific), all according to the manufacturer's protocol. Other samples  
28 were processed as follows: Briefly, PBS homogenized bat feces was enriched for viral  
29 particles using a discontinuous sucrose gradient (2, 3). Total RNA was extracted with the  
30 QIAamp viral RNA Mini Kit (QIAGEN) except carrier RNA (poly-A) was omitted from Buffer  
31 AVL and genomic DNA was removed with DNase I digestion prior to RNA extraction as  
32 previously described (2). Random RT-PCR amplification and double-stranded cDNA was  
33 prepared as previously described (2) except K8N random primers were replaced with (5'  
34 GTTCCAGTAGGTCTC NNN NNNN-3') for cDNA synthesis and 5'-A\*G\*C\*A\*C  
35 TGTAGGTTCCAGTAGGTCTC-3' for double-stranded cDNA amplification (4). Sequencing

## SI Appendix

1 libraries were generated using Illumina Nextera-XT library construction, sequencing was  
2 performed on the Illumina MiSeq platform, and bioinformatics including FASTQ paired-  
3 end read quality control and *de novo* assembly was performed as described previously  
4 (3). KoRV-related viral contigs were identified by BLASTn and BLASTx analysis using the  
5 assembled contigs as query sequences against the NCBI nucleotide collection database  
6 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) using default parameters.

7

8 Isolation and assembly of the HPG genome sequence

9 Partial HPG sequences were initially identified in the metagenomic analysis of a *P. alecto*  
10 scat sample obtained in Hervey bay, in 2011, using the method described above in  
11 “Metagenomic analysis of viral nucleic acid in bat samples”, employing random RT PCR  
12 amplification. To generate the complete HPG genome sequence we used a modified  
13 single-cell whole transcriptome amplification (WTA) procedure for detecting ultra-low-  
14 copy viral RNA, and *de novo* sequence assembly pipeline.

15 Total RNA was purified from the same fecal sample and using the same procedure as  
16 described in the initial metagenomics study with the exception that DNase I digestion was  
17 performed after extraction of total RNA with the QIAamp viral RNA mini kit (QIAGEN) and  
18 final purified total RNA was eluted in a total volume of 20  $\mu$ L. Concentration of both DNA  
19 and total RNA was determined with the Qubit HS DNA and HS RNA assays (Invitrogen)  
20 read on the Qubit 3.0 fluorimeter (Invitrogen) and was below the level of detection for  
21 both assays, < 0.5 ng/mL and < 20 ng/mL, respectively.

22 Eight microliters of purified total RNA was converted to cDNA, ligated, and then  
23 isothermally amplified using the REPLI-g WTA Single Cell kit (QIAGEN), according to the  
24 manufacturer’s protocol, except the amplicons were purified using the Genomic DNA  
25 Clean and Concentrator-10 kit (Zymo Research, Irvine, USA). Briefly, 60  $\mu$ L of amplified  
26 cDNA was diluted in a total volume of 100  $\mu$ L with 40  $\mu$ L of 10 mM Tris pH 8.5 (QIAGEN),  
27 to which was added 200  $\mu$ L of DNA binding buffer and processed according to the  
28 manufacturer’s protocol. Purified cDNA was eluted sequentially with 20  $\mu$ L and 15  $\mu$ L of  
29 70°C pre-heated 10 mM Tris pH 8.5 (QIAGEN) buffer for 2 min prior to elution. The total  
30 amount of amplified product was 3.15  $\mu$ g (89.95 ng/ $\mu$ L), as determined with the Qubit  
31 dsDNA BR Assay Kit (Invitrogen).

32 Dual indexed libraries were prepared according to the Illumina Nextera XT DNA Library  
33 Prep Kit (Illumina). Library concentration was determined with the Qubit HS dsDNA assay  
34 (Invitrogen). Library quality and distribution was determined by loading 3 ng of sample  
35 on an Agilent Technology 2100 Bioanalyzer using the Agilent High Sensitivity DNA assay.

## SI Appendix

1 Libraries were normalized, denatured then diluted to a final concentration of 10 pM with  
2 HT1 buffer (Illumina) and spiked with 1% PhiX control library (Illumina). Libraries were  
3 sequenced on the Illumina MiSeq platform, using the MiSeq Reagent v2 kit (300 cycles),  
4 generating 150 bp paired end reads.

5 Illumina FASTQ paired-end reads (8,162,956) were imported into CLC Genomics  
6 Workbench v10.1.1 using default Illumina import parameters, and then trimmed for size,  
7 quality, and ambiguous bases using default parameters except for the following: Quality  
8 Limit = 0.01, Ambiguous limit = 2 and Minimum number of nucleotides in reads = 30. Host  
9 reads (*Pteropus alecto* draft genome assembly GenBank assembly accession  
10 GCA\_000325575.1 and mitochondrion Genbank accession NC\_023122) were removed by  
11 read mapping using default settings on the CLC Genomics Grid Worker v7.0.1 except  
12 "Length" and "Similarity" fractions were both set to 0.9.

13 Host subtracted, trimmed, FASTQ paired end reads (6,836,522) were imported into  
14 Geneious v10.2.2 with "Read Technology" set to Illumina and the default paired end insert  
15 size selected (500 nt). Reads were error corrected and normalised with the Kmer-based  
16 tool, BBNorm v37.25 (<https://sourceforge.net/projects/bbmap/>), using default settings  
17 except "Minimum Depth" normalization was increased from 6 to 40. Normalized paired-  
18 end reads (295,939) were *de novo* assembled using default settings for SPAdes v3.10.0 (5,  
19 6) selecting the "Multi Cell" Data Source option with error correction. The *de novo*  
20 assembled contig (8,040 bp) was verified by mapping trimmed reads to obtain the final  
21 genome sequence (8,030 bp).

22

23 Identification and assembly of KoRV related viruses in publicly available databases  
24 To identify KoRV-related gammaretroviruses in public databases, data from the Sequence  
25 Read Archive (SRA) derived from bat RNA and DNA were subjected to SRA-BLAST analysis  
26 ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&BLAST\\_SPEC=SRA&LINK\\_LOC=blasttab](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&BLAST_SPEC=SRA&LINK_LOC=blasttab)). The genome sequences of the assembled Hervey  
28 pteropid gammaretrovirus (HPG) was used as the query sequence. The algorithm  
29 parameters set were to: Program = blastn, Max target sequences = 1000, Expect threshold  
30  $< 1 \times 10^{-10}$ , word size = 11, match score = 2, mismatch cost = -3, gap costs = existence 5  
31 extension 2, no filtering or masking. SRA that contained reads aligning to the query  
32 sequences were from the Chinese microbats *Rhinolophus hipposideros* (Genbank:  
33 SRX1059482 & SRX1059481) and *Hipposideros larvatus* (Genbank: SRX1059446).  
34 Sequencing reads aligning to the query sequences were downloaded and assembled into  
35 the partially complete genomes of RhGRV and HIGRV as follows: Reads were downloaded  
36 and assembled using the CLC Genomics Workbench 11.0 (QIAGEN, Aarhus, Denmark)

## SI Appendix

1 "Assemble Sequences" tool into a contiguous consensus sequence using the following  
2 parameters: Minimum aligned read length = 20, alignment stringency = high, conflicts =  
3 Vote (A, C, G, T). Assembled contigs were subsequently used as a new query in an  
4 otherwise identical BLASTn search against the same SRA. This process was iteratively  
5 repeated until all contigs could be extended out until they overlapped with each other or  
6 reached a region of zero read coverage. The extended and overlapping contigs were  
7 assembled by alignment against the reference/query HPG genome sequence in CLC  
8 Genomics Workbench.

9

### 10 Annotation of Retroviral Genomes

11 Bat retroviral genome sequences were annotated using CLC Genomics Workbench by  
12 alignment using MUSCLE, and comparison against the genomes of KoRV-A (Genbank:  
13 AF151794) and M-MLV (Genbank: NC001501).

14

### 15 HPG specific analysis of Pteropid genomes

16 To determine whether HPG could be identified as an endogenous retrovirus within the  
17 genomes of Pteropid bats, we performed *in vitro* and *in silico* analyses. For the *in silico*  
18 analysis, we performed a BLAST analysis using CLC Genomics Workbench, of the genomes  
19 of *Pteropus alecto* (Genbank: PRJNA232518) and *P. vampyrus* (Genbank: PRJNA275879)  
20 using the HPG genome as the query sequence, with the algorithm parameters: Expect  
21 threshold =  $1 \times 10^{-10}$ ; word size = 11; Low complexity regions filtered.

22 For the *in vitro* analysis, two sources of *P. alecto* genomic DNA were analyzed by PCR for  
23 the presence of HPG. The first source of *P. alecto* genomic DNA was extracted from pooled  
24 heart and muscle tissue of a male bat captured in Brisbane QLD, November 2008. DNA  
25 was extracted using the QIAGEN Genra Puregene Tissue Kit (QIAGEN) and further  
26 purified with the MO BIO Powerclean DNA clean up kit (MO BIO, Carlsbad, USA) and then  
27 AMPure XP beads (Beckman Coulter, Brea, USA), all of which were performed according  
28 to the manufacturer's protocol. The second source of *P. alecto* genomic DNA was  
29 extracted from a primary kidney (PaKi) cell line (7) using the QIAamp DNA Mini kit  
30 (QIAGEN), following the manufacturer's protocol.

31 HPG positive controls were derived from two regions within HPG [1.34 & 1.55 kb in length  
32 (Supplementary Figure 3), which were identified in the NGS metagenomics analysis  
33 described above in "Metagenomic analysis of viral nucleic acid in bat samples". The two  
34 sequences were amplified by PCR with the QIAGEN HotStar HiFidelity polymerase

## SI Appendix

1 according to the manufacturer's protocol from random RT PCR amplified bat scat sample  
2 used for the initial metagenomics NGS using primers designed to amplify each sequence  
3 (Supplementary Table 5). The two amplicons were cloned into the pCR4-TOPO  
4 (Invitrogen) vector using the Zero Blunt TOPO PCR cloning kit (Invitrogen) according to  
5 the manufacturer's protocol. Primer sensitivity tests were conducted to determine the  
6 template copy number required for a band to be present in this analysis using the 1.3 &  
7 1.5 kb sequence primers (Supplementary Table 5) based on plasmid controls. Primers  
8 specific for the 1.3 & 1.5 kb HPG sequences were found to be sensitive to  $1.4 \times 10^8$  and  
9  $9.0 \times 10^3$  copies, respectively.

10

### 11 Phylogenetic analysis

12 To determine the evolutionary relationships among KoRV-related gammaretroviruses, we  
13 performed phylogenetic analyses using aligned complete genome nucleotide sequences  
14 (Supplementary Table 2). Accordingly, a multiple sequence alignment of 19 complete  
15 genomes was performed using a combination of MAFFT (8) and MUSCLE algorithms (9).  
16 Following alignment, regions of ambiguous and uncertain alignment were removed using  
17 Gblocks (10). This resulted in final alignment of 6,925 nt that was used to infer  
18 evolutionary relationships. A phylogenetic tree of these data was estimated using the  
19 maximum likelihood (ML) method available in the PhyML program (11), assuming a GTR  
20 model of nucleotide substitution with a proportion of invariant sites (I) and a gamma  
21 distribution of among-site rate variation ( $\Gamma$ ). To determine the robustness of each node,  
22 a bootstrap resampling analysis (1,000 replications) was performed using the same  
23 nucleotide substitution model. The *Mus caroli* ERV, McERV (Supplementary Table 2),  
24 sequence was used as an outgroup to root the tree.

25

### 26 Generation of HPG proviral sequence and synthesis of HPG proviral expression 27 construct

28 To generate a synthetic HPG provirus *in silico*, with the 5' and 3' long terminal repeats  
29 (LTRs) necessary for retroviral gene expression, the HPG genome sequence was modified  
30 by copying the unique 5' (U5) region and inserting it immediately following the 3' terminal  
31 repeat (R) region, and copying the unique 3' (U3) region and inserting it immediately prior  
32 to the 5' terminal R region. The HPG proviral sequence was chemically synthesized  
33 (GenScript, Nanjing, China) and inserted within the pCC1BAC cloning plasmid (GenBank:  
34 EU140750) at the *EcoRI* 333 site, generating the pCC1-HPG proviral expression construct.

35

## SI Appendix

### 1 Cell cultures

2 Human embryonic kidney (HEK) 293T cells (kindly provided by Richard Axel, Columbia  
3 University), human epithelial cervical adenocarcinoma (HeLa) cells (NIH AIDS Research  
4 and Reference Reagent Program, Division of AIDS, NIAID, NIH), NIH Swiss mouse embryo  
5 (NIH/3T3) fibroblast cells (American Type Culture Collection), a *P. alecto* kidney (PaKi)  
6 cell line (7), and HEK cells that express M-MLV Gag and Pol polyproteins (GP2-293 cells;  
7 Takara Bio) were utilized. 293T and HeLa cells were authenticated using the Promega  
8 GenePrint 10 system performed by the Australian Genome Research Facility (AGRF). The  
9 short tandem repeat (STR) profile was used to search the ATCC STR  
10 database [https://www.atcc.org/en/STR\\_Database.aspx](https://www.atcc.org/en/STR_Database.aspx) and the DSMZ-German  
11 Collection of Microorganisms and Cell cultures  
12 database [https://www.dsmz.de/services/human-and-animal\\_cell-lines/online-str-](https://www.dsmz.de/services/human-and-animal_cell-lines/online-str-analysis)  
13 [analysis](https://www.dsmz.de/services/human-and-animal_cell-lines/online-str-analysis). All cell cultures were maintained at 37°C, 5% CO<sub>2</sub> in Dulbecco's modified  
14 Eagle's medium (DMEM; Thermo Fisher Scientific) supplemented with heat-inactivated  
15 fetal bovine serum (100mL/l; Invitrogen), glutamine (292 mg/mL; Invitrogen), and the  
16 antibiotics penicillin (100 units/mL; Invitrogen) and streptomycin (100 units/mL;  
17 Invitrogen), with the exception of the PaKi cells for which DMEM was substituted with  
18 DMEM/F-12 1:1 medium (Thermo Fisher Scientific).

19

### 20 Transfection of 293T cells for generation of HPG & M MLV viral particles

21 HPG and M-MLV viral particles were generated by transfection of 293T cells with the  
22 pCC1-HPG and pNCS (Addgene: 17362) plasmids, respectively. 293T cells were  
23 transfected at 50% confluency with 20 µg or pCC1-HPG or 10 µg of pNCS, using  
24 Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's  
25 instructions. Untransfected cells and cells transfected with 20 µg of either the empty  
26 plasmid pCR2.1 (Thermo Fisher Scientific) or the empty plasmid pcDNA3.1 (Invitrogen),  
27 were used as controls. Transfected cells were incubated at 37°C, 5% CO<sub>2</sub> for 48 h, and  
28 then virion-containing supernatants were collected and clarified by centrifugation at 200  
29 x g for 5 min. Virus production was determined by quantifying virion-associated RT  
30 activity, as previously described (12).

31

### 32 Electron Microscopy and viral particle morphology

33 HPG, M-MLV, and pcDNA3.1 transfected cell cultures were generated as described above,  
34 in "Transfection of 293T cells for generation of HPG & M-MLV viral particles", using  
35 Lipofectamine 2000 (Invitrogen). Untransfected cells were used as a control. For thin

## SI Appendix

1 section electron microscopy (EM), cells were pelleted and immersed in 2.5%  
2 glutaraldehyde in 0.1 M Sorenson's phosphate buffer pH 7.2, (300 mOsmol/kg) for 1 h at  
3 room temperature. After washing with Sorensen's phosphate buffer the cells were fixed  
4 with 1% osmium tetroxide for 1 h and dehydrated in a graded ethanol series at room  
5 temperature. Samples were embedded in Spurr's resin (ProSciTech, Australia) according  
6 to the manufacturer's protocol. Ultrathin sections were obtained using a Leica ultracut  
7 UCT Microtome and stained with saturated uranyl acetate in 50% ethanol and lead  
8 citrate. All prepared grids were examined using a Philips CM120 or JEOL JEM-1400  
9 transmission electron microscope at 120kV. Electron micrographs representing negative  
10 control untransfected and mock transfected cells did not reveal the presence of viral-like  
11 particles (Supplementary Figure 8).  
12

### 13 Reverse transcriptase divalent cation preference

14 To evaluate the divalent cation preference of HPG reverse transcriptase (RT), we  
15 performed a virion associated RT assay using the gammaretroviral RT co-factor,  
16 manganese ( $Mn^{2+}$ ), and the lentiviral RT co-factor, magnesium ( $Mg^{2+}$ ). We compared HPG  
17 RT activity to HIV-1 RT activity in the presence of each co-factor. 293T cells were co-  
18 transfected with different quantities of HPG (pCC1-HPG; 0.04 - 0.22 pmol), M-MLV (pNCS;  
19 0.04 - 0.22 pmol), or HIV-1 (pNL4-3; 0.01 - 0.04 pmol), with the total mass of transfected  
20 plasmid DNA equalized by the addition of the empty expression vector pCR2.1 (Thermo  
21 Fisher Scientific). Cells transfected only with pCR2.1 (0.04 - 0.22 pmol) were used as  
22 controls. Transfections were performed using Lipofectamine 2000 according to the  
23 manufacturer's protocol. Transfected cell cultures were incubated at 37°C, 5% CO<sub>2</sub> for 48  
24 h, and then virion-containing supernatants were collected and clarified by centrifugation  
25 at 200 x g for 5 min. To assess magnesium co-factor usage by the viral reverse  
26 transcriptase, we performed a virion-associated RT activity, as previously described (12).  
27 To assess manganese co-factor usage, the virion-associate RT activity assay was modified  
28 by the replacement of magnesium with 0.1M manganese.

29

### 30 Replication kinetics assay

31 To determine the cell tropism of HPG compared to ecotropic M-MLV, HPG and M-MLV  
32 transfected cell cultures were generated as described in the previous section  
33 "Transfection of 293T cells for generation of HPG & M-MLV viral particles". Virion-  
34 containing supernatants were normalized by virion-associated RT activity, using  
35 manganese as the cofactor, as described in (12) except with the above-mentioned  
36 modifications. Virion containing supernatants were used to infect human HeLa, mouse

## SI Appendix

1 3T3, and bat PaKi cells. Untransfected 293T cell culture supernatant was used as a control.  
2 Cells were seeded in a 96-well plate (Sarstedt, Nümbrecht, Germany) at a density of  
3 15,000 cells/well. When cells reached ~50% confluency, the media was replaced with 225  
4  $\mu\text{L}$  of normalized HPG or M-MLV virion containing supernatant with the addition of DEAE-  
5 Dextran (Sigma-Aldrich) at a final concentration of 10  $\mu\text{g}/\text{mL}$ . Cells were incubated for 6  
6 h, and then the supernatant was removed, cells were washed twice with phosphate  
7 buffered saline (PBS), and 250  $\mu\text{L}$  of DMEM medium was added. To assess cell culture  
8 supernatants for the release of viral particles, 20  $\mu\text{L}$  samples were collected from the  
9 supernatant of each well at 24 h intervals for 5 days. To measure the presence of HPG or  
10 M-MLV virions in the cell culture supernatant samples, we performed a virion-associated  
11 RT activity assay.

12

13 Generation of gammaretroviral *env* gene expression constructs for pseudotyping  
14 Cloning constructs pUC57-GALV-*env* and pUC57-KoRV-A *env* encoding the Envelope  
15 proteins of GALV (Genbank: KT724048) and KoRV-A (Genbank: NC039228), respectively,  
16 were chemically synthesized (GenScript). The Envelope sequences were enzymatically cut  
17 from the cloning plasmids using *Bam*HI and *Xho*I enzymes (New England Biolabs) and  
18 ligated into the mammalian expression vector pcDNA3.1 (Invitrogen), using T4 DNA ligase  
19 (New England Biolabs) following the manufacturer's protocols, generating the expression  
20 vectors pcD-GALV-*env* & pcD-KoRV-A-*env*. To generate an expression plasmid for the HPG  
21 Envelope protein, the HPG *env* gene was amplified from the pCC1-HPG plasmid using  
22 primers (HPG-*env*-F and HPG-*env*-R; Supplementary Table 5) designed to anneal  
23 upstream of the cytoplasmic accumulation element (13) and downstream of *env* stop  
24 codon. To facilitate directional cloning, *Eco*RI and *Xba*I restriction sequences were  
25 incorporated into the forward and reverse primers, respectively. The HPG *env* gene was  
26 amplified using the Phusion High-Fidelity PCR Kit (New England Biolabs) according to the  
27 manufacturer's instructions and using 50 ng of pCCI-HBPG template and 0.5  $\mu\text{M}$  of each  
28 forward and reverse primer in a 20  $\mu\text{L}$  reaction. The HPG *env* amplicon was ligated into  
29 the pcDNA3.1 vector using T4 DNA ligase (New England Biolabs) following the  
30 manufacturer's protocols, using the restriction enzymes *Eco*RI and *Xba*I (New England  
31 Biolabs), generating the expression plasmid pcD-HPG-*env*. The sequences of all  
32 expression plasmids were confirmed by Sanger sequencing. Expression plasmids for other  
33 Envelope proteins including VSV-G (pVSV-G), ecotropic MLV (pEco), 4070A amphotropic  
34 MLV (pAmpho), and 10A1 amphotropic MLV ('dualtropic' MLV, p10A1) were obtained  
35 from the Retro-X Universal Packaging System (Takara Bio).

36

## SI Appendix

1 Generation of pseudotyped retroviruses and assay of host cell tropism

2 To determine the tropism of HPG in comparison to M-MLV, pseudotyped viral particles  
3 were produced using the Retro-X Universal Packaging System (Takara Bio). To generate  
4 viral particles pseudotyped with each gammaretroviral Envelope protein, GP2-293 cells  
5 that express M MLV Gag and Pol were transfected with the reporter vector, pQCLIN, and  
6 the Envelope expression vector, pVSV-G, pEco, pAmpho, p10A1, pcD-HPG-env, pcd-KoRV-  
7 A env, or pcD-GALV-env, to generate viral particles pseudotyped with Env derived from  
8 VSV-G, Ecotropic MLV, Amphotropic MLV, HPG, KoRV\_A, and GALV, respectively. T75  
9 tissue culture flasks (Nunc/Thermo Fisher Scientific) were seeded with  $2.1 \times 10^6$  GP2-293  
10 cells. Cells were incubated for 16 h and then transfected with 5  $\mu\text{g}$  of pQCLIN and 10  $\mu\text{g}$   
11 of the Env expression plasmid using Lipofectamine 2000 (Thermo Fisher Scientific)  
12 according to the manufacturer's protocol. Transfected cells were incubated for 48 h, and  
13 then viral particles were collected from clarified supernatants and concentrated using the  
14 Lenti-X Concentrator (Takara Bio).

15 Human HeLa and mouse 3T3 cells were seeded in 96-well plates (Nunc) at a density of  
16 20,000 cells/well, and incubated for 16 h. Following incubation, equal volumes of  
17 pseudotyped viral particles containing DEAE-Dextran (Sigma Aldrich) at a final  
18 concentration of 10  $\mu\text{g}/\text{mL}$  was added to the cells. Cells were incubated for 48 h and then  
19 viral cell entry was determined by the presence of blue cell-forming units as previously  
20 described (12).

21

22 RT qPCR for detecting the presence of gammaretroviral nucleic acids in bat samples  
23 Nucleic acids from 50  $\mu\text{L}$  PBS resuspended bat scat samples were extracted on a  
24 KingFisher Flex Purification System (Thermo Fisher Scientific) using the Applied  
25 Biosystems MagMAX-96 Viral RNA Isolation Kit (Thermo Fisher Scientific) according to  
26 manufacturer's protocol except final purified nucleic acids were eluted in 50  $\mu\text{L}$  nuclease  
27 free water instead of 90  $\mu\text{L}$  elution buffer. No DNase I digestion was performed. 'Broad'  
28 primers (HPG-rel-F and HPG-rel-R) were designed to bind to HPG, FFRV1, MmGRV, and  
29 SaGRV, in the region upstream of the *gag* gene. Primers specific for HPG (HPG gag F and  
30 HPG-gag-R) were designed to bind to within the *gag* gene. Sequences for all primers are  
31 provided in (Supplementary Table 5). The presence of viral nucleic acids was determined  
32 by reverse transcriptase quantitative PCR (RT-qPCR) using the Power SYBR Green RNA to-  
33 CT 1-Step Kit (Thermo Fisher Scientific), with a cycle threshold of 36. Reaction mixtures  
34 contained 4.5  $\mu\text{L}$  of purified RNA, 200 nM of each primer, and RT enzyme mix. Reactions  
35 were performed in either 384- or 96-well plates on the QuantStudio 7 Flex qPCR machine  
36 (Thermo Fisher Scientific). To determine if nucleic acid amplification was from RNA or

## SI Appendix

1 DNA, an identical reaction was performed where the RT enzyme mix was excluded, to  
2 prevent amplification from RNA. Cycling conditions were as follows: 1x cycle of 48°C for  
3 30 min, then 95°C for 10 min, and 40x cycles of 95°C for 15 s then 60°C for 1 min.

4

5 Generation of HPG Envelope protein for serological assays

6 A codon-optimized (*Homo sapiens*) synthetic gene, encoding the predicted HPG retrovirus  
7 Env ectodomain (Glu<sup>38</sup>-Ser<sup>603</sup>), was chemically synthesized (GeneArt, Regensburg,  
8 Germany). The synthetic gene incorporated an in-frame 5' *NheI* site, a C-terminal His<sub>6</sub> tag  
9 followed by a termination codon and 3' *XbaI* site. In addition, the putative SU-TM cleavage  
10 site, Arg<sup>473</sup>LeuLysArg, was ablated by substitution with Ser<sup>473</sup>LeuGlnSer. The synthetic  
11 gene was ligated downstream of the tissue plasminogen activator leader sequence in the  
12 pcDNA3-based vector, pcE2<sup>661</sup>myc (14) to give pcHPG-Env<sup>603</sup>. For expression, 293-F cells  
13 were transfected with pcHPG-Env<sup>603</sup> using 293fectin (Thermo Fisher Scientific). At 24 h  
14 post-transfection, 0.5% (w/v) lupin peptone and 0.02% (w/v) pluronic F-68 were added to  
15 cells. Proteins were harvested following 3-5 days of incubation by centrifugation at 1,500  
16 x *g* for 5 min, followed by filtration through a 0.45 μM filter before storage at 4°C. The  
17 Envelope protein was purified using Talon metal affinity resin (Takara) and 250 mM  
18 imidazole/PBS as the elution buffer. The protein was exchanged into PBS and  
19 concentrated using an Amicon centrifugal filter device (Merck Millipore, Burlington, USA).  
20 SDS-PAGE in the presence and absence of β-mercaptoethanol revealed a single diffuse  
21 band with a molecular weight range of ~80-90 kDa (Supplementary Figure 8), consistent  
22 with the molecular weight predicted from the amino acid sequence (62,805 Da) with 6 N-  
23 linked glycans (~ 18 kDa).

24

25 Generation of anti-HPG Envelope sera

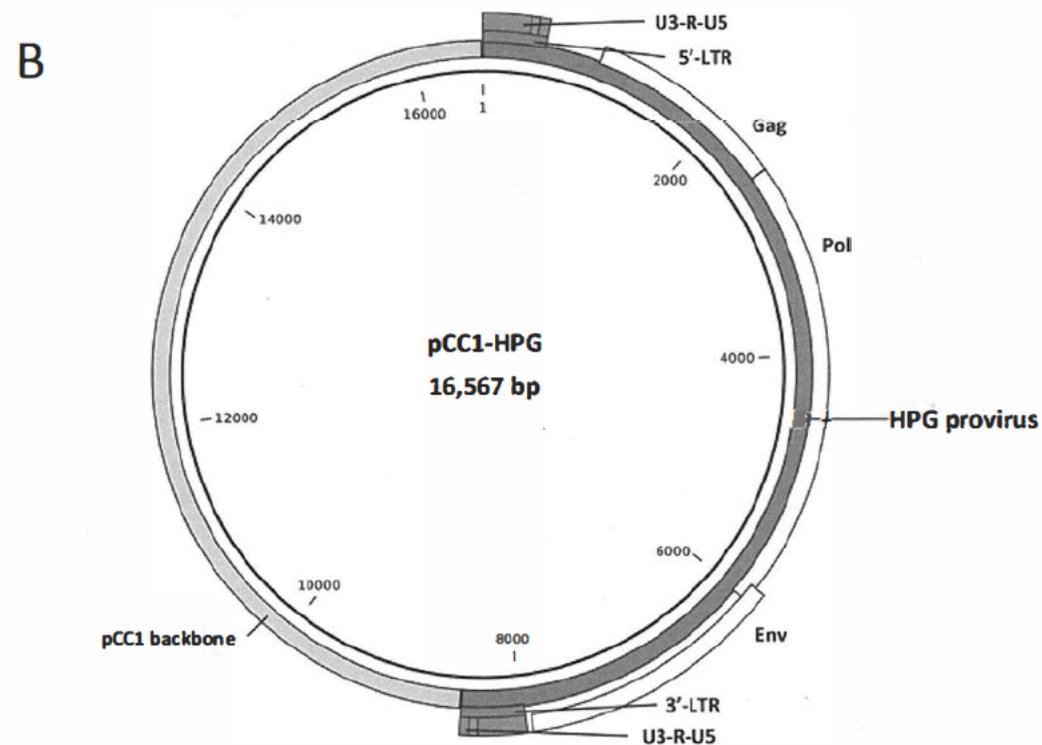
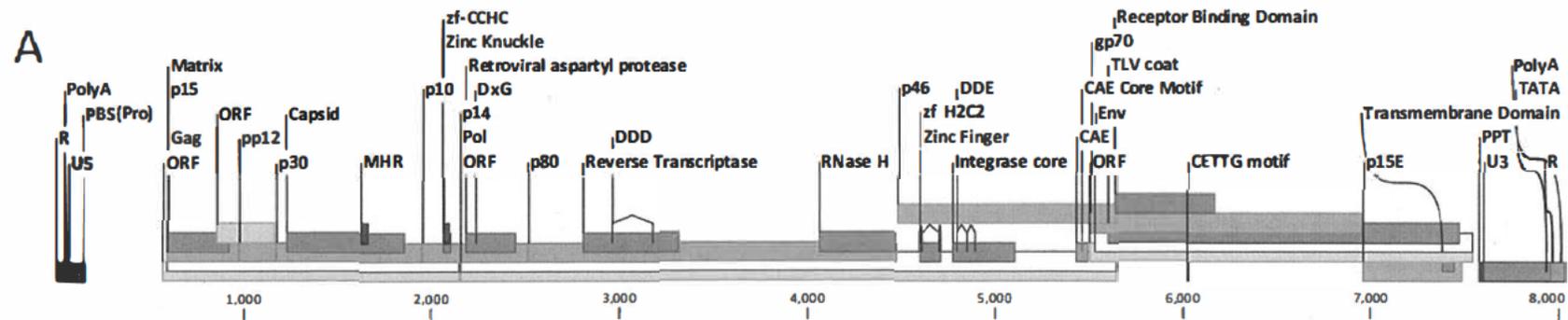
26 Rabbit polyclonal anti-HPG-Envelope sera was generated by the Antibody Services at the  
27 Walter and Eliza Hall Institute Biotechnology Centre (Melbourne, Australia), using the  
28 HPG Envelope protein described above in "Generation of HPG Envelope protein for  
29 serological assays" as the antigen. Rabbits were immunized with 200 μg of the HPG  
30 Envelope protein three times with a 4-week interval between immunizations. Sera was  
31 collected over the course of 68 days and included the collection of pre-bleed sera as a  
32 control.

33

## SI Appendix

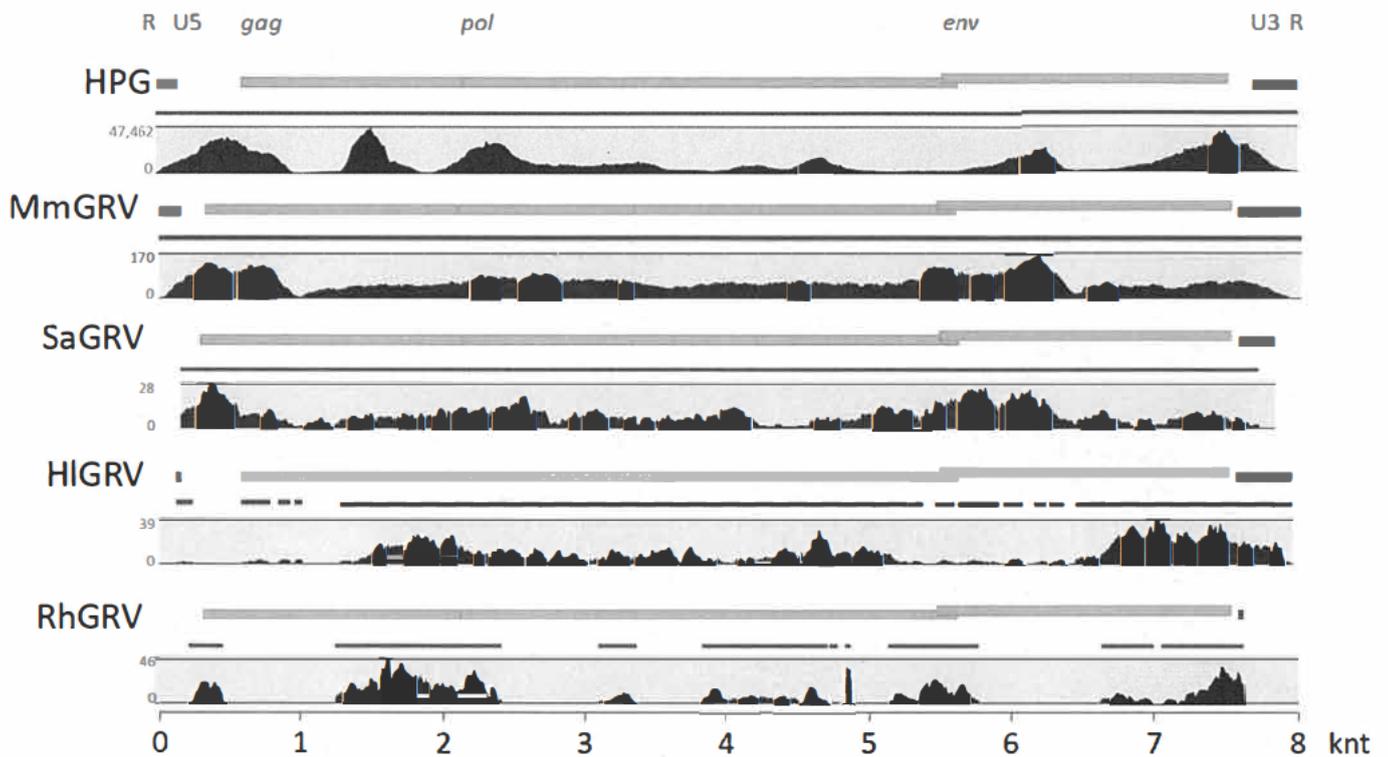
1 Serological assay for the presence of anti HPG antibodies in bats  
2 To assess bat sera for the presence of anti-HPG-Envelope immunoglobulin, 88 bat samples  
3 collected from the East coast of Australia and the Daintree rainforest between 2007 and  
4 2014 were analyzed using a serological assay as previously described (15), with the  
5 following modifications: HPG Envelope gp120 proteins were coupled to magnetic  
6 Luminex beads (Thermo Fisher Scientific) using a mixture of Protein A/Protein G-  
7 biotinylated, and samples were analyzed using a Bio-Plex instrument (Bio-Rad). HIV  
8 SOSIP-Envelope proteins were utilized as a negative control (16). Previous studies  
9 published using this platform have used a threshold of at least three times the mean MFI  
10 of negative sera from other bat species with values below 250 MFI considered negative  
11 (17-20). The same principle was used here to establish a threshold based on an MFI of  
12 250 corresponding to a negative sample with sample MFIs above 1000 considered  
13 positive.  
14

## Supplementary Figures



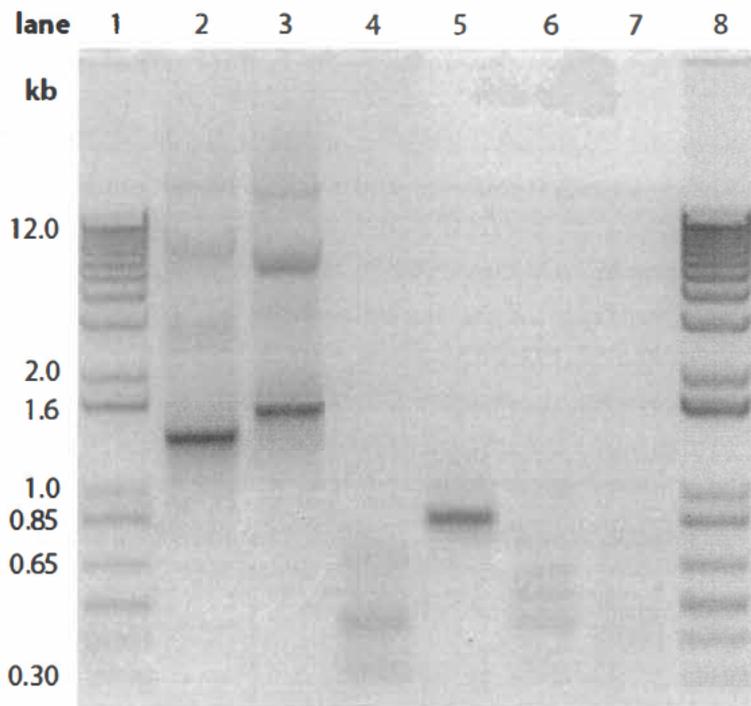
**Supplementary Figure 1.** Schematic diagrams of the Hervey pteropid gammaretrovirus (HPG) expression construct and genome. (A) Annotated HPG genome. Colors represent open reading frames (ORF; grey), nucleotide motifs (dark purple), polyprotein coding sequences (yellow), putative mature proteins (orange), structural and enzymatic protein domains (red), unique 5'/3' regions (U3/U5; green), and repeated regions (R) at both ends of the RNA genome (salmon). Scale is in nucleotides. PBS, primer binding site; PPT, polypurine tract; MHR, major homology region; CAE, cytoplasmic accumulation element; PolyA, polyadenylation signal. (B) The HPG expression plasmid pCC1-HPG. Colors indicate the plasmid backbone (grey), inserted proviral sequence (blue), HPG polyprotein coding sequences (yellow), long terminal repeats (LTRs; pink), unique 5'/3' (U3/U5) regions (green), and repeated regions (R) at both ends of RNA genome (salmon). Scale is in base pairs.

SI Appendix

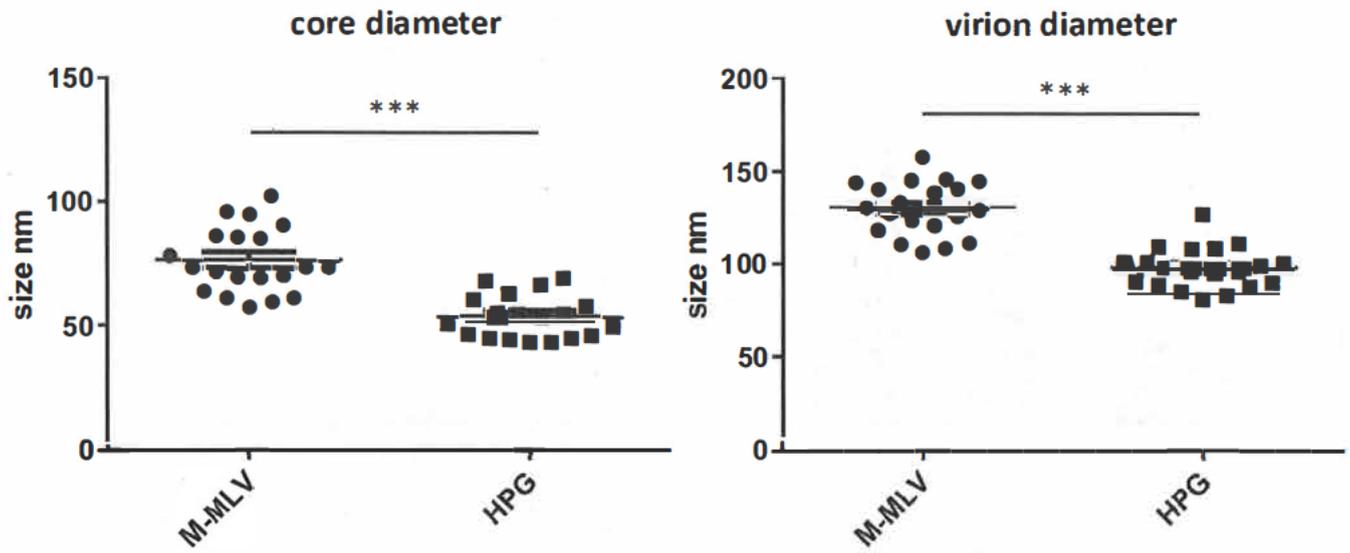


**Supplementary Figure 2.** Sequencing read maps for KoRV-related retroviruses in Australian bats. Colored bars indicate the genomic regions. Red: R, repeated regions at both ends of RNA genome; Dark blue: U5/U3, unique 5' & 3' regions; Light blue: retroviral genes *gag*, *pol*, and *env*. Black graphs represent the read coverage across the retroviral genomes, and gaps in the black lines above the graphs represent regions of zero coverage. Red numbers indicated the minimum and maximum read coverage. HPG, Hervey pteropid gammaretrovirus; MmGRV, *Macroglossus minimus* gammaretrovirus; SaGRV, *Syconycteris australis* gammaretrovirus; HIGRV, *Hipposideros larvatus* gammaretrovirus; RhGRV, *Rhinolophus hipposideros* gammaretrovirus.

SI Appendix

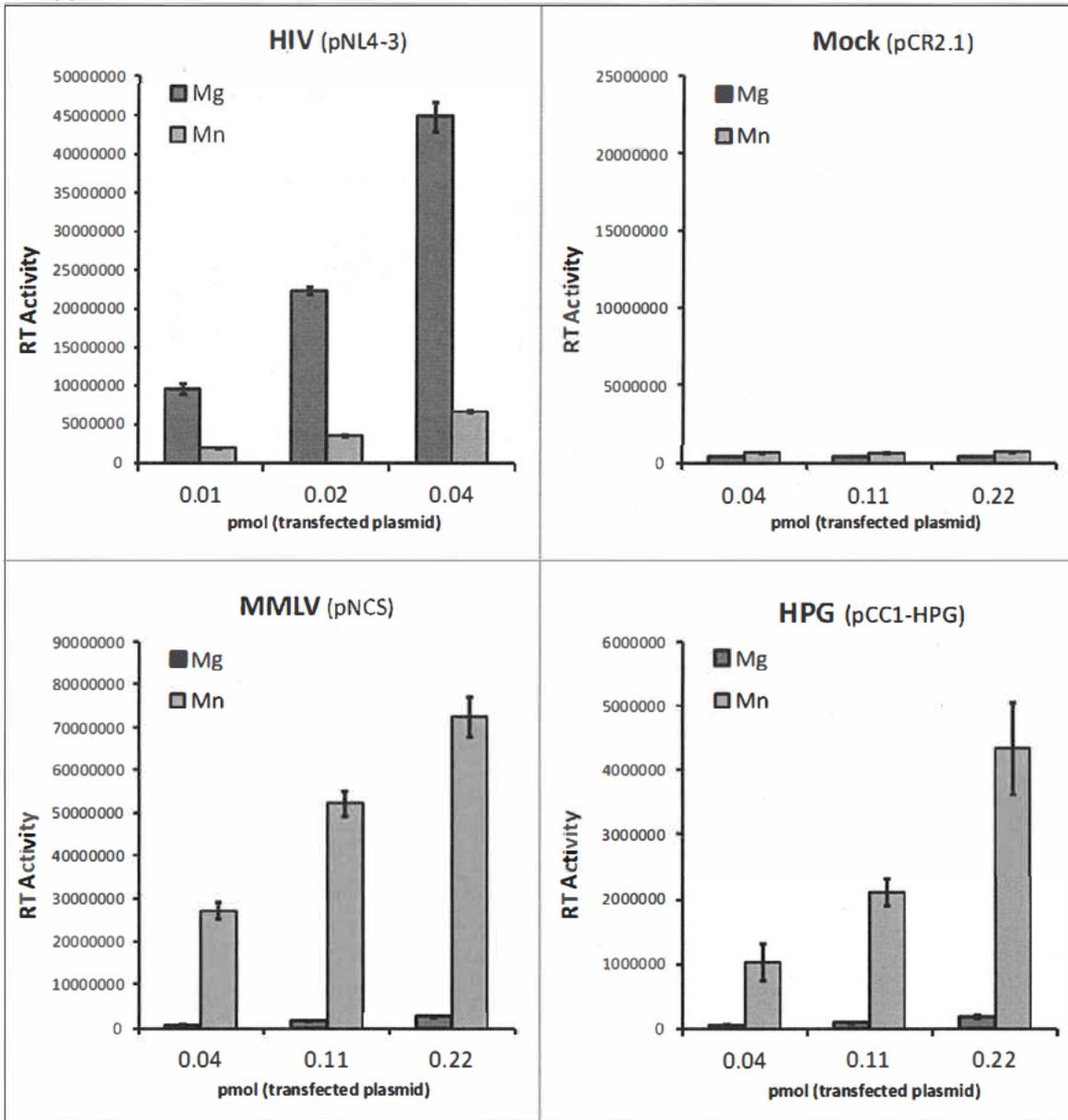


**Supplementary Figure 3.** Assessment of the presence of endogenous Hervey pteropid gammaretrovirus (HPG) in the *Pteropus alecto* genome. PCR primers (Supplementary Table 5) were used to amplify HPG DNA to generate HPG amplicons of 1.34 kb or 1.55 kb and were visualised on a 1% (w/v) agarose gel. Lane 1, 1 Kb Plus DNA Ladder; Lane 2, 80 ng of HPG plasmid was used as template to amplify a 1.34 kb HPG amplicon; Lane 3, 80 ng of HPG plasmid used as a template to amplify a 1.55 kb HPG amplicon; Lane 4, negative control where *P. alecto* genomic DNA (gDNA) was subjected to PCR using M13 primers (M13F/R, Supplementary Table 5); Lane 5, positive control for amplification of a single-copy gene, *APOBEC3Z3* (824 bp) using *P. alecto* gDNA as template; Lane 6, *P. alecto* gDNA subjected to PCR amplification using primers that generate the 1.34 kb segment of HPG; Lane 7, *P. alecto* gDNA subjected to PCR amplification using primers for the 1.55 kb segment of HPG; Lane 8, 1 Kb Plus DNA Ladder. Template gDNA was derived from *P. alecto* tissue and a kidney cell line. A representative gel using *P. alecto* kidney cell line gDNA as the template is shown.



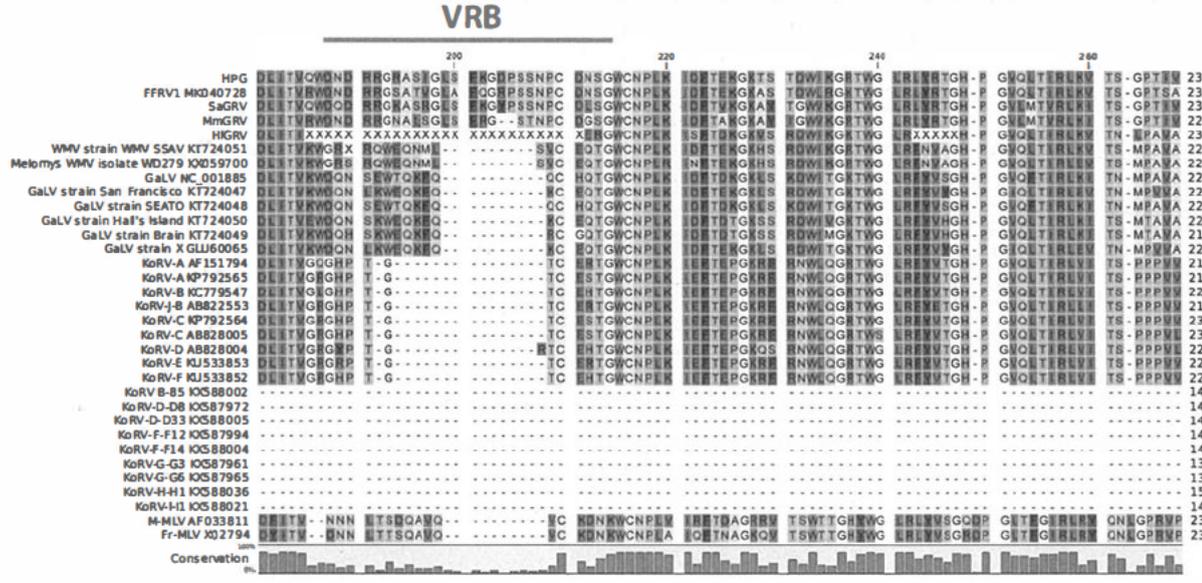
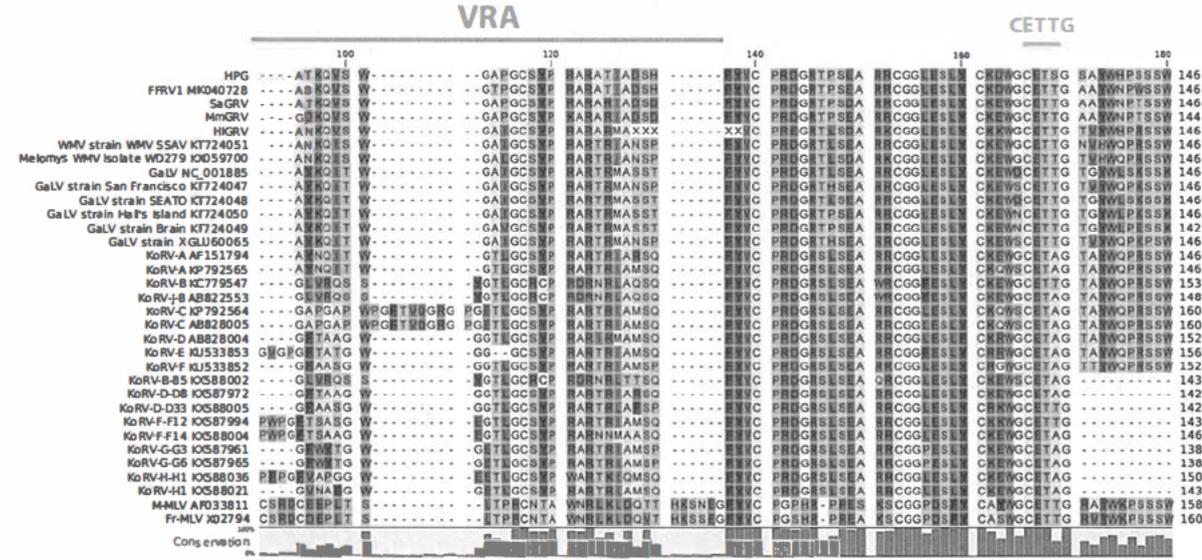
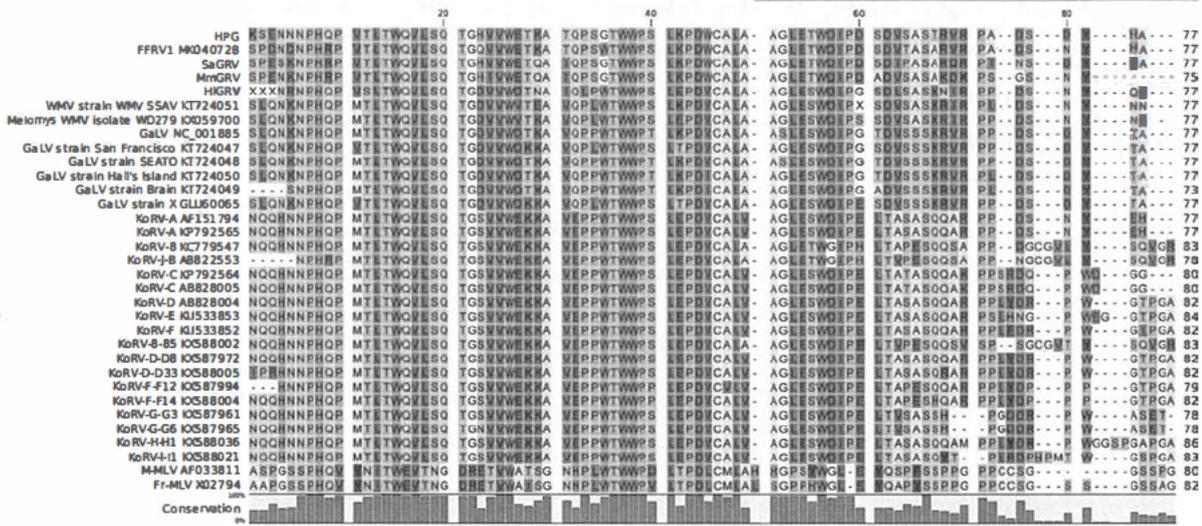
**Supplementary Figure 4.** The diameters of the virion core and virion of Hervey pteropid gammaretrovirus (HPG) and Moloney murine leukemia virus (M-MLV) viral particles. Statistical significance was calculated using the Mann-Whitney test. \*\*\*p value < 0.001, N = 20.

SI Appendix



**Supplementary Figure 5.** Divalent cation preferences of Human immunodeficiency virus (HIV), Moloney murine leukemia virus (MMLV), and Hervey pteropid Gammaretrovirus (HPG) virion-associated reverse transcriptase. Human or murine cell lines were transfected with infectious molecular clones of HPG, MMLV, and HIV (indicated in brackets), or the empty plasmid pCR2.1. Virions collected from the cell culture supernatant were assessed for virion-associated reverse transcriptase (RT) activity assay. Mg, magnesium; Mn, manganese. Error bars represent the standard error of the mean (SE, n = 3).

VRA



Supplementary Figure 6. Multiple sequence alignment of the receptor binding domains of KoRV related viruses.

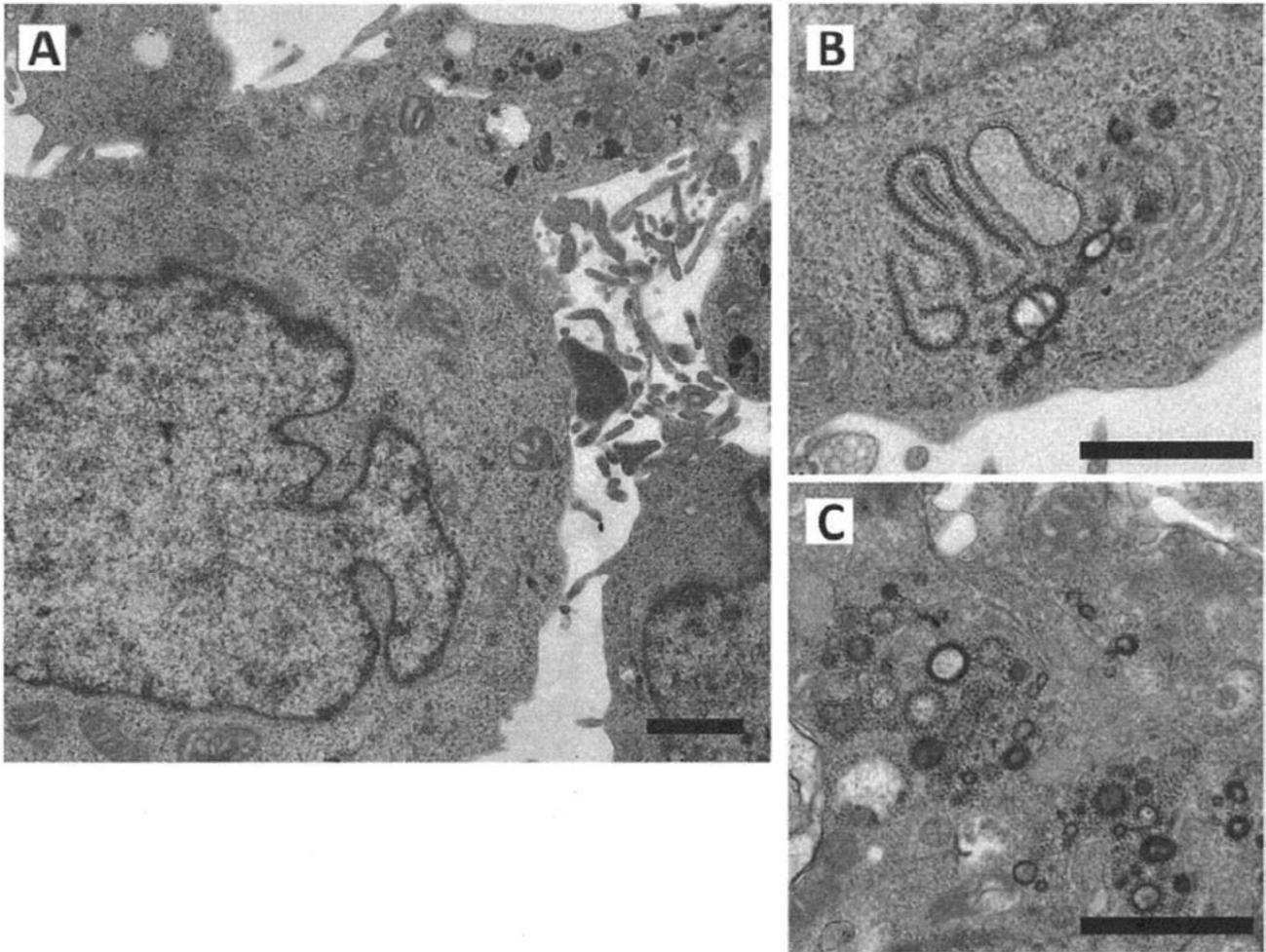
## SI Appendix

The alignment was generated using MUSCLE (9). The non-KoRV-related gammaretroviruses M-MLV and Fr-MLV are included for comparison. Highly variable regions A (VRA) and B (VRB) indicated by the green and red lines, respectively. The CETTG motif is denoted with an orange line. For HIGRV, 'X' indicates regions of zero sequence coverage. The sequence of the region downstream of the CETTG motif, which includes the VRB, is not available for some KoRV sequences.

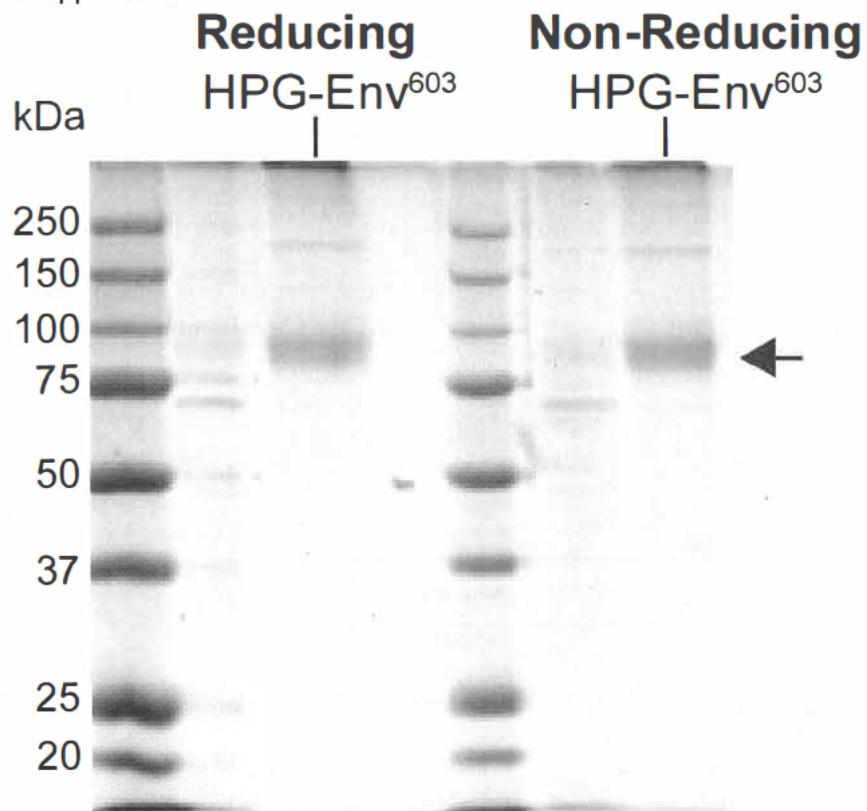
SI Appendix

		550
PiT-1 (Homo sapiens; Human)	ALYLVY	DTGDVSSKVATP IW
PiT-1 (Hylobates lar; Gibbon)	ALYLVY	DTGDVSSKVATP IW
PiT-1 (Phascolarctos cinereus; Koala)	ALYLVY	ETGDVASKVATP IW
PiT-1 (Pteropus vampyrus; Large flying fox)	ALYLVY	DTGDVSSKVATP IW
PiT-1 (Pteropus alecto; Black flying fox)	ALYLVY	DTGDVSSKVATP IW
PiT-1 (Felis catus; Cat)	ALYLVY	DTGDVSSKVATP IW
PiT-1 (Rattus norvegicus; Rat)	ALYLVY	ETRDVTTKEATP IW
PiT-1 (Mus musculus; Mouse)	ALYLVY	KQ - EASTKAATP IW

**Supplementary Figure 7.** Multiple sequence alignment of residues of the PiT-1 Region A of mammals permissive and resistant to GALV infection. The Region A motif of mammalian PiT-1 (SLC20A1) is shown in the red box (amino acid positions 550-557). Residues highlighted in blue and red denote residues in GALV-infection susceptible and resistant mammalian PiT-1 homologs, respectively (21).



**Supplementary Figure 8.** *Electron micrographs (EM) of control untransfected cells and cells mock transfected with the empty vector pcDNA3.1.* A) Transmission electron micrograph of untransfected 293T cells. Cell morphology and ultrastructure is consistent with healthy cells in tissue culture. The cells have typical filopodia extensions of the plasma membrane. Cut in transverse, the filopodia appear round and are in the same size range as retrovirus particles. However, they can be clearly distinguished containing cytoplasmic material (including ribosomes). B – C) Transmission electron micrograph of 293T cells mock transfected with pcDNA3.1. Cells appear healthy and have typical filopodia. There are a few unusual structures (arrows) with appearance of altered endoplasmic reticulum membranes displaying an ordered density of ribosomes either attached to, or in close proximity to these structures. The structures do not show viral morphology and could be formed as an artefact from the transfection procedure. Scale bar represents 1  $\mu\text{m}$ .



**Supplementary Figure 9.** *Hervey pteropid gammaretrovirus recombinant envelope protein expressed from the pCHPG-Env<sup>603</sup> construct.* SDS-PAGE in the presence (reducing) and absence (non-reducing) of  $\beta$  mercaptoethanol and Coomassie blue staining revealed a major diffuse band with a molecular weight range of ~80-90 kDa.

## Supplementary Tables

**Supplementary Table 1.** The sources of KoRV-like viral genomes identified within Australian and Asian bats.

<b>Virus</b>	<b>Bat species</b>	<b>Location</b>	<b>Sample type</b>
HPG	<i>Pteropus alecto</i>	Hervey Bay, Australia	Fecal
MmGRV	<i>Macroglossus minimus</i>	Daintree Rainforest, Australia	Pooled oral and urine
SaGRV	<i>Syconycteris australis</i>	Daintree Rainforest, Australia	Pooled oral and urine
HIGRV	<i>Hipposideros larvatus</i>	Guangxi, China	Pooled fecal and pharyngeal
RhGRV	<i>Rhinolophus hipposideros</i>	Sichuan, China	Pooled fecal and pharyngeal

HPG, Hervey pteropid gammaretrovirus; -GRV = gammaretrovirus

**Supplementary Table 2.** GenBank accession numbers

<b>Abbreviation</b>	<b>Name</b>	<b>Accession</b>
<b>FFRV1</b>	Flying fox retrovirus isolate FFRV1	MK040728
<b>GALV</b>	Gibbon ape leukemia virus	NC_001885
<b>GALV</b>	Gibbon ape leukemia virus strain Brain	KT724049
<b>GALV</b>	Gibbon ape leukemia virus strain Hall's Island	KT724050
<b>GALV</b>	Gibbon ape leukemia virus strain San Francisco	KT724047
<b>GALV</b>	Gibbon ape leukemia virus strain SEATO	KT724048
<b>GALV</b>	Gibbon ape leukemia virus strain X	GLU60065
<b>HIGRV</b>	Hipposideros larvatus gammaretrovirus	
<b>HPG</b>	Hervey pteropid gammaretrovirus	
<b>KoRV-A</b>	Koala retrovirus	AF151794
<b>KoRV-A</b>	Koala retrovirus clone KV522	AB721500
<b>KoRV-A</b>	Koala retrovirus isolate Pci-maex1738	KF786281
<b>KoRV-A</b>	Koala retrovirus isolate Pci-SN265	KF786285
<b>KoRV-B</b>	Koala retrovirus isolate Br2-1CETTG	NC_021704
<b>McERV</b>	Mus caroli endogenous virus	KC460271
<b>MmGRV</b>	Macroglossus minimus gammaretrovirus	
<b>RhGRV</b>	Rhinolophus hipposideros gammaretrovirus	
<b>SaGRV</b>	Syconycteris australis gammaretrovirus	
<b>WMV</b>	Woolly monkey virus strain WMV SSAV	KT724051

SI Appendix

**Supplementary Table 3.** Summary results of HPG seroprevalence

Bat species	Samples tested	$\alpha$ -HPG-Env Ig-positive (MFU > 1000)	
<i>Hipposideros ater</i>	3	0	0.0%
<i>Hipposideros diadema</i>	1	0	0.0%
<i>Macroglossus minimus</i>	10	1	10.0%
<i>Macroglossus syconycters</i>	1	0	0.0%
<i>Nyctimene robinsoni</i>	4	0	0.0%
<i>Nyctophilus bifax</i>	1	0	0.0%
<i>Pteropus alecto</i>	32	12	37.5%
<i>Pteropus conspicillatus</i>	2	2	100.0%
<i>Rhinolophus megaphyllus</i>	23	0	0.0%
<i>Syconycteris australis</i>	11	0	0.0%

HPG, Hervey pteropid gammaretrovirus; MFU, mean fluorescence units

SI Appendix

Supplementary Table 4. Luminex data

Fluorescence intensity				Fluorescence intensity			
Sample	HPG Env <sup>a</sup>	HIV Env <sup>b</sup>	Species / Control	Sample	HPG Env <sup>a</sup>	HIV Env <sup>b</sup>	Species / Note
	1277.5 <sup>c</sup>	270	Positive 1:50000 <sup>d</sup>	45	1076.5	343.5	<i>Pteropus alecto</i>
	6397	269.5	Positive 1:5000 <sup>d</sup>	46	543.5	305	<i>Pteropus alecto</i>
	184	246	Negative sera <sup>e</sup>	47	517.5	332	<i>Pteropus alecto</i>
1	162.5	184	<i>Hipposideros ater</i>	48	481	394	<i>Pteropus alecto</i>
2	91	252	<i>Hipposideros ater</i>	49	861	375	<i>Pteropus alecto</i>
3	334.5	265.5	<i>Hipposideros ater</i>	50	1217	321	<i>Pteropus alecto</i>
4	286.5	252	<i>Hipposideros diadema</i>	51	1429	311	<i>Pteropus alecto</i>
5	129.5	175	<i>Macroglossus minimus</i>	52	1448	304	<i>Pteropus alecto</i>
6	149	187	<i>Macroglossus minimus</i>	53	1581	191	<i>Pteropus conspicillatus</i>
7	141.5	183	<i>Macroglossus minimus</i>	54	4153	162	<i>Pteropus conspicillatus</i>
8	154	188	<i>Macroglossus minimus</i>	55	299.5	165.5	<i>Rhinolophus megaphyllus</i>
9	157	155	<i>Macroglossus minimus</i>	56	110	199	<i>Rhinolophus megaphyllus</i>
10	183	228	<i>Macroglossus minimus</i>	57	285.5	192.5	<i>Rhinolophus megaphyllus</i>
11	229	226.5	<i>Macroglossus minimus</i>	58	165	201	<i>Rhinolophus megaphyllus</i>
12	408	210	<i>Macroglossus minimus</i>	59	144	201	<i>Rhinolophus megaphyllus</i>
13	484	251	<i>Macroglossus minimus</i>	60	185	197.5	<i>Rhinolophus megaphyllus</i>
14	1349	216	<b><i>Macroglossus minimus</i></b>	61	128	182	<i>Rhinolophus megaphyllus</i>
15	196	177	<i>Macroglossus syconycters</i>	62	158	178	<i>Rhinolophus megaphyllus</i>
16	259	216	<i>Nyctimene robinsoni</i>	63	164.5	163	<i>Rhinolophus megaphyllus</i>
17	90.5	107	<i>Nyctimene robinsoni</i>	64	212	173	<i>Rhinolophus megaphyllus</i>
18	160	184	<i>Nyctimene robinsoni</i>	65	342	424	<i>Rhinolophus megaphyllus</i>
19	176	197	<i>Nyctimene robinsoni</i>	66	145.5	214	<i>Rhinolophus megaphyllus</i>
20	147	171	<i>Nyctophilus bifax</i>	67	151	200	<i>Rhinolophus megaphyllus</i>
21	599.5	167	<i>Pteropus alecto</i>	68	159	184.5	<i>Rhinolophus megaphyllus</i>
22	158.5	152	<i>Pteropus alecto</i>	69	168	223.5	<i>Rhinolophus megaphyllus</i>
23	153.5	156	<i>Pteropus alecto</i>	70	98	160	<i>Rhinolophus megaphyllus</i>
24	2813	200.5	<i>Pteropus alecto</i>	71	171	225.5	<i>Rhinolophus megaphyllus</i>
25	1473.5	140	<i>Pteropus alecto</i>	72	117	263.5	<i>Rhinolophus megaphyllus</i>
26	190	142	<i>Pteropus alecto</i>	73	176	202	<i>Rhinolophus megaphyllus</i>
27	178.5	152	<i>Pteropus alecto</i>	74	134	196	<i>Rhinolophus megaphyllus</i>
28	1715.5	160.5	<i>Pteropus alecto</i>	75	163.5	215	<i>Rhinolophus megaphyllus</i>
29	474.5	144	<i>Pteropus alecto</i>	76	120	173	<i>Rhinolophus megaphyllus</i>
30	1359.5	150.5	<i>Pteropus alecto</i>	77	133	210	<i>Rhinolophus megaphyllus</i>
31	318.5	193	<i>Pteropus alecto</i>	78	167	151	<i>Syconycteris australis</i>
32	835	157.5	<i>Pteropus alecto</i>	79	132	161	<i>Syconycteris australis</i>
33	341	164.5	<i>Pteropus alecto</i>	80	185	251	<i>Syconycteris australis</i>
34	193	198	<i>Pteropus alecto</i>	81	138	163	<i>Syconycteris australis</i>
35	516	191	<i>Pteropus alecto</i>	82	133	197	<i>Syconycteris australis</i>
36	218.5	138	<i>Pteropus alecto</i>	83	150.5	187	<i>Syconycteris australis</i>
37	115	128	<i>Pteropus alecto</i>	84	185.5	184	<i>Syconycteris australis</i>
38	574	348.5	<i>Pteropus alecto</i>	85	261	316.5	<i>Syconycteris australis</i>
39	768.5	349.5	<i>Pteropus alecto</i>	86	235.5	206	<i>Syconycteris australis</i>
40	853.5	259	<i>Pteropus alecto</i>	87	265.5	196.5	<i>Syconycteris australis</i>
41	1157.5	306	<i>Pteropus alecto</i>	88	428.5	397	<i>Syconycteris australis</i>
42	1925	248	<i>Pteropus alecto</i>				
43	3435.5	304	<i>Pteropus alecto</i>				
44	3669.5	282	<i>Pteropus alecto</i>				

<sup>a</sup>HPG Env, Hervey pteropid gammaretrovirus trimeric Envelope protein coupled to beads; <sup>b</sup>HIV Env, Human immunodeficiency virus SOSIP Envelope coupled to beads; <sup>c</sup>Red highlighted numbers define positive values defined by fluorescence values >1000; <sup>d</sup>Positive sera, rabbit anti HPG Env; <sup>e</sup>Negative sera, rabbit prebled serum.

Supplementary Table 5. PCR amplification primers used in this study

Target	Primers	Coordinates	Primer sequence (5' > 3')
pCR4-TOPO plasmid	M13F		GTAAAACGACGGCCAG
	M13R		CAGGAAACAGCTATGAC
Mammalian cytB gene	FM-up		CCCCHCCHCAYATYAARCCM
	FM-down		TCRACDGGNTGYCCTCCDATT
<i>Pteropus alecto</i> APOBEC3Z gene	A3Z3F	(2300..2317)	CAGCTCCGAGTCAAAAAG
	A3Z3R	(3104..3123)	AGCGGATCTTGTGATAAAG
HPG <i>pol</i> 1.34 kb sequence	HPG-pol-F1	(4184..4204)	GAACTCATCGCCTTGACTCAG
	HPG-pol-R1	(5521..5500)	AGCAATACCGTCGACCTTTACC
HPG <i>pol</i> 1.55 kb sequence	HPG-pol-F2	(2967..2988)	TCTTCTGCCTCAAACCTGCATCC
	HPG-pol-R2	(4511..4491)	CTGTGGTTTCAGCCAGTACTC
HPG <i>env</i> gene	HPG-env-F	(5354..5375)	GGAAGAATTCAAAGAGGTATACAGACCTGG
	HPG-env-R	(7998..8020)	GCATTCTAGAAGAGGTTTATTAGGTACACGGG
HPG <i>gag</i> 'specific'	HPG-gag-F	(512..532)	AACTCGCTACCGCTTCCATT
	HPG-gag-R	(683..664)	CTTCCACGGACAGGTTGTGA
HPG-related leader 'broad'	HPG-rel-F	(192..212)	CCATCGACGGGAGGTAAGC
	HPG-rel-R	(389..373)	CTGATCCTGGGGCGTCC

cytB, cytochrome B; HPG, Hervey pteropid gammaretrovirus

**Supplementary Table 6.** RT-qPCR survey of Australian bat scat for HPG and related viruses

Site (North to South in descending order)	Samples collected	1st Assay <sup>1</sup> (DNA and RNA amplification)				2nd Assay (DNA amplification only)			
		Positive for HPG-related nucleic acids (DNA/RNA)		Positive for HPG-specific nucleic acids (DNA/RNA)		Positive for HPG-related RNA <sup>2</sup>		Positive for HPG-specific RNA <sup>2</sup>	
Hervey Bay	76	8	10.5%	5	6.6%	2	2.6%	5	6.6%
Dalby	16	1	6.3%	0	0.0%	0	0.0%	0	0.0%
Redcliffe	17	2	11.8%	0	0.0%	0	0.0%	0	0.0%
Sandgate	37	1	2.7%	0	0.0%	0	0.0%	0	0.0%
Boonah	31	12	38.7%	5	16.1%	5	16.1%	5	16.1%
Byron Bay	32	2	6.3%	1	3.1%	1	3.1%	1	3.1%
Alstonville	84	24	28.6%	12	14.3%	4	4.8%	12	14.3%
Nambucca Heads	80	7	8.8%	2	2.5%	1	1.3%	2	2.5%
<b>Total samples</b>	<b>373</b>	<b>57</b>	<b>15.3%</b>	<b>25</b>	<b>6.7%</b>	<b>13</b>	<b>3.5%</b>	<b>25</b>	<b>6.7%</b>

<sup>1</sup>57 samples that were positive in the 1st assay for HPG-related nucleic acid sequences (DNA or RNA) were prioritized for analysis to detect the presence of HPG-specific nucleic acids (DNA or RNA). 25 of which tested positive and were then analyzed in the 2nd assay.

<sup>2</sup>Samples are inferred as RNA positive through a failure to generate amplicons in the absence of reverse transcriptase.

HPG, Hervey pteropid gammaretrovirus

## Supplementary References

1. Van den Hurk AF, Smith IL, & Smith GA (2007) Development and evaluation of real-time polymerase chain reaction assays to identify mosquito (diptera: Culicidae) bloodmeals originating from native australian mammals. *J Med Entomol* 44(1):85-92.
2. Kohl C, *et al.* (2015) Protocol for metagenomic virus detection in clinical specimens. *Emerging Infect Dis* 21(1):48.
3. Kohl C, *et al.* (2018) Hervey virus: Study on co-circulation with henipaviruses in pteropid bats within their distribution range from australia to africa. *PLoS One* 13(2):e0191933.
4. Marsh GA, *et al.* (2012) Cedar virus: A novel henipavirus isolated from australian bats. *PLoS Path* 8(8):e1002836.
5. Bankevich A, *et al.* (2012) Spades: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19(5):455-477.
6. Nurk S, *et al.* (2013) Assembling genomes and mini-metagenomes from highly chimeric reads. *Annual International Conference on Research in Computational Molecular Biology*, (Springer), pp 158 170.
7. Crameri G, *et al.* (2009) Establishment, immortalisation and characterisation of pteropid bat cell lines. *PLoS One* 4(12):e8266.
8. Katoh K & Standley DM (2013) Mafft multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30(4):772-780.
9. Edgar RC (2004) Muscle: Multiple sequence alignment with high accuracy and high throughput. *Nuc Acid Res* 32(5):1792-1797.
10. Talavera G & Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* 56(4):564-577.
11. Guindon S & Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52(5):696-704.
12. Wapling J, Moore KL, Sonza S, Mak J, & Tachedjian G (2005) Mutations that abrogate human immunodeficiency virus type 1 reverse transcriptase dimerization affect maturation of the reverse transcriptase heterodimer. *J Virol* 79(16):10247-10257.
13. Machinaga A, Ishihara S, Shirai A, & Takase-Yoden S (2016) Splicing of friend murine leukemia virus env-mrna enhances its ability to form polysomes. *Front Microbiol* 7:160.

## SI Appendix

14. Drummer HE, Boo I, Maerz AL, & Pountourios P (2006) A conserved gly436 trp-leu ala-gly-leu-phe-tyr motif in hepatitis c virus glycoprotein e2 is a determinant of cd81 binding and viral entry. *J Virol* 80(16):7844-7853.
15. Bossart KN, *et al.* (2007) Neutralization assays for differential henipavirus serology using bio-plex protein array systems. *J Virol Methods* 142(1-2):29-40.
16. Sanders RW, *et al.* (2002) Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. *J Virol* 76(17):8875-8889.
17. Hayman DT, *et al.* (2008) Evidence of henipavirus infection in west african fruit bats. *PLoS One* 3(7):e2739.
18. Hayman DT, *et al.* (2011) Antibodies to henipavirus or henipa-like viruses in domestic pigs in ghana, west africa. *PLoS One* 6(9):e25256.
19. Plowright RK, *et al.* (2008) Reproduction and nutritional stress are risk factors for hendra virus infection in little red flying foxes (*Pteropus scapulatus*). *Proceedings of the Royal Society B: Biological Sciences* 275(1636):861-869.
20. Breed AC, *et al.* (2010) Prevalence of henipavirus and rubulavirus antibodies in pteropid bats, papua new guinea. *Emerging Infect Dis* 16(12):1997.
21. Farrell KB, Russ JL, Murthy RK, & Eiden MV (2002) Reassessing the role of region a in pit1-mediated viral entry. *J Virol* 76(15):7683-7693.

s22

s22

**From:** Edward Holme  
**Sent:** Thursday, 5 September 2019 1:19 PM  
**To:** Joshua Hayward  
**Cc:** mary.tachedjian; Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; Solymosi, Philip; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos; garycrameri  
**Subject:** Heidi Drummer; Andy Poubourios; Glenn.Marsh; Michelle Baker; Wang Linfa; <Ina.Smith>; Gilda Tachedjian  
Re: HPG Paper Fwd: PNAS MS#2019-15400 Submitted

Great work Josh!

Eddie

Professor Edward C. Holmes FAA FRS  
The University of Sydney

Duplicate Email - Removed

**Robinson, Karen (Governance, Black Mountain)**

---

**From:** Gary Cramer; [REDACTED] s22  
**Sent:** Thursday, 5 September 2019 2:41 PM  
**To:** Joshua Hayward  
**Cc:** mary.tachedjian; Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; Solymosi, Philip; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos [REDACTED] s22 Heidi Drummer; Andy Poubourios; Glenn.Marsh; Michelle Baker; Edward Holmes; Wang Linfa; <Ina.Smith [REDACTED] s22 Gilda Tachedjian  
**Subject:** Re: HPG Paper Fwd: PNAS MS#2019-15400 Submitted

Terrific Josh thanks for all your hard work  
Cheers  
Gary

Duplicate Email - Removed



s22

**From:** Heidi Drummer [s22]  
**Sent:** Thursday, 5 September 2019 3:05 PM  
**To:** Joshua Hayward  
**Cc:** mary.tachedjian; Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; Solymosi, Philip; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos [s22]; Gary Cramer; Andy Poubourios; Glenn.Marsh; Michelle Baker; Edward Holmes; Wang Linfa; <Ina.Smith [s22]>; Gilda Tachedjian  
**Subject:** Re: HPG Paper Fwd: PNAS MS#2019 15400 Submitted

Thanks to everyone for their contributions  
Good Luck  
H

On Thu, 5 Sep 2019 at 15:00, Joshua Hayward [s22] wrote:  
Hi all,  
Just a heads-up that I will also be submitting an abstract regarding this work for the AVS10 meeting in December.  
Cheers,  
Josh

**Joshua Hayward PhD**

**Research Officer  
Retroviral Biology and Antivirals Research Laboratory**

**Burnet Institute  
Disease Elimination and Maternal & Child Health Programs**

[s22]

GPO Box 2284, Melbourne, Victoria, Australia 3001

[s22]

The Macfarlane Burnet Institute for Medical Research and Public Health Ltd ABN 49 007 349 984

Duplicate Email - Removed

s22

**From:** Smith, Ina (H&B, Black Mountain) [REDACTED] s22  
**Sent:** Thursday, 5 September 2019 3:22 PM  
**To:** 'Joshua Hayward'; Tachedjian, Mary (H&B, Geelong AAHL); Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; Solymosi, Philip; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos [REDACTED] s22; garycrameri [REDACTED] s22 Heidi Drummer; Andy Poubourios; Marsh, Glenn (H&B, Geelong AAHL); Baker, Michelle (H&B, Geelong AAHL); Edward Holmes; Wang Linfa; Gilda Tachedjian  
**Subject:** RE: HPG Paper Fwd: PNAS MS#2019-15400 Submitted

Well done Josh

Thanks everyone.

Cheers  
:na

Ina Smith, PhD  
Senior Research Scientist | Risk Evaluation and Preparedness Program | Health and Biosecurity | CSIRO, Clunies Ross St, Black Mountain ACT 2601 Ph. 02 6218 3579

Duplicate Email - Removed

s22

s22

**From:** Gilda Tachedjian  
**Sent:** Thursday, 10 October 2019 7:53 AM  
**To:** Edward Holmes  
**Cc:** Joshua Hayward; Muzza Work; Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; philip daniel.solymosi; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos; Gary Cramer; Heidi Drummer; Andy Poubourios; Glenn.Marsh; Michelle Baker; Wang Linfa; Black Mountain Smith Ina  
**Subject:** Re: PNAS MS# 2019 15400 Decision Notification

Terrific thanks Eddie!

On 10 Oct 2019, at 7:12 am, Edward Holmes wrote:

Perfect Josh! You send em, I'll run em.

Cheers,

Eddie

Professor Edward C. Holmes FAA FRS  
The University of Sydney

On 9 Oct 2019, at 10:04 pm, Joshua Hayward wrote:

Hi Gilda,  
I had done those using amino acid sequences however. For consistency we may consider generating nucleotide gag, pol and env trees with the same methodology and software as the full genome tree. Would that be ok, Eddie? I'm happy to split the processed alignment into the relevant genes.

Cheers,  
Josh

**Joshua Hayward PhD**

**Research Officer  
Retroviral Biology and Antivirals Research Laboratory**

**Burnet Institute  
Disease Elimination and Maternal & Child Health Programs**

s22

GPO Box 2284, Melbourne, Victoria, Australia 3001

s22

The Macfarlane Burnet Institute for Medical Research and Public Health Ltd ABN 49 007 349 984

On Wed, 9 Oct 2019 at 21:52, Gilda Tachedjian

s22

wrote:

Thanks Eddie,

Josh has done the ENV and POL the former looks like the data in the paper, the latter lacks resolution (low boot strap values). Are you happy to do the recombination analysis?

Good to know the evolutionary time scale is irrelevant!

Cheers

Gilda

Sent from my iPad

On 9 Oct 2019, at 8:48 pm, Edward Holmes

s22

wrote:

Sounds good Gilda.

1. The time issue. Irrelevant. The evolutionary time scale depicted in this phylogeny is clearly far older than the time span of sampling. Accordingly, sampling viruses a decade apart will no impact. This can just be argued with a statement such as this.

2. Individual tree genes. Easy and we sort of have most of these anyway. We could add gag. Easy to upload as a Supplementary Figure. If there is any phylogenetic movement this will be due to a lack of signal/resolution. Can't believe there will be much/any evidence of recombination. Easy to deal with though.

Cheers,

Eddie

---

**PROFESSOR EDWARD C. HOLMES FAA FRS**

ARC Australian Laureate Fellow

**THE UNIVERSITY OF SYDNEY**

Marie Bashir Institute for Infectious Diseases & Biosecurity,

Charles Perkins Centre,

School of Life & Environmental Sciences and Sydney Medical School,

The University of Sydney | Sydney | NSW | 2006 | Australia

s22

On 9 Oct 2019, at 5:58 pm, Gilda Tachedjian

s22

wrote:

Dear All,

Please see attached reviews on our manuscript which overall are favourable!!

Reviewer #2 has requested additional experiments - which we are already doing in anticipation (i.e Pit1 receptor).

Eddie, appreciate your input on the queries re: phylogeny i.e. How does time impact relatedness, request for RT and Env trees I.

Josh in response to the query re: HPG replication in cell culture - didnt you take soup from infected cells and use them to reinfect fresh cells?

Ina/Heidi/Andy - for the query re: HPG seropositivity we could modify what we are claiming or express MLV envelope and repeat the serology?

The revision is due 7th of December, although we can request for more time.

Cheers

Gilda

**Professor Gilda Tachedjian BSc (Hons) PhD**  
Head, Life Sciences Discipline  
Group Head, Retroviral Biology and Antivirals Lab  
NHMRC Senior Research Fellow

**Burnet Institute**

s22

The Macfarlane Burnet Institute for Medical Research and Public Health Ltd,  
85 Commercial Road, Melbourne, VIC 3004, Australia  
[burnet.edu.au](http://burnet.edu.au)

*Equity through better health*

Duplicate Email - Removed

s22

**From:** Edward Holmes [REDACTED] s22  
**Sent:** Tuesday, 3 December 2019 7:13 AM  
**To:** Gilda Tachedjian  
**Cc:** Joshua Hayward; Muzza Work; Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; Solymosi, Philip; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.konto [REDACTED] s22; Gary Cramer; Heidi Drummer; Andy Poubourios; Glenn.Marsh; Michelle Baker; Wang Linfa; <Ina.Smith [REDACTED] s22  
**Subject:** Re: PNAS MS# 2019 15400 Decision Notification

Many thanks for handling all this Gilda.

Cheers,

Eddie

---

**PROFESSOR EDWARD C. HOLMES FAA FRS**  
 ARC Australian Laureate Fellow

**THE UNIVERSITY OF SYDNEY**  
 Marie Bashir Institute for Infectious Diseases & Biosecurity,  
 Charles Perkins Centre,  
 School of Life & Environmental Sciences and Sydney Medical School,  
 The University of Sydney | Sydney | NSW | 2006 | Australia

s22

On 3 Dec 2019, at 5:47 am, Gilda Tachedjian [REDACTED] s22 wrote:

Dear all,

A short note to let you know that we have made excellent progress toward completing additional experiments requested by the reviewers however I have requested and have been granted an extension to submit our revised manuscript early next year. See email correspondence below for details.

Best regards

Gilda

**Professor Gilda Tachedjian BSc (Hons) PhD**  
 Head, Life Sciences Discipline  
 Group Head, Retroviral Biology and Antivirals Lab  
 NHMRC Senior Research Fellow

**Burnet Institute**

s22

The Macfarlane Burnet Institute for Medical Research and Public Health Ltd,  
 [REDACTED] s22  
[burnet.edu.au](http://burnet.edu.au)

Begin forwarded message:

**From:** "Myers, Zoe" [REDACTED] s22  
**Subject:** RE: PNAS MS# 2019-15400 Decision Notification  
**Date:** 3 December 2019 at 5:00:23 am NZDT  
**To:** "gilda.tachedjian" [REDACTED] s22

Dear Dr. Tachedjian,

Thank you for your email. It will not be a problem to grant your extension until February 8, 2020. We understand it is a busy time of year and we look forward to receiving your revised manuscript.

Best wishes,

Zoe Myers

PNAS Editorial Office  
[REDACTED] s22

---

**From:** PNAS [REDACTED] s22  
**Sent:** Monday, December 2, 2019 10:56 AM  
**To:** Myers, Zoe [REDACTED] s22  
**Cc:** PNAS [REDACTED] s22  
**Subject:** FW: PNAS MS# 2019 15400 Decision Notification

-JN

---

**From:** Gilda Tachedjian [REDACTED] s22  
**Sent:** Friday, November 29, 2019 4:54 PM  
**To:** [REDACTED] s22 [REDACTED] pnascentral.org; PNAS [REDACTED] s22  
**Cc:** Gilda Tachedjian [REDACTED] s22 [REDACTED] Joshua Hayward [REDACTED] s22 [REDACTED] Smith, Ina (H&B, Black Mountain) [REDACTED] s22 [REDACTED] Mary Tachedjian [REDACTED] s22 [REDACTED] Adam Johnson [REDACTED] s22  
**Subject:** Re: PNAS MS# 2019-15400 Decision Notification

Dear PNAS Editor-in-Chief,

I would like to request a 9 week extension from the current due date of our revised manuscript (requested new due date 8th Feb 2020)

We have been asked to undertake additional experiments by the reviewers which we agree will improve the manuscript.

While we are making excellent progress towards completion of these studies, we will need more time due to circumstances out of our control.

These include key staff being on leave including having to serve on a jury for almost a month and the upcoming holiday season where our institute is closed

over the break and staff are expected to take their annual leave up to mid January.

Thank you for considering our request.

We look forward to hearing from you soon.

Best Regards

Gilda

**Professor Gilda Tachedjian BSc (Hons) PhD**  
Head, Life Sciences Discipline  
Group Head, Retroviral Biology and Antivirals Lab  
NHMRC Senior Research  
Fellow

**Burnet Institute**

s22

The Macfarlane Burnet Institute for Medical Research and Public Health Ltd,

s22

[burnet.edu.au](http://burnet.edu.au)

*Equity through better health*

On 9 Oct 2019, at 6:56  
am, s22@pnascentral.org wrote:

October 8, 2019

**Title: "Infectious KoRV-related retroviruses circulating in Australian bats "**  
Tracking #: 2019-15400  
Authors: Hayward et al.

Dear Dr. Tachedjian,

The expert who is serving as editor for your manuscript [MS# 2019-15400] has obtained 2 reviews, which are included below. The editor requests that you constructively address the concerns of the reviewers in a revised manuscript. Please note that multiple revisions are rarely permitted and there is no guarantee that the paper will be accepted.

PNAS allows 60 days to submit a revision. Your revision is due by December 7, 2019. If you require additional time, please contact the PNAS office.

When submitting revised materials, we require that you include a cover letter with a point-by-point response to the reviewers' comments. If you submitted a single PDF at initial submission, you must submit individual publication-ready files (e.g., Word

file for manuscript text; EPS, TIFF, or high-resolution PDF for figures; Word file for tables; etc.)

Please note that statements such as "data not shown" and "personal communication" cannot be used to support claims in the work and should be removed prior to submission. Authors are encouraged to use supporting information to show all necessary data, or to deposit their data in a publicly accessible database if posting as supporting information is not possible. Authors should include a statement in their methods section describing how readers will be able to access the data, associated protocols, code, and materials.

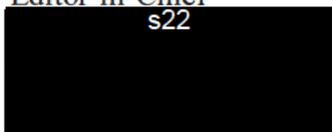
You may submit your revised manuscript here: [https://www.pnascentral.org/cgi-bin/main.plex?el\\_A2B7DVBw3A1BbFN2I7A9ftdMpCfRHgvFpmlE6PZCqVwQZ](https://www.pnascentral.org/cgi-bin/main.plex?el_A2B7DVBw3A1BbFN2I7A9ftdMpCfRHgvFpmlE6PZCqVwQZ).

\*\*\*Adding, removing, or reordering your author list requires approval from all coauthors before we can proceed. If you wish to add an additional corresponding author, please note this in the "Comments for Editorial Staff" box when completing your revision.\*\*\*

We recommend that authors submit ORCID IDs. If you provide your ORCID ID when you submit your manuscript, you can opt in to have your ORCID record automatically updated if your article is published. Watch for an email from Crossref in your ORCID inbox requesting permission to access your ORCID record.

Thank you for submitting to PNAS. We look forward to receiving your revision.

Sincerely yours,  
May R. Berenbaum  
Editor-in-Chief



\*\*\*\*\*

Editor Comments:

Our reviewers were overall positive about the paper, but had specific suggestions for improvement. Both made very good points. Some weakening of claims (about possible relation to endogenous viruses, for example) might be in order. Reviewer #2 had several requests for additional experiments, and many were not difficult. Adding as many as feasible would strengthen the paper.

We would look forward to a revised draft that addresses as many as possible of the reviewers' points.

Reviewer Comments:

Reviewer #1:

Sufficient General Interest?:

Yes

Conclusions Justified?:

Yes

Clearly Written?:

Yes

Procedures Described?:

Yes

Supplemental Material Warranted?:

Yes

Willingness to Re-review?:

Yes

Comments:

Here Hayward et al. identify and characterize the a full-length, replication-competent gammaretrovirus genome isolated from a bat in Australia, called HPG. In addition, four other related, partial retroviral genomes were isolated from other bat species in Australia and China. These retroviruses are closely related to koala retrovirus (KoRV) and gibbon ape leukemia virus (GALV) with similar sequence, tropism, and structure. HPG envelope antibodies were detected in multiple bat serum samples and HPG-like nucleic acids were detected in multiple bat feces samples. The authors propose that this novel gammaretrovirus is actively replicating in bats and that bats having overlapping habitats with koalas and gibbons may have led to interspecies transmission. The work is solid and novel, but a few key questions were not fully addressed in the study that could strengthen the manuscript.

Major Comments:

1. It appears that KoRV-related retroviruses were only detected in bat mucosal excretions and not in blood. Is it known where these viruses may replicate in koalas (or bats based on receptor expression) and if this tropism would explain this finding? Related to this, how do the authors envision that interspecies transmission occurred? This would be particularly useful to include in the Discussion, as it is not clear when and how transmission(s) between the 2 species occurred.
2. On page 9, line 15 and 17, the authors state, " contain endogenous HPG-related sequences" and "...suggesting evidence of endogenization or latent infection with HPG-related viruses." However, on page 7, they report that HPG is

unlikely to be an endogenous virus. First, these statements are contradictory and should be reconciled. Second, in the BLAST analysis, presumably bat ERVs were identified. What sequence identity do they have with HPG? The authors state that "sequences with high percent nucleotide similarity" were not observed. However, "high" is not defined.

3. Regarding phylogenetic analysis with related gammaretroviruses, how much does time impact relatedness? For example, HPG was isolated from a bat obtained in 2011, but it is unclear when the other viruses were obtained and how this could affect evolution, particularly in new host species (i.e. species adaptation after transmission from bats to koalas).

4. On page 9, the authors describe 12 samples that were positive only for HPG DNA and not HPG RNA and state that they represent animals that are "latently infected with other HPG-related virus(es) or contain endogenous HPG-related sequences." An alternative explanation would be low quantity and/or low quality RNA present in these samples. It is unclear that RNA and DNA quantity or quality were controlled in the analysis.

#### Minor Comments:

1. The Introduction (page 5, line 1) and the Discussion (page 10, line 3) state "the east coast of Australia including the Daintree rainforest," which implies that the Daintree rainforest is part of the Australian east coast. However, the Results section (page 6, lines 6-7) state "373 bats along the east coast of Australia and 106 bats from the Daintree Rainforest (Queensland)" that seems to imply that the Daintree rainforest is separate from the east coast. The text should be consistent throughout the manuscript.

2. Supplementary Table 2: it is assumed that Genbank accession numbers for the bat retroviruses will be forthcoming and included in the final version of the manuscript.

3. It is recommended that data presented in Supplementary Figure 7 be described in the Results section after Figure 5 (page 8, 2nd paragraph).

4. Supplementary Figure 9 is not discussed anywhere in the text. It is suggested that it be removed or appropriately described in the text.

#### Reviewer #2:

Sufficient General Interest?:

Yes

Conclusions Justified?:

No

Clearly Written?:

Yes

Procedures Described?:

Yes

Supplemental Material Warranted?:

Yes

Willingness to Re-review?:

Yes

Comments:

This is an interesting paper submitted by Hayward and colleagues, describing the discovery and biological characterization of KoRV-related gammaretrovirus sequences in samples from different Australian bat species. The bulk of the results focuses on sequences obtained from scat of *P. Alecto*, and referred to here as Hervey pteropid gammaretrovirus (HPG). A consensus is used to reconstruct a full HPG viral genome, and a variety of biochemical methods and EM is used to confirm production of gamma-like virions. If true, this may be the first description of an exogenous gammaretrovirus of bats. While the study is likely to be of broad interest, there are several caveats to interpretation that should be addressed, as well as some minor points.

Page 7, results first paragraph - An important caveat is that this could also be a recent, rare and unfixed ERV insertion, similar to many KoRV loci in Koalas, especially in southern koalas - that is to say, a similar approach in southern koalas might "miss" detecting a rare enKoRV sequence. The intact nature of the HPG is also consistent with something that could be present in both exogenous and endogenous forms. Effectively, this doesn't change the impact of the manuscript - either its an exogenous gammaretrovirus of bats, a very recently endogenized gammaretrovirus of bats, or both. My suggestion is to stay open to all possibilities - present it as an exogenous virus, but acknowledge that the actual samples might have detected a germline insertion (ERV).

Page 7, results - phylogenetic analysis. While a tree based on the full genomes is potentially robust, the authors should also analyze RT and env separately. Do they give the same/similar results as one another, and are they consistent with the tree based on the entire genomes? Recombination can obscure phylogenetic relationships, especially when one part of the genome is more divergent or has had a very different evolutionary trajectory. For example, it could be one gene, such as env, that separates one branch from the others, but is the result of a single recombination event and not of divergence over time. Gene-specific phylogenies could be added to supplemental data, and wouldn't be necessary in the main text (unless they reveal a more complex phylogenetic history, in which case the authors will want to make it part of the story).

Page 8, serological analysis - the negative control (HIV Env) rules out general background, but does not rule out cross-reactivity with other gamma-type retroviruses or ERV expression. How specific is this assay? Since the claim is "HPG seropositivity", it should include Env proteins from a distant relative (GaLV, KoRV) and even a different gamma lineage altogether (e.g., MLV Env). The conclusion could then be "HPG-seropositivity" or "KoRV-related retrovirus seropositivity" depending - either result fits the story being described in the manuscript. But as is, it's not clear they can claim specificity for HPG.

page 17, figure 2 - Need to explain the ratios at the nodes (e.g., 1/100, 1/97, etc). Are these bootstrap values, and if so, why are they presented this way (usually they are given as percentages)? This info should be in the figure legend.

Page 18, Figure 3 and related results section - The negative control (which is shown in Supp Fig 8) to confirm that the particles are produced by the transfected plasmid is not mentioned here or referred to in the main text. Authors also need to include some indication of how many images or fields were needed to detect the particles or, more specifically, to explain with what certainty the negative controls can be said to have less/no particles compared to the composites in figure 3. For example, were sufficient fields analyzed or similar numbers of cells visualized/ is this the result of comparing similar numbers of images/cells for transfected and control cells?

Page 20, figure 4: the HPG result is distinctive - the MLV infection results in the expected plateau consistent with ongoing replication, whereas HPG replication peaks and drops quickly to background. This raises the possibility that the data don't represent ongoing replication cycles, but rather a burst of production from initially infected cells. Another possibility is that the HPG retrovirus is replicating, but is toxic to cells, similar to lentivirus replication in cell culture. An experiment to examine these possibilities and to definitively establish successive rounds of replication is important (e.g., passaging filtered supe to a second plate/flask followed by RT assay, or replication with and without inhibiting RT, etc).

General - The manuscript proposes that HPG uses the same receptor as KoRV-A and GaLV, and even includes a supplemental figure depicting the conserved binding site motif in PiT-1 of the relevant host species. This is presented as part of the argument in referring to this as a "KoRV-related retrovirus", as in the title of the manuscript and elsewhere in the text. Given how easy it is to do, why not formally prove this? It should be straightforward, and there is plenty of precedent in the literature - either by adding PiT-1 expression to null cells (such as the NIH3T3 cells used in figure 5), or by

means of a standard superinfection cross-interference assay. Either experiment can be done with existing reagents in a relatively short period, and would strengthen the manuscript.

Page 32, line 21 supplemental methods refers to "Supp Figure 8" but probably is supposed to refer to Supp Figure 9.

Page 43, Supp figure 9 - legend could use some additional info - Are there control lanes (non-transfected or mock transfected) in the image? If so, are the controls the basis for establishing that the indicated bands are HPG Env? The lanes should be labeled or mentioned in the legend. Alternatively, If there are no control lanes, how can the authors claim that this isn't an unfortunate background band?

\*\*\*\*\*

#### PUBLICATION INFORMATION

High resolution figure files are required for the final version of your manuscript. PNAS figure preparation guidelines state that no specific feature within an image may be enhanced, obscured, moved, removed, or introduced. The grouping or consolidation of images from multiple sources must be made explicit by the arrangement of the figure and in the figure legend. Adjustments of brightness, contrast, or color balance are acceptable if they are applied to the whole image and if they do not obscure, eliminate, or misrepresent any information present in the original, including backgrounds. Please note that our production editors may flag figures that are not in compliance with our figure policy, resulting in delays. For more information on submitting high resolution figures please review the

s22

s22

**From:** Wang Linfa  
**Sent:** Tuesday, 3 December 2019 11:22 AM  
**To:** Gilda Tachedjian  
**Cc:** Joshua Hayward; Muzza Work; Kohl, Claudia; Adam Johnson; Dearnley, Megan (AAHL, Geelong AAHL); Brianna Jesaveluk; Christine Langer; Solymosi, Philip; Hille, Georg; Andreas; Cecilia Sanchez; Adam Werner; dimitri.kontos s22 Gary Crameri; Heidi Drummer; Andy Poubourios; Glenn.Marsh; Michelle Baker; Edward Holmes; <Ina.Smith s22  
**Subject:** Re: PNAS MS# 2019 15400 Decision Notification

Dear Gilda,

Thanks for the update and good luck!

Linfa

Sent from my iPhone

Duplicate Email - Removed