s22

From: Joshua Hayward s22
Sent: Friday, 31 January 2020 9:21 PM
To: 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 Crameri; Heidi Drummer; Andy Poumbourios; Glenn.Marsh; Michelle Baker; Edward Holmes; Wang Linfa; s22 Gilda Tachedjian
Subject: HPG Paper - PNAS Submission - Response to reviewers and revised manuscript
Attachments: Attachments not available to print

Dear all,

Please find attached our draft of the revised manuscript & SI, and our draft response to viewers. I'd like to invite you all to view the changes and our response and welcome any feedback/suggested edits you may wish to provide.

Apologies that the timeline for the final edits is a little tight. As we have a deadline of Feburary 8, I will need to get any feedback by CoB on Wednesday 5th of Feb (next week). If I haven't heard from you by then, I will assume you are happy with the manuscript and response to reviewers as it is.

Changes relative to our original submission are highlighted in yellow.

The major changes:
i) We have added a 'Secondary infection assay' to demonstrate that HPG can establish successive rounds of infection.
ii) We have added phylogenetic analyses of individual genes
iii) We have added a 'Superinfection interference assay', which has allowed us to confirm the receptor usage (turns out HPG uses both PiT-1 and PiT-2).
iv) We have superseded the luminex assay with a VRA peptide and protein binding assay, which has allowed us to reveal more details regarding the specificities of the positive sera, showing that in the majority of cases while cross reactivity with KoRV and GALV is possible, reactivities are strongest against HPG.

do need a couple of things from some authors:

Heidi: Can you please check my interpretation of the VRA data
Heidi/Andy: Can you please provide a line describing the anti-MLV sera (source etc.)
Megan: Can you please look over the response to the question about the EM
Eddie: Can you please check my descriptions of the new phylogenetics
Adam: As discussed, just need that PERT data and the nucleic acid concentration stuff.

Have a great weekend!

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

[redacted]

From: Edward Holme [redacted]
Sent: Saturday, 1 February 2020 11:21 AM
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] Gary Crameri; Heidi [redacted] bourios; Glenn.Marsh; Michelle Baker; Wang Linfa; [redacted] Gilda Tachedjian
Subject: Re: HPG Paper PNAS Submission Response to reviewers and revised manuscript
Attachments: Response to Reviewers v1.2.EH.docx; HPG paper SI Appendix Revised v2.0.EH.docx; HPG paper Manuscript Revised v2.1.EH.docx
Attachments Removed

Hi Josh,

Here you go just some very minor edits from me. Fabulous stuff.

All the best,

Eddie

PROFESSOR EDWARD C. HOLMES FAA FRS
ARC Australian Laureate Fellow

THE UNIVERSITY OF SYDNEY
Marie Bashir Institute for Infectious Diseases & Biosecurity,
School of Life & Environmental Sciences and School of Medical Sciences,
The University of Sydney | Sydney | NSW | 2006 | Australia
[redacted]

On 31 Jan 2020, at 9:24 pm, Joshua Hayward [redacted] wrote:

As is tradition, I immediately notice an error please replace the manuscript draft with the one attached now!

Joshua Hayward PhD

Research Officer
Retroviral Biology and Antivirals Research Laboratory

Burnet Institute
Disease Elimination and Maternal & Child Health Programs

[redacted]
GPO Box 2284, Melbourne, Victoria, Australia 3001
[redacted]

Duplicate Email - Removed

<HPG paper - Manuscript Revised v2.1.docx>

7th February 2020

Dear Editor,

Thank you for considering our manuscript entitled "**Infectious KoRV-related retroviruses circulating in Australian bats**" by Joshua A. Hayward, Mary Tachedjian, Claudia Kohl, Adam Johnson, Megan Dearnley, Brianna Jesaveluk, Christine Langer, Philip D. Solymosi, Georg Hille, Andreas Nitsche, Cecilia A. Sánchez, Adam Werner, Dimitri Kontos, Gary Crameri, Heidi E. Drummer, Pantelis Poumbourios, Glenn A. Marsh, Michelle L. Baker, Edward C. Holmes, Lin Fa Wang, Ina Smith, Gilda Tachedjian for publication in ***Proceedings of the National Academy of Sciences***.

Please find our responses to the reviewer's questions below. Revisions within the manuscript are highlighted in yellow.

Thank you for your consideration

Yours Sincerely,

Prof. Gilda Tachedjian
NHMRC, Senior Research Fellow
Head, Life Sciences Discipline
Head, Retroviral Biology and Antivirals Laboratory
Burnet Institute
s22
[redacted]

Editor's 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.

Reviewer #1:

General 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:

Comment 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.*

Response: Although appropriate blood samples were not available for analysis, we were able to detect KoRV related sequences in feces and urine. While KoRV has been detected in the blood of infected koalas (Tarlinton, 2006, Nature, 442:7098; Simmons, 2012, Aus. Vet. J. 90:10; Waugh, 2017, Sci. Rep. 7:1), KoRV nucleic acids and/or proteins have also been identified in various tissues, including sperm (Tarlinton, 2006, Nature, 442:7098), breast milk (Xu, 2013, PNAS, 110:28; Morris, 2016, Sci. Rep. 6) as well as feces (Wedrowicz, 2016,

Conserv. Genet. Resour. 8:4). The closely related Gibbon ape leukemia virus (GALV) has additionally been identified in the feces and urine of gibbons (Kawakami, 1977, J. Natl. Cancer. Inst. 268:5619).

KoRV A and GALV utilize the PiT 1 (SLC20A1) receptor (reviewed in Denner, 2016, Viruses, 8:12), which is a phosphate transport protein ubiquitously expressed at variable levels throughout the mammalian body (Kavanaugh, 1994, PNAS, 91:15; Johann, 1992, J. Virol. 66(3); also see expression database entries:

Expression Atlas (Petryszak, 2015, *Nucleic acids research,* D746 D752)
https://www.ebi.ac.uk/gxa/genes/ensg00000144136?bs=%7B%22homo%20sapiens%22%3A%5B%22ORGANISM_PART%22%5D%7D#baseline

Bgee (Bastian, 2008, In *International Workshop on Data Integration in the Life Sciences*, Springer, Berlin, Heidelberg)
https://bgee.org/?page=gene&gene_id=ENSG00000144136)

The highest expression levels of PiT1 are in locations including the colon, testes, breast, bladder, placenta, and brain. Our experimental results (Fig 5A), including new data from infection interference assays (Fig 5B) suggests that HPG utilizes the same cell receptor, PiT 1, similar to KoRV A and GALV. Taken together, these observations are consistent with detecting KoRV related retroviruses (i.e. HPG) in the feces of bats as well as other pooled tissue samples.

Regarding potential routes for interspecies transmission we would like to clarify that we do not propose that a specific species to species transmission from bats to Koalas/Gibbon apes occurred for the KoRV related retroviruses identified in our study. Rather, our phylogenetic analysis suggests that there are likely retroviruses **more closely** related to KoRV/GaLV yet to be discovered. This was stated in the discussion section [redacted] as follows:

"Hence, bat communities could in theory provide a route of transmission for KoRV related viruses between Asia and Australia, although the immediate ancestor of KoRV remains uncertain and it is clear that additional animal species need to be sampled. Indeed, there are likely to be other currently unidentified species infected with KoRV-related viruses linking the habitats of R. hipposideros and Australian bats. The long phylogenetic branch length linking the KoRV clade to its closest known relatives in the GALV/WMV clade indicates that the

phylogenetic picture remains incomplete, with additional as yet unknown viruses and host species existing between the KoRV and GALV/WMV lineages of gammaretroviruses."

Further regarding interspecies transmission, in general, given the diversity of body fluids within which KoRV and GALV (and by extension, KoRV related viruses) might be found, a number of possible scenarios may be reasonably speculated. These include transmission via blood during fighting/predation, and contamination of food sources by feces and urine.

To address the comments raised by the reviewer, we have included the following paragraph in the Discussion at [ ]:

"KoRV and GALV utilize the PiT-1 receptor for cell entry (19, 27, 28). This receptor is almost ubiquitously expressed throughout the mammalian body at variable levels (40 43), and is highly expressed in many tissues including the colon, breast, testes, bladder, placenta, and brain (40, 41). KoRV and GALV have been detected in numerous tissues and body fluids including blood, sperm, breast milk, feces, and urine (5, 27, 44 49). Given the wide distribution of PiT-1 expression and the detection of KoRV and GALV in body fluids including blood, urine, and feces, it is possible that interspecies transmission might occur along routes including blood during fighting/predation, and contamination of food sources by feces and urine."

Comment 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.*

Response: We thank the reviewer for raising this point and can appreciate why our statements appear contradictory. We agree that we cannot rule out the possibility that HPG is not endogenous in some proportion of the bat gene pool since we have only sampled a small number of bats. Accordingly, we have modified the manuscript as follows to soften our claims:

We have modified our result heading from

"HPG is not an endogenous retrovirus"

to

"HPG sequences were not detected in the genomes of pteropid bats" on [page, line].

Furthermore, within the discussion section, we have modified the text to read as follows on [page, line]:

"We searched carefully for the presence of HPG in the genomes of P. alecto and P. vampyrus using molecular analyses, and more broadly for KoRV related viruses in the SRA: we were unable to detect these viral sequences in any currently available bat genome sequence. While these data suggest that bat KoRV related viruses are not endogenous, we cannot fully exclude the possibility as we have only sampled a small proportion of bats within each species. In this regard, KoRV endogenization in koalas is relatively recent, and accordingly is not represented across the entire koala gene pool (44); existing in both endogenous and exogenous forms (5, 27, 50). Thus, given that HPG specific sequences have been identified across several bat species, either HPG is an exogenous virus or it is undergoing endogenization in real time. A possible example of the latter is FFRV1 (14), which was recently discovered in the brain tissue of a P. alecto bat, but which we were not able to identify within the genome of P. alecto or other bats. ~~*We searched carefully for the presence of HPG in the genomes of P. alecto and P. vampyrus using molecular analyses, and more broadly for KoRV related viruses in the SRA and were unable to detect these viral sequences in the genome of any bat species whose genome is currently available. While these data suggest that bat KoRV related viruses are not endogenous, we cannot rule out the possibility as we have only sampled a small proportion of bats within each species. In this regard, KoRV endogenization in koalas is relatively recent, and accordingly is not represented across the entire koala gene pool (44); existing in both endogenous and exogenous forms (5, 27, 50). Thus, given that HPG specific sequences have been identified across several bat species, either HPG is an exogenous virus or it is undergoing endogenization in real time. A possible example of the latter is FFRV1 (14), which was recently discovered in the brain tissue of a P. alecto bat, but which we were not able to identify within the genome of P. alecto or other bats."*~~

Formatted: Font: Italic

Formatted: Font: Italic

Regarding our BLAST search within pteropid genomes for HPG sequences, when we analysed the genome of *P. alecto* for sequences similar to HPG, our BLAST analysis revealed that HPG and closely related retroviral

sequences were not present. The closest identified hit against the HPG sequence in this analysis was a 546 nt sequence aligning to the *pol* gene of HPG, with a nucleotide identity of 69%.

To address this matter, we have replaced the sentence on [page, line]:

"No sequences with high percent nucleotide similarity to HPG were identified"
with
"No sequences matching HPG were identified. The closest identified hit against the HPG sequence in this analysis was a 546 nt sequence within the genome of P. alecto, aligning to the pol gene of HPG, with an e value of 5.0 x 10^{-46} and a nucleotide identity of 69%."

Comment 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).*

Response: We apologize that the sampling period was not clear. The sampling period for these viruses was only over seven years (as stated on pg x Lines Y) and accordingly is highly unlikely to have an impact on relatedness in the context of the evolutionary time scale depicted in our phylogeny which is almost certainly on the scale of thousands to millions of years (Holmes EC. (2009). The Evolution and Emergence of RNA Viruses. Oxford Series in Ecology and Evolution, Oxford University Press, Oxford.).

Comment 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.*

Response: To clarify, the samples referenced in this sentence tested positive for HPG *specific* RNA, but not HPG *related* RNA indicating that the quality of the RNA in these samples was adequate for this analysis. While we did not detect HPG related RNA in these samples, we did detect HPG related DNA.

[Add info about determination of concentrations] During the qRT PCR analysis, bat samples were classified as positive or negative based on their fluorescence signal compared a standard curve generated using $1x10^0$ – $1x10^7$ copies of the HPG proviral plasmid. In the standard curve, signal was only generated down to a threshold of $1x10^1$ copies, the signal for which appeared at cycle 36 (CT 36), and this CT value served as the cut off for determining a positive result.

Commented [JH1]: We will include summary info regarding nucleic acid concentrations for the 373 samples prior to resubmission.

While we cannot exclude that some amount of HPG related RNA was present in the samples and below the limits of detection, we did utilize a highly sensitive kit (Thermo Power SYBR Green RNA to CT Kit) that is capable of detecting specific targets from sub picogram levels of total RNA.

To address the reviewer's alternative explanation, we have modified this sentence to now read as [page, line]:

"While we cannot rule out that some or all of these samples may have contained HPG related RNA below the limit of detection of this assay, these data suggests that 12 bat samples were actively infected with HPG and were either latently infected with other HPG related virus(es) or contain endogenous HPG related sequences."

Within the supplementary methods [redacted] we have added the following details:

"Bat samples were classified as positive or negative based on their fluorescence signal compared against a standard curve generated using $1x10^0$ – $1x10^7$ copies of the HPG proviral plasmid. The cutoff for determining a positive result was a cycle threshold of 36, which correlated to $1x10^1$ copies of the HPG provirus."

Minor Comments:

Comment 5: *The Introduction (page 5, line 1) and the Discussion (page 10, line 3) state "the Daintree inforest," 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.*

Response: To clarify this statement we have modified the text at various locations to read as follows:

[location]: *"we collected bat samples (feces, blood, urine, and oral swabs) from towns and the Daintree rainforest along the east coast of Australia"*

[location]: *"To identify KoRV related viruses in bats, samples were collected from the east coast of Australia, including feces, oral swabs, blood, and urine. 373 samples were collected from towns in New South Wales and Queensland and 106 from the Daintree Rainforest (Queensland)."*

[location]: *"To determine whether KoRV-related viruses are present in Australian bats, we collected samples from bats on the east coast of Australia"*

Comment 6: *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.*

Response: The Genbank accession numbers for the bat retroviruses have been included in the updated version of Supplementary Table [redacted].

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

Response: We thank the reviewer for this suggestion and have included this description as follows in the results section:

"A complementary alignment of the binding motif within mammalian PiT 1 genes further supports this result as the binding sites within P. alecto and P. vampyrus PiT 1 share the permissive amino acid residues, which are distinct from the non permissive motif within mouse PiT 1 (Supplementary Figure [redacted])." on [page, line].

Comment 8: *Supplementary Figure 9 is not discussed anywhere in the text. It is suggested that it be removed or appropriately described in the text.*

Response: A mislabeled reference to this Supplementary Figure within the text of the Supplementary Methods section has been corrected, and now reads:

"SDS PAGE in the presence and absence of β-mercaptoethanol revealed a single diffuse band with a molecular weight range of ~80 90 kDa (Supplementary Figure [redacted]), consistent with the molecular weight predicted from the amino acid sequence (62,805 Da) with 6 N linked glycans (~ 18 kDa)." on [page, line].

Reviewer #2:

General 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.

Comment 1: *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).*

Response: We agree with the comment made by the reviewer, and apologize that this was not clear in the manuscript. We have addressed this comment in our response to Reviewer 1's Major Comment 2.

Comment 2: *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).*

Response: As suggested by the reviewer, we have now provided individual phylogenies for the *env*, *pol* and *gag* genes (Supplementary Figure [redacted]). As can be seen, the tree topologies for *env* and *pol* genes are the same as that for the complete viral genomes (Figure 2). A slightly different topology was observed in the *gag* gene phylogeny, however, as all the relevant bootstrap values were very low (35%, 41%, 48%), a history of genomic recombination cannot be safely inferred since the difference in tree topology in the *gag* gene lacks phylogenetic resolution.

We have included this additional analysis as Supplementary Figure [redacted], on [page, line].

We have included the corresponding text within the Results section:

"This analysis is supported by phylogenetic analyses of the individual pol and env genes, which reveal the same branching pattern. While analysis of the gag gene resulted in a slightly different branching pattern, this is likely as a result of low phylogenetic resolution~~,~~ as indicated by low bootstrap support for ~~this individual~~key nodes on this tree (Supplementary Figure [redacted])." on [page, line].

We have also updated the Supplementary methods section to include description of the phylogenetic analysis as follows on [page, line]:

"To determine the evolutionary relationships among KoRV-related gammaretroviruses, we performed phylogenetic analyses using aligned complete genome nucleotide sequences (Supplementary Table 2) and individual gene sequences. Accordingly, a multiple sequence alignment of 19 complete genomes was performed using a combination of MAFFT (8) and MUSCLE algorithms (9). Following alignment, regions of ambiguous and uncertain alignment were removed using Gblocks (10). For the complete genomes, this resulted in final alignment of 6,925 nt that was used to infer evolutionary relationships. Subsets of this alignment covering the gag, pol, and env gene regions were used for the individual gene analyses. Phylogenetic trees of these data were estimated using the maximum likelihood (ML) method available in the PhyML program (11), assuming a GTR model of nucleotide substitution with a proportion of invariant sites (I) and a gamma distribution of among site rate variation (Γ). To determine the robustness of each node a bootstrap resampling analysis (1,000 replications) was performed using the same nucleotide substitution model. For the complete genome tree (Figure 2), a

Shimodaira Hasegawa (SH) test was also conducted, providing additional nodal support. The Mus caroli ERV, McERV (Supplementary Table 2), sequence was used as an outgroup to root the tree."

Comment 3: *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.*

Response: The reviewer has raised an important issue with regards to HPG specific seropositivity across the tested bat samples. To address this issue we have undertaken a peptide binding analysis in a solid phase enzyme immunoassay to assess the seroreactivity of bat samples against short peptide sequences specific to HPG, KoRV, GALV, and the more distantly related MLV.

These new serology results have been included (Supplementary Figure) and supersede the luminex data within the Results section, which now reads as follows, on [page, line]:

"To assess Australian bats for exposure to HPG or KoRV-related viruses, we tested for the presence of bat antibodies reactive against the HPG Env protein. To determine whether bat samples that were seropositive for HPG might be cross reactive against the closely related gammaretroviruses, KoRV-A and GALV, or the more distantly related gammaretrovirus, MLV, we performed a peptide binding analysis using short peptide sequences derived from the Variable Region A within the Env protein of these viruses. We also tested for the presence of HPG specific nucleic acid in bat fecal samples.

Bat sera (87 samples) were screened for the presence of antibodies reactive to the HPG VRA region of Env in a solid phase enzyme immunoassay. Where available sample material permitted, analysis was additionally conducted on the full range of VRA peptides and the HPG Env trimer ectodomain (Glu^{38} Ser^{603}) (Supplementary Figure).

Of the 87 bat samples, 18 (20.7%) were reactive to the HPG VRA peptide [P. Alecto (n = 16), P. conspicillatus (n = 1), Rhinolopus megaphyllus (n = 1)]. All of these samples were also reactive to the HPG Env ectodomain. Nine (50%) of the HPG VRA positive samples were also cross reactive for KoRV A and GALV VRA peptides. However,

only a single sample (#20 P. alecto) was more strongly cross reactive, against the GALV VRA peptide. Two samples (#7 P. alecto and #8 P. alecto) were reactive against the KoRV A and GALV VRA peptide, respectively but not reactive against the HPG VRA peptide. Only a single bat (#27 P. alecto) demonstrated cross reactivity to MLV, and this bat was also reactive against the VRA of HPG, KoRV A, GALV, and HPG Env.

Hence, t~~T~~hese results reveal that 28% of bat samples were seropositive for HPG or KoRV related protein sequences. They also indicate that while some cross reactivity is observable, reactivity among the tested bat samples is almost entirely strongest against HPG, and cross reactivity generally does not extend to distant gammaretroviral relatives such as MLV."

The methodology in the supplementary methods section [redacted] *"Serological assay for the presence of anti PG antibodies in bats"* now reads:

"Bat sera were screened for the presence of antibodies reactive to the VRA region of Env using synthetic peptides in a solid phase enzyme immunoassay. N terminal biotinylated synthetic peptide encoding the HPG VRA region (LETWDIPDSDVSASTRVRPADSD, Genscript, USA) was added to Avidin coated plates (Nunc, Maxisorb) at 5 µg/ml followed by the addition of serially diluted bat serum in PBS containing 2.5 mg/ml bovine serum albumin and Tween 20 (0.05%). Bound antibodies were detected with horseradish peroxidase labelled Protein A/G (Thermo Scientific, Rockford) followed by 3,3',5,5' Tetramethylbenzidine (TMB) substrate (Sigma, USA). Sera that displayed above 10x background levels of binding were further screened for reactivity to biotinylated synthetic peptides of the equivalent regions of KoRV-A (LESWDIPELTASASQQARPPDSN), GALV (LESWDIPGTDVSSSKRVRPPDSD), and MLV (PSYWGLEYQSPFSSPPGPPCCS) in the same way."

Comment 4: *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.*

Response: We apologise that these values were not clearly explained. The values are not fractions, but the ~~combination~~ depiction of two different measures of nodal support : SH like branch support to the left and bootstrap support to the right

i.e. *'SH value/Bootstrap value'*

We have updated the figure legend to clarify this, and it now reads as follows:

"...All branches are scaled according to the number of nucleotide substitutions per site, and branches representing bat retroviruses are shown in red. Support for key nodes on the phylogeny are shown in the form SH like branch support/bootstrap support. Silhouettes represent the host species..."

Comment 5: Page 18, Figure 3 and related results secion 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?

Response: In thin section electron microscopy, the analysis is taking place on a single plane (70 90 nm) section of the cell, the height of which can be up to 20 µm in a cell monolayer and larger in suspension depending on the cell's orientation and morphology. As such, we feel that detection of viral particles measured by EM analysis alone is not the most efficient approach to measuring % infection and/or providing 100% confidence of finding (or not finding) viral particles in a test or control sample. For this reason, as described in the Supplementary Methods sections "Transfection of 293T cells for generation of HPG & M MLV viral particles" and "Electron Microscopy and viral particle morphology", we complimented our EM analysis with a virion associated reverse transcriptase activity (RT) assay on supernatant from each of the cell suspension samples that was to be analyzed by electron microscopy. A positive reading was obtained for MLV and HPG but not for the negative controls (cells with no transfection and the empty plasmid) indicating virus budding and activity in the test samples only. In sample blocks containing MLV 293T cells and HPG 293T cells, viral particles were readily observed budding from the cell membrane or in inclusion bodies within the cells, indicating a relative abundance of virus in the cells. For the negative controls, which were untransfected 293T cells and 293T cells mock transfected with the pcDNA3.1 plasmid, no virus was observed in cells following extensive examination across numerous fields of view, multiple sections and on two separate grids. Furthermore, cell morphology and ultrastructure of control samples was consistent with healthy cells in tissue culture. In contrast, cells in the population that had transfected with MLV and HPG showed morphological indictors of infection such as

fragmented cell and organelle membranes, extracellular debris (membrane) and in some instances, cytoplasmic or nuclear condensation. Together these data increase our confidence that control samples were truly negative for virus, whilst the MLV 293T cells and HPG 293T cells did contain replicating virus particles.

We have included the quality control data from the virion associated RT assay alongside the EM negative controls in Supplementary Figure [redacted]. The legend for this figure now reads:

"Supplementary Figure [redacted]. Electron micrographs (EM) of control untransfected cells and cells mock transfected with the empty vector pcDNA3.1, and virion associated reverse transcriptase (RT) assay. The graph displays the result of a virion associated RT assay, supporting the result that cells transfected with proviral M MLV and HPG expression plasmids generate retroviral particles, while untransfected and mock pcDNA3.1 transfected cells do not. A) Transmission electron micrograph..."

We appreciate the reviewer's suggestion that we make appropriate mention of the negative EM controls in the main text and have modified our reference to these data in the results section [redacted], which now reads:

"Transfection of human 293T cells with a plasmid construct carrying the HPG provirus resulted in the generation and release of viral particles morphologically similar to ecotropic Moloney murine leukemia virus (M MLV), as determined by electron microscopy (Figure 3), in contrast to untransfected and mock transfected 293T cells (Supplementary Figure [redacted]). These data are supported by virion associated reverse transcriptase analysis of the samples analyzed by electron microscopy (Supplementary Figure [redacted])."

The following sentence has been appended to the end of the legend of Figure 3 [redacted]:

"Negative transfection controls were untransfected cells and cells mock transfected with the empty vector pcDNA3.1. These controls were not observed to contain or produce viral particles (Supplementary Figure [redacted])."

Comment 6: 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).

Response: To address this possibility, and as suggested by the reviewer, we conducted a 'secondary infection assay', in which we established successive rounds of replication, and is included as Supplementary Figure [redacted]. In brief, 293T cells were transfected with the HPG proviral plasmid; cell culture supernatant was later harvested and clarified. This clarified supernatant was used to establish a primary infection in 293T cells in the same manner as our original infectivity assay. We then collected the clarified supernatant of these cells and repeated the process, and successfully established a secondary infection in 293T cells, as determined by a virion associated PERT assay. The data from this experiment confirms that successive rounds of replication can be established by HPG in 293T cells. In contrast to our original experiments we used a PERT assay (RT qPCR) to confirm the presence of HPG for these new experiments, as we have ceased using the radiolabeled virion associated reverse transcriptase assay due to increased and prohibitive costs of radiolabeled nucleotides in our region.

We have added this analysis to the Results section [redacted]:

"HPG establishes successive rounds of replication through a secondary infection assay (Supplementary Figure [redacted])."

The methodology for this assay has been added to the Supplementary Methods as "Secondary infection assay" [redacted], and reads:

"To confirm that HPG was capable of establishing successive rounds of infection, 239T cells were transfected with an infectious molecular clone of HPG as described in "Transfection of 293T cells for generation of HPG & M MLV viral particles". To establish a primary infection from HPG virions, 293T cells were infected as described in "Replication kinetics assay". To establish a secondary infection, clarified supernatant harvested 48 h following the primary infection was collected and used to establish a second round infection in 293T cells as described in "Replication kinetics assay", except that for the second round infection neat HPG supernatant was used; 5 µL

samples were collected at inoculation, 6 h following inoculation and washing (t = 0), and at 48 h following washing (t = 48). Collected samples were analyzed for the presence of virion associated reverse transcriptase activity by PERT assay, as described in "Generation of HeLa cells persistently infected with HPG". Values derived from the PERT assay represent arbitrary units of RT activity in comparison to a dilution series of HPG virions which were generated in house, as described in "Transfection of 293T cells for generation of HPG & M MLV viral particles"."

Other comments

Comment 7: 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.

Response: We thank the reviewer for this suggestion and have undertaken a superinfection interference assay, included as an additional panel (B) in Figure 5. Briefly, in this experiment we generated persistently HPG infected HeLa cells, then challenged these cells with infection by Envelope pseudotyped reporter retroviruses representing HPG, KoRV A, GALV, Amphotropic MLV, Dualtropic MLV, and the vesicular stomatitis virus (VSV). Compared against uninfected HeLa cells, HPG infected HeLa cells were strongly resistant to superinfection from HPG, KoRV A, and GALV Env pseudotyped viral particles; they were moderately resistant to infection by amphotropic and dualtropic MLV (which respectively use the PiT 2 and PiT 1 & PiT 2 cell receptors [Feldman, 2004, J. Virol. 78:2; Miller, 1996, J. Virol. 70:8]) Env pseudotyped particles; almost no impact was observed on susceptibility to infection by VSV. The ecotropic MLV used in our infection kinetics assay was not utilized in this assay as it is incapable of infecting human cells. These results suggest that HPG utilizes the PiT 1 and PiT 2 receptors for cell entry.

Commented [JH2]: We will add PERT data demonstrating maintenance of HPG infection over the course of passaging.

We have added the following description of this analysis to the Results section [redacted]:

"To further investigate receptor usage by HPG, we performed a superinfection interference assay (Figure 5B). In this assay, HeLa cells persistently infected with HPG became strongly resistant to superinfection with a reporter virus pseudotyped with the envelope proteins of KoRV A, GALV, or HPG (97.8 98.6% reduction in infectivity). Infections with retroviral particles pseudotyped with dualtropic or amphotropic MLV Env were also moderately inhibited (34.5% and 47.1% reduction in infectivity, respectively). Dualtropic MLV uses both PiT 1 & Pit 2 (SLC20A2) cell receptors (27), while amphotropic MLV exclusively uses PiT 2 (28). In contrast, superinfection by particles pseudotyped with the unrelated vesicular stomatitis virus (VSV) envelope G protein was not restricted. These data indicate that HPG utilizes the PiT 1 and PiT 2 cell receptors for cell entry."

Given that this analysis revealed inhibition of superinfection by amphotropic and dualtropic MLV, both of which utilize the PiT 2 receptor, we have included these viruses in an updated receptor binding domain alignment (Supplementary Figure [redacted]). This analysis revealed that as with HPG, amphotropic and dualtropic MLV also contained a significant insertion in the VRB domain relative to KoRV, GALV, and ecotropic MLV.

We have updated our description of the results of the RBD alignment to read as follows [redacted]:

"An alignment of the receptor binding (RBD) domain (31) of HPG against other KoRV related viruses reveals numerous differences in the variable regions (VRA and VRB) within the RBD (Supplementary Figure 6). Within this region, the pathologically important CETTG motif within the RBD (32), that is conserved in all other bat KoRV-related viruses, contains a threonine to serine mutation in HPG, resulting in a CETSG motif. HPG is more similar to GALV than to KoRV across both the VRA and VRB, where the RBD amino acid identities for HPG compared to GALV and KoRV are 66% and 62%, respectively. However, all of the KoRV-related bat gammaretroviruses analyzed contain a large insertion within the VRB of 10 and 16 amino acids, respectively relative to GALV and KoRV. Amphotropic and dualtropic MLV also contain several insertions within the VRB, increasing the length of their VRB region by 17 and 23 amino acids, relative to GALV and KoRV. These insertions are not present within ecotropic M MLV, which utilizes the mouse CAT1 (SLC7A1) cell receptor (33, 34). A complementary alignment of the binding motif within mammalian PiT 1 genes further supports this result as the binding sites within P. alecto and P. vampyrus PiT 1 share the permissive amino acid residues, which are distinct from the non permissive motif within mouse PiT 1 (Supplementary Figure [redacted]). Some gammaretroviruses that utilize PiT 1 for cell entry also utilize the related protein, PiT-2, and this has been attributed to subtle differences in the composition and length of amino acid sequences within the VRA and VRB

regions of the viral Env protein (35, 36). Taken together, these results indicate that HPG may share a similar host range as KoRV-A and GALV, with the caveat that the specific determinants of receptor usage and cell tropism for PiT 1 and PiT 2 are complex (35 37), and further investigation will be required to more accurately delineate the host range and cell tropism of HPG."

The following paragraph has been added to the discussion section [redacted]:

"Infection of cells with a retrovirus can restrict the subsequent superinfection by viruses that use the same receptor by various mechanisms including down regulation of the receptor, and blocking the binding site on the cell receptor, preventing penetration or adsorption of the virus (54, 55). This method has been used to demonstrate the shared use of the PiT 1 receptor between KoRV A and GALV (56). We undertook a superinfection interference assay which demonstrated that infection with HPG restricts superinfection by a reporter virus pseudotyped with the envelope protein of KoRV A, GALV, amphotropic MLV, and dualtropic MLV. KoRV-A and GALV utilize the PiT 1 receptor (19, 27, 28), while amphotropic MLV utilizes PiT-2 (57), and dualtropic MLV utilizes both PiT 1 and PiT 2 (29). These results indicate that HPG utilizes the PiT-1 and PiT-2 receptors for cell entry."

The following section has been added to the Supplementary Methods [redacted] as "Generation of HeLa cells persistently infected with HPG":

"HPG virion containing supernatants were generated as previously described in, "Transfection of 293T cells for generation of HPG & M MLV viral particles", and used to infect HeLa cells. Cells were seeded at a density of $7x10^5$ cells per T25 tissue culture flask (BD Biosciences, Bedford MA). Once cells reached 50% confluency, media was replaced with a mix of 4 ml DMEM, 1 ml HPG virion containing supernatant and DEAE Dextran (Sigma Aldrich) at a final concentration of 10 µg/mL. Cells were incubated for 16 h at which point the supernatant was removed, cells were washed twice in PBS and 5 ml of fresh DMEM was added. At 48 hours post infection, cells were passaged at a concentration of 1:5 into a new T25 flask. Cells were routinely passaged 1:5 twice weekly for three weeks and supernatants were tested for the presence of virion associated RT activity by a Product Enhanced Reverse Transcriptase (PERT) Assay, as previously described (13), except using a PrecisionPLUS qPCR SYBR Master Mix (Primer Design, Chandler's Ford, UK) and analysed on a QuantStudio 7 Flex Real Time PCR machine (Thermo Fischer Scientific)".

Commented [JH3]: Double-check this interval with AJ

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

Response: We thank the reviewer for noticing this. All of the supplementary figure labels have been updated in the revised manuscript.

Comment 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?

Response: We thank the reviewer for picking this up. The lane preceding HPG in both the reducing and non reducing conditions is a control lane containing expressed supernatant before binding/column purification of the polyhistidine tagged (His_6 tag) HPG Env ectodomain protein. We have modified the figure and legend (Supplementary Figure [redacted]) to include this information [redacted].

****

[redacted]

From: Joshua Hayward [redacted]
Sent: Saturday, 1 February 2020 12:30 PM
To: Wang Linfa
Cc: Edward Holmes; 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] Gary Crameri; Heidi Drummer; Andy Poumbourios; Glenn.Marsh; Michelle Baker; [redacted] Gilda Tachedjian
Subject: Re: HPG Paper PNAS Submission Response to reviewers and revised manuscript

Thanks, Linfa, agree it's serendipitous timing to be putting out a bat virus paper!

Joshua Hayward PhD

Research Officer
Retroviral Biology and Antivirals Research Laboratory

Burnet Institute
Disease Elimination and Maternal & Child Health Programs

[redacted]
GPO Box 2284 Melbourne Victoria Australia 3001
[redacted]

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

On Sat, 1 Feb 2020 at 12:24, Wang Linfa [redacted] wrote:

> Hi Josh,
>
> I scanned through Eddie's files and all ok with me. Absolutely flat out with the other "potentially bat virus (2019 nCoV)".
>
> Good luck and the timing is good: bats/viruses are headlines, again!
>
> LF
>
> ***Linfa (Lin-Fa) WANG, PhD FTSE***
>
> **Professor & Director**
>
> **Programme in Emerging Infectious Disease**

Duke-NUS Medical School,

s22

From: Joshua Hayward s22
Sent: Saturday, 1 February 2020 8:47 AM
To: Edward Holmes s22
Cc: mary.tachedjian s22 Kohl, Claudia s22 Adam Johnson s22 Dearnley, Megan (AAHL, Geelong AAHL) < s22 Brianna Jesaveluk s22 ; Christine Langer s22 Solymosi, Philip s22 Hille, Georg s22 Andreas s22 Cecilia Sanchez s22 Adam Werner s22 dimitri.kontos s22 Gary Crameri ; Heidi Drummer s22 Andy Poumbourios s22 Glenn.Marsh s22 Michelle Baker s22 Wang Linfa s22 <Ina.Smith s22 s22 Gilda Tachedjian >
Subject: Re: HPG Paper - PNAS Submission - Response to reviewers and revised manuscript

- External Email -

Much appreciated, Eddie, cheers!

Duplicate Email - Removed

[redacted]

From: Heidi Drummer [redacted]
Sent: Tuesday, 4 February 2020 10:59 AM
To: Joshua Hayward
Cc: Wang Linfa; Edward Holmes; 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.konto [redacted] Gary Crameri; Andy Poumbourios; Glenn.Marsh; Michelle Baker; [redacted] Gilda Tachedjian
Subject: Re: HPG Paper PNAS Submission Response to reviewers and revised manuscript
Attachments: Response to Reviewers v1.2_HD.docx; HPG paper Manuscript Revised v2.1 HD.docx; HPG paper - SI Appendix Revised v2.0 HD.docx

Hi Josh,
Please find attached my suggestions for the paper. I will go through and re-check all the data from the serology table today and tomorrow. As discussed, all the sera were screened against both Env trimer and HPG VRA peptide.
Overall the paper is in great shape. Well done!!
Cheers
Heidi

[redacted]
Duplicate Email - Removed

7th February 2020

Dear Editor,

Thank you for considering our manuscript entitled "**Infectious KoRV-related retroviruses circulating in Australian bats**" by Joshua A. Hayward, Mary Tachedjian, Claudia Kohl, Adam Johnson, Megan Dearnley, Brianna Jesaveluk, Christine Langer, Philip D. Solymosi, Georg Hille, Andreas Nitsche, Cecilia A. Sánchez, Adam Werner, Dimitri Kontos, Gary Crameri, Heidi E. Drummer, Pantelis Poumbourios, Glenn A. Marsh, Michelle L. Baker, Edward C. Holmes, Lin Fa Wang, Ina Smith, Gilda Tachedjian for publication in ***Proceedings of the National Academy of Sciences***.

Please find our responses to the reviewer's questions below. Revisions within the manuscript are highlighted in yellow.

Thank you for your consideration

Yours Sincerely,

Prof. Gilda Tachedjian
NHMRC, Senior Research Fellow
Head, Life Sciences Discipline
Head, Retroviral Biology and Antivirals Laboratory
Burnet Institute
s22

Editor's 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.

Reviewer #1:

General comments:

Here Hayward et al. identify and characterize the a full length, replication competent gammaretrovirus genome isolated from a bot in Australia, called HPG. In addition, four other related, partial retroviral genomes were isolated from other bot 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:

Comment 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.*

Response: Although appropriate blood samples were not available for analysis, we were able to detect KoRV related sequences in feces and urine. While KoRV has been detected in the blood of infected koalas (Tarlinton, 2006, Nature, 442:7098; Simmons, 2012, Aus. Vet. J. 90:10; Waugh, 2017, Sci. Rep. 7:1), KoRV nucleic acids and/or proteins have also been identified in various tissues, including sperm (Tarlinton, 2006, Nature, 442:7098), breast milk (Xu, 2013, PNAS, 110:28; Morris, 2016, Sci. Rep. 6) as well as feces (Wedrowicz, 2016,

Conserv. Genet. Resour. 8:4). The closely related Gibbon ape leukemia virus (GALV) has additionally been identified in the feces and urine of gibbons (Kawakami, 1977, J. Natl. Cancer. Inst. 268:5619).

KoRV A and GALV utilize the PiT 1 (SLC20A1) receptor (reviewed in Denner, 2016, Viruses, 8:12), which is a phosphate transport protein ubiquitously expressed at variable levels throughout the mammalian body (Kavanaugh, 1994, PNAS, 91:15; Johann, 1992, J. Virol. 66(3); also see expression database entries:

Expression Atlas (Petryszak, 2015, *Nucleic acids research,* D746 D752)
https://www.ebi.ac.uk/gxa/genes/ensg00000144136?bs=%7B%22homo%20sapiens%22%3A%5B%22ORGANISM_PART%22%5D%7D#baseline

Bgee (Bastian, 2008, In *International Workshop on Data Integration in the Life Sciences*, Springer, Berlin, Heidelberg)
https://bgee.org/?page=gene&gene_id=ENSG00000144136)

The highest expression levels of PiT 1 are in locations including the colon, testes, breast, bladder, placenta, and brain. Our experimental results (Fig 5A), including new data from infection interference assays (Fig 5B) suggests that HPG utilizes the same cell receptor, PiT 1, similar to KoRV A and GALV. Taken together, these observations are consistent with detecting KoRV related retroviruses (i.e. HPG) in the feces of bats as well as other pooled tissue samples.

Regarding potential routes for interspecies transmission we would like to clarify that we do not propose that a specific species to species transmission from bats to Koalas/Gibbon apes occurred for the KoRV related retroviruses identified in our study. Rather, our phylogenetic analysis suggests that there are likely retroviruses **more closely** related to KoRV/GaLV yet to be discovered. This was stated in the discussion section [redacted] as follows:

"Hence, bat communities could in theory provide a route of transmission for KoRV-related viruses between Asia and Australia, although the immediate ancestor of KoRV remains uncertain and it is clear that additional animal species need to be sampled. Indeed, there are likely to be other currently unidentified species infected with KoRV related viruses linking the habitats of R. hipposideros and Australian bats. The long phylogenetic branch length linking the KoRV clade to its closest known relatives in the GALV/WMV clade indicates that the

phylogenetic picture remains incomplete, with additional as yet unknown viruses and host species existing between the KoRV and GALV/WMV lineages of gammaretroviruses."

Further regarding interspecies transmission, in general, given the diversity of body fluids within which KoRV and GALV (and by extension, KoRV related viruses) might be found, a number of possible scenarios may be reasonably speculated. These include transmission via blood during fighting/predation, and contamination of food sources by feces and urine.

To address the comments raised by the reviewer, we have included the following paragraph in the Discussion at [ ]:

"KoRV and GALV utilize the PiT-1 receptor for cell entry (19, 27, 28). This receptor is almost ubiquitously expressed throughout the mammalian body at variable levels (40 43), and is highly expressed in many tissues including the colon, breast, testes, bladder, placenta, and brain (40, 41). KoRV and GALV have been detected in numerous tissues and body fluids including blood, sperm, breast milk, feces, and urine (5, 27, 44 49). Given the wide distribution of PiT-1 expression and the detection of KoRV and GALV in body fluids including blood, urine, and feces, it is possible that interspecies transmission might occur along routes including blood during fighting/predation, and contamination of food sources by feces and urine."

Commented [PHD1]: Similarly here, if you change discussion remember to change this

Comment 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.*

Response: We thank the reviewer for raising this point and can appreciate why our statements appear contradictory. We agree that we cannot rule out the possibility that HPG is not endogenous in some proportion of the bat gene pool since we have only sampled a small number of bats. Accordingly, we have modified the manuscript as follows to soften our claims:

We have modified our result heading from

"*HPG is not an endogenous retrovirus*"

to

"*HPG sequences were not detected in the genomes of pteropid bats*" on [page, line].

Furthermore, within the discussion section, we have modified the text to read as follows on [page, line]:

"*We searched carefully for the presence of HPG in the genomes of P. alecto and P. vampyrus using molecular analyses, and more broadly for KoRV-related viruses in the SRA and were unable to detect these viral sequences in the genome of any bat species whose genome is currently available. While these data suggest that bat KoRV related viruses are not endogenous, we cannot rule out the possibility as we have only sampled a small proportion of bats within each species. In this regard, KoRV endogenization in koalas is relatively recent, and accordingly is not represented across the entire koala gene pool (44); existing in both endogenous and exogenous forms (5, 27, 50). Thus, given that HPG specific sequences have been identified across several bat species, either HPG is an exogenous virus or it is undergoing endogenization in real time. A possible example of the latter is FFRV1 (14), which was recently discovered in the brain tissue of a P. alecta bat, but which we were not able to identify within the genome of P. alecto or other bats.*"

Regarding our BLAST search within pteropid genomes for HPG sequences, when we analysed the genome of *P. alecto* for sequences similar to HPG, our BLAST analysis revealed that HPG and closely related retroviral sequences were not present. The closest identified hit against the HPG sequence in this analysis was a 546 nt sequence aligning to the *pol* gene of HPG, with a nucleotide identity of 69%.

To address this matter, we have replaced the sentence on [page, line]:

"No sequences with high percent nucleotide similarity to HPG were identified"

with

"No sequences matching HPG were identified. The closest identified hit against the HPG sequence in this analysis was a 546 nt sequence within the genome of P. alecto, aligning to the pol gene of HPG, with an e value of 5.0 x 10^{46} and a nucleotide identity of 69%."

Comment 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).*

Response: We apologize that the sampling period was not clear. The sampling period was only over seven years (as stated on pg x Lines-Y) and accordingly is highly unlikely to have an impact on relatedness in the context of the evolutionary time scale depicted in our phylogeny which is thousands to millions of years (Holmes EC. (2009). The Evolution and Emergence of RNA Viruses. Oxford Series in Ecology and Evolution, Oxford University Press, Oxford.).

Comment 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.*

Response: To clarify, the samples referenced in this sentence tested positive for HPG *specific* RNA, but not HPG *related* RNA indicating that the quality of the RNA in these samples was adequate for this analysis. While we did not detect HPG related RNA in these samples, we did detect HPG related DNA.

[Add info about determination of concentrations]. During the qRT PCR analysis, bat samples were classified as positive or negative based on their fluorescence signal compared a standard curve generated using $1x10^0$ $1x10^7$ copies of the HPG proviral plasmid. In the standard curve, signal was only generated down to a threshold of $1x10^1$ copies, the signal for which appeared at cycle 36 (CT 36), and this CT value served as the cut off for determining a positive result.

Commented [JH2]: We will include summary info regarding nucleic acid concentrations for the 373 samples prior to resubmission.

While we cannot exclude that some amount of HPG related RNA was present in the samples and below the limits of detection, we did utilize a highly sensitive kit (Thermo Power SYBR Green RNA to CT Kit) that is capable of detecting specific targets from sub picogram levels of total RNA.

To address the reviewer's alternative explanation, we have modified this sentence to now read as [page, line]:

"While we cannot rule out that some or all of these samples may have contained HPG related RNA below the limit of detection of this assay, these data suggests that 12 bat samples were actively infected with HPG and were either latently infected with other HPG related virus(es) or contain endogenous HPG related sequences."

Within the supplementary methods [redacted] we have added the following details:

"Bat samples were classified as positive or negative based on their fluorescence signal compared against a standard curve generated using $1x10^0 – 1x10^7$ copies of the HPG proviral plasmid. The cut off for determining a positive result was a cycle threshold of 36, which correlated to $1x10^1$ copies of the HPG provirus."

Minor Comments:

Comment 5: *The Introduction (page 5, line 1) and the Discussion (page 10, line 3) state "the Daintree rainforest," which implies that the Dointree 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.*

Response: To clarify this statement we have modified the text at various locations to read as follows:

[location]: *"we collected bat samples (feces, blood, urine, and oral swabs) from towns and the Daintree rainforest along the east coast of Australia"*

[location]: *"To identify KoRV-related viruses in bats, samples were collected from the east coast of Australia, including feces, oral swabs, blood, and urine. 373 samples were collected from towns in New South Wales and Queensland and 106 from the Daintree Rainforest (Queensland)."*

[location]: *"To determine whether KoRV related viruses are present in Australian bats, we collected samples from bats on the east coast of Australia"*

Comment 6: *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.*

Response: The Genbank accession numbers for the bat retroviruses have been included in the updated version of Supplementary Table [redacted].

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

Response: We thank the reviewer for this suggestion and have included this description as follows in the results section:

"A complementary alignment of the binding motif within mammalian PiT 1 genes further supports this result as the binding sites within P. alecto and P. vampyrus PiT 1 share the permissive amino acid residues, which are distinct from the non permissive motif within mouse PiT 1 (Supplementary Figure [redacted])." on [page, line].

Comment 8: *Supplementary Figure 9 is not discussed anywhere in the text. It is suggested that it be removed or appropriately described in the text.*

Response: A mislabeled reference to this Supplementary Figure within the text of the Supplementary Methods section has been corrected, and now reads:

"SDS PAGE in the presence and absence of β-mercaptoethanol revealed a single diffuse band with a molecular weight range of ~80 90 kDa (Supplementary Figure [redacted]), consistent with the molecular weight predicted from the amino acid sequence (62,805 Da) with 6 N linked glycans (~ 18 kDa)." on [page, line].

Reviewer #2:

General 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.

Comment 1: *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).*

Response: We agree with the comment made by the reviewer, and apologize that this was not clear in the manuscript. We have addressed this comment in our response to Reviewer 1's Major Comment 2.

Comment 2: *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).*

Response: As suggested by the reviewer, we have now provided individual phylogenies for the *env*, *pol* and *gag* genes (Supplementary Figure). As can be seen, the tree topologies for *env* and *pol* genes are the same as that for the complete viral genomes (Figure 2). A slightly different topology was observed in the *gag* gene phylogeny, however, as all the relevant bootstrap values were very low (35%, 41%, 48%), a history of genomic recombination cannot be safely inferred since the difference in tree topology in the *gag* gene lacks phylogenetic resolution.

We have included this additional analysis as Supplementary Figure , on [page, line].

We have included the corresponding text within the Results section:

"This analysis is supported by phylogenetic analyses of the individual pol and env genes, which reveal the same branching pattern. While analysis of the gag gene resulted in a slightly different branching pattern, this is likely as a result of low phylogenetic resolution, as indicated by low bootstrap support for this individual tree (Supplementary Figure)." on [page, line].

We have also updated the Supplementary methods section to include description of the phylogenetic analysis as follows on [page, line]:

"To determine the evolutionary relationships among KoRV-related gammaretroviruses, we performed phylogenetic analyses using aligned complete genome nucleotide sequences (Supplementary Table 2) and individual gene sequences. Accordingly, a multiple sequence alignment of 19 complete genomes was performed using a combination of MAFFT (8) and MUSCLE algorithms (9). Following alignment, regions of ambiguous and uncertain alignment were removed using Gblocks (10). For the complete genomes, this resulted in final alignment of 6,925 nt that was used to infer evolutionary relationships. Subsets of this alignment covering the gag, pol, and env gene regions were used for the individual gene analyses. Phylogenetic trees of these data were estimated using the maximum likelihood (ML) method available in the PhyML program (11), assuming a GTR model of nucleotide substitution with a proportion of invariant sites (I) and a gamma distribution of among site rate variation (Γ). To determine the robustness of each node a bootstrap resampling analysis (1,000 replications) was performed using the same nucleotide substitution model. For the complete genome tree (Figure 2), a

Shimodaira Hasegawa (SH) test was conducted, providing additional nodal support. The Mus caroli ERV, McERV (Supplementary Table 2), sequence was used as an outgroup to root the tree."

Comment 3: *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.*

Response: The reviewer has raised an important issue with regards to HPG specific seropositivity across the tested bat samples. To address this issue we have undertaken a peptide binding analysis in a solid phase enzyme immunoassay to assess the seroreactivity of bat samples against short peptide sequences from the VRA region specific to HPG, KoRV, GALV, and the more distantly related MLV, in addition to the HPG Env trimer. The assay was validated using high titre immune serum raised to HPG Env in rabbits, which showed specific binding to HPG VRA peptide, but not KORV, GALV or MLV VRA peptides. In addition, a macague immune serum raised to MLV only showed reactivity to MLV VRA peptide. Whilst we cannot exclude that antibodies that develop in bats infected with HPG can cross react with peptides from KORV and GALV, the data strongly suggest that 32% of bats are infected with HPG or a KoRV related viruses, with 27% of P Alecto bats showing seropositivity.

These new serology results have been included (Supplementary Figure) and supersede the luminex data within the Results section, which now reads as follows, on [page, line]:

"To assess Australian bats for exposure to HPG or KoRV related viruses, we tested for the presence of bat antibodies reactive against the HPG Env protein. To determine whether bat samples that were seropositive for HPG might be cross reactive against the closely related gammaretroviruses, KoRV A and GALV, or the more distantly related gammaretrovirus, MLV, we performed a peptide binding analysis using short peptide sequences derived from the Variable Region A within the Env protein of these viruses. We also tested for the presence of HPG specific nucleic acid in bat fecal samples.

Bat sera (87 samples) were screened for the presence of antibodies reactive to the HPG VRA region of Env in a solid phase enzyme immunoassay. Where available sample material permitted, analysis was additionally conducted on the full range of VRA peptides and the HPG Env trimer ectodomain (Glu^{38} Ser^{603}) (Supplementary Figure X).

Of the 87 bat samples, 18 (20.7%) were reactive to the HPG VRA peptide [P. Alecto (n = 16), P. conspicillatus (n = 1), Rhinolopus megaphyllus (n = 1)]. All of these samples were also reactive to the HPG Env ectodomain. Nine (50%) of the HPG VRA positive samples were also cross reactive for KoRV A and GALV VRA peptides. However, only a single sample (#20 P. alecto) was more strongly cross reactive, against the GALV VRA peptide. Two samples (#7 P. alecto and #8 P. alecto) were reactive against the KoRV A and GALV VRA peptide, respectively but not reactive against the HPG VRA peptide. Only a single bat (#27 P. alecto) demonstrated cross reactivity to MLV, and this bat was also reactive against the VRA of HPG, KoRV-A, GALV, and HPG Env.

These results reveal that 28% of bat samples were seropositive for HPG or KoRV related protein sequences. They also indicate that while some cross reactivity is observable, reactivity among the tested bat samples is almost entirely strongest against HPG, and cross reactivity generally does not extend to distant gammaretroviral relatives such as MLV."

Commented [PHD3]: Make it consistent with suggested new text.

The methodology in the supplementary methods section [redacted] "*Serological assay for the presence of onti HPG antibodies in bats*" now reads:

"Bat sera were screened for the presence of antibodies reactive to the VRA region of Env using synthetic peptides in a solid phase enzyme immunoassay. N terminal biotinylated synthetic peptide encoding the HPG VRA region (LETWDIPDSDVSASTRVRPADSD, Genscript, USA) was added to Avidin coated plates (Nunc, Maxisorb) at 5 µg/ml followed by the addition of serially diluted bat serum in PBS containing 2.5 mg/ml bovine serum albumin and Tween 20 (0.05%). Bound antibodies were detected with horseradish peroxidase labelled Protein A/G (Thermo Scientific, Rockford) followed by 3,3',5,5' Tetramethylbenzidine (TMB) substrate (Sigma, USA). Sera that displayed above ~~10x~~ Sx background levels of binding were further screened for reactivity to biotinylated synthetic peptides of the equivalent regions of KoRV-A (LESWDIPELTASASQQARPPDSN), GALV (LESWDIPGTDVSSSKRVRPPDSD), and MLV (PSYWGLEYQSPFSSPPGPPCCS) in the same way."

Comment 4: *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.

Response: We apologise that these values were not clearly explained. The values are not fractions, but the combination of two different measures of nodal support SH like branch support and bootstrap support i.e. *'SH value/Bootstrap value'*
We have updated the figure legend to clarify this, and it now reads as follows:

"...All branches are scaled according to the number of nucleotide substitutions per site, and branches representing bat retroviruses are shown in red. Support for key nodes on the phylogeny are shown in the form ' like branch support/bootstrap support. Silhouettes represent the host species..."

Comment 5: 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?

Response: In thin section electron microscopy, the analysis is taking place on a single plane (70 90 nm) section of the cell, the height of which can be up to 20 µm in a cell monolayer and larger in suspension depending on the cell's orientation and morphology. As such, we feel that detection of viral particles measured by EM analysis alone is not the most efficient approach to measuring % infection and/or providing 100% confidence ' finding (or not finding) viral particles in a test or control sample. For this reason, as described in the Supplementary Methods sections "Transfection of 293T cells for generation of HPG & M MLV viral particles" and "Electron Microscopy and viral particle morphology", we complimented our EM analysis with a virion associated reverse transcriptase activity (RT) assay on supernatant from each of the cell suspension samples that was to be analyzed by electron microscopy. A positive reading was obtained for MLV and HPG but not for the negative controls (cells with no transfection and the empty plasmid) indicating virus budding and activity

in the test samples only. In sample blocks containing MLV 293T cells and HPG 293T cells, viral particles were readily observed budding from the cell membrane or in inclusion bodies within the cells, indicating a relative abundance of virus in the cells. For the negative controls, which were untransfected 293T cells and 293T cells mock transfected with the pcDNA3.1 plasmid, no virus was observed in cells following extensive examination across numerous fields of view, multiple sections and on two separate grids. Furthermore, cell morphology and ultrastructure of control samples was consistent with healthy cells in tissue culture. In contrast, cells in the population that had transfected with MLV and HPG showed morphological indictors of infection such as fragmented cell and organelle membranes, extracellular debris (membrane) and in some instances, cytoplasmic or nuclear condensation. Together these data increase our confidence that control samples were truly negative for virus, whilst the MLV 293T cells and HPG 293T cells did contain replicating virus particles.

We have included the quality control data from the virion associated RT assay alongside the EM negative controls in Supplementary Figure [redacted]. The legend for this figure now reads:

"Supplementary Figure [redacted]. Electron micrographs (EM) of control untransfected cells and cells mock transfected with the empty vector pcDNA3.1, and virion associated reverse transcriptase (RT) assay. The graph displays the result of a virion associated RT assay, supporting the result that cells transfected with proviral M MLV and HPG expression plasmids generate retroviral particles, while untransfected and mock pcDNA3.1 transfected cells do not. A) Transmission electron micrograph..."

We appreciate the reviewer's suggestion that we make appropriate mention of the negative EM controls in the main text and have modified our reference to these data in the results section [redacted], which now reads:

"Transfection of human 293T cells with a plasmid construct carrying the HPG provirus resulted in the generation and release of viral particles morphologically similar to ecotropic Moloney murine leukemia virus (M MLV), as determined by electron microscopy (Figure 3), in contrast to untransfected and mock transfected 293T cells (Supplementary Figure [redacted]). These data are supported by virion associated reverse transcriptase analysis of the samples analyzed by electron microscopy (Supplementary Figure [redacted])."

The following sentence has been appended to the end of the legend of Figure 3 [redacted]:

"Negative transfection controls were untransfected cells and cells mock transfected with the empty vector pcDNA3.1. These controls were not observed to contain or produce viral particles (Supplementary Figure)."

Comment 6: 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 llowed by RT assay, or replication with and without inhibiting RT, etc).

Response: To address this possibility, and as suggested by the reviewer, we conducted a 'secondary infection assay', in which we established successive rounds of replication, and is included as Supplementary Figure . In brief, 293T cells were transfected with the HPG proviral plasmid; cell culture supernatant was later harvested and clarified. This clarified supernatant was used to establish a primary infection in 293T cells in the same manner as our original infectivity assay. We then collected the clarified supernatant of these cells and repeated the process, and successfully established a secondary infection in 293T cells, as determined by a virion associated PERT assay. The data from this experiment confirms that successive rounds of replication can be established by HPG in 293T cells. In contrast to our original experiments we used a PERT assay (RT qPCR) to confirm the presence of HPG for these new experiments, as we have ceased using the radiolabeled virion associated reverse transcriptase assay due to increased and prohibitive costs of radiolabeled nucleotides in our region.

We have added this analysis to the Results section [ ]:

"HPG establishes successive rounds of replication through a secondary infection assay (Supplementary Figure)."

The methodology for this assay has been added to the Supplementary Methods as "Secondary infection assay" [ ], and reads:

"To confirm that HPG was capable of establishing successive rounds of infection, 239T cells were transfected with an infectious molecular clone of HPG as described in "Transfection of 293T cells for generation of HPG & M MLV viral particles". To establish a primary infection from HPG virions, 293T cells were infected as described in "Replication kinetics assay". To establish a secondary infection, clarified supernatant harvested 48 h following the primary infection was collected and used to establish a second round infection in 293T cells as described in "Replication kinetics assay", except that for the second round infection neat HPG supernatant was used; 5 µL samples were collected at inoculation, 6 h following inoculation and washing (t = 0), and at 48 h following washing (t = 48). Collected samples were analyzed for the presence of virion associated reverse transcriptase activity by PERT assay, as described in "Generation of HeLa cells persistently infected with HPG". Values derived from the PERT assay represent arbitrary units of RT activity in comparison to a dilution series of HPG virions which were generated in house, as described in "Transfection of 293T cells for generation of HPG & M MLV viral particles"."

Other comments

Comment 7: 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.

Response: We thank the reviewer for this suggestion and have undertaken a superinfection interference assay, included as an additional panel (B) in Figure 5. Briefly, in this experiment we generated persistently HPG infected HeLa cells, then challenged these cells with infection by Envelope pseudotyped reporter retroviruses representing HPG, KoRV A, GALV, Amphotropic MLV, Dualtropic MLV, and the vesicular stomatitis virus (VSV). Compared against uninfected HeLa cells, HPG infected HeLa cells were strongly resistant to superinfection from HPG, KoRV A, and GALV Env pseudotyped viral particles; they were moderately resistant to infection by amphotropic and dualtropic MLV (which respectively use the PiT 2 and PiT 1 & PiT 2 cell

Commented [JH4]: We will add PERT data demonstrating maintenance of HPG infection over the course of passaging.

receptors [Feldman, 2004, J. Virol. 78:2; Miller, 1996, J. Virol. 70:8]) Env pseudotyped particles; almost no impact was observed on susceptibility to infection by VSV. The ecotropic MLV used in our infection kinetics assay was not utilized in this assay as it is incapable of infecting human cells. These results suggest that HPG utilizes the PiT 1 and PiT 2 receptors for cell entry.

We have added the following description of this analysis to the Results section [redacted]:

"To further investigate receptor usage by HPG, we performed a superinfection interference assay (Figure 5B). In this assay, HeLa cells persistently infected with HPG became strongly resistant to superinfection with a reporter virus pseudotyped with the envelope proteins of KoRV A, GALV, or HPG (97.8 98.6% reduction in fectivity). Infections with retroviral particles pseudotyped with dualtropic or amphotropic MLV Env were also moderately inhibited (34.5% and 47.1% reduction in infectivity, respectively). Dualtropic MLV uses both PiT 1 & Pit 2 (SLC20A2) cell receptors (27), while amphotropic MLV exclusively uses PiT 2 (28). In contrast, superinfection by particles pseudotyped with the unrelated vesicular stomatitis virus (VSV) envelope G protein was not restricted. These data indicate that HPG utilizes the PiT 1 and PiT 2 cell receptors for cell entry."

Given that this analysis revealed inhibition of superinfection by amphotropic and dualtropic MLV, both of which utilize the PiT 2 receptor, we have included these viruses in an updated receptor binding domain alignment (Supplementary Figure [redacted]). This analysis revealed that as with HPG, amphotropic and dualtropic MLV also contained a significant insertion in the VRB domain relative to KoRV, GALV, and ecotropic MLV.

We have updated our description of the results of the RBD alignment to read as follows [redacted]:

"An alignment of the receptor binding (RBD) domain (31) of HPG against other KoRV related viruses reveals numerous differences in the variable regions (VRA and VRB) within the RBD (Supplementary Figure 6). Within this region, the pathologically important CETTG motif within the RBD (32), that is conserved in all other bat KoRV related viruses, contains a threonine to serine mutation in HPG, resulting in a CETSG motif. HPG is more similar to GALV than to KoRV across both the VRA and VRB, where the RBD amino acid identities for HPG compared to GALV and KoRV are 66% and 62%, respectively. However, all of the KoRV related bat gammaretroviruses analyzed contain a large insertion within the VRB of 10 and 16 amino acids, respectively relative to GALV and KoRV. Amphotropic and dualtropic MLV also contain several insertions within the VRB,

increasing the length of their VRB region by 17 and 23 amino acids, relative to GALV and KoRV. These insertions are not present within ecotropic M MLV, which utilizes the mouse CAT1 (SLC7A1) cell receptor (33, 34). A complementary alignment of the binding motif within mammalian PiT 1 genes further supports this result as the binding sites within P. alecto and P. vampyrus PiT 1 share the permissive amino acid residues, which are distinct from the non permissive motif within mouse PiT 1 (Supplementary Figure [redacted]). Some gammaretroviruses that utilize PiT 1 for cell entry also utilize the related protein, PiT 2, and this has been attributed to subtle differences in the composition and length of amino acid sequences within the VRA and VRB regions of the viral Env protein (35, 36). Taken together, these results indicate that HPG may share a similar host range as KoRV A and GALV, with the caveat that the specific determinants of receptor usage and cell tropism for PiT 1 and PiT 2 are complex (35 37), and further investigation will be required to more accurately delineate the host range and cell tropism of HPG."

The following paragraph has been added to the discussion section [redacted]:

"Infection of cells with a retrovirus can restrict the subsequent superinfection by viruses that use the same receptor by various mechanisms including downregulation of the receptor, and blocking the binding site on the cell receptor, preventing penetration or adsorption of the virus (54, 55). This method has been used to demonstrate the shared use of the PiT-1 receptor between KoRV-A and GALV (56). We undertook a superinfection interference assay which demonstrated that infection with HPG restricts superinfection by a reporter virus pseudotyped with the envelope protein of KoRV A, GALV, amphotropic MLV, and dualtropic MLV. KoRV A and GALV utilize the PiT-1 receptor (19, 27, 28), while amphotropic MLV utilizes PiT-2 (57), and dualtropic MLV utilizes both PiT 1 and PiT 2 (29). These results indicate that HPG utilizes the PiT 1 and PiT 2 receptors for cell entry."

The following section has been added to the Supplementary Methods [redacted] as "Generation of HeLa cells persistently infected with HPG":

"HPG virion containing supernatants were generated as previously described in, "Transfection of 293T cells for generation of HPG & M MLV viral particles", and used to infect HeLa cells. Cells were seeded at a density of $7x10^5$ cells per T25 tissue culture flask (BD Biosciences, Bedford MA). Once cells reached 50% confluency, media was replaced with a mix of 4 ml DMEM, 1 ml HPG virion containing supernatant and DEAE Dextran (Sigma Aldrich) at a final concentration of 10 µg/mL. Cells were incubated for 16 h at which point the supernatant was removed,

cells were washed twice in PBS and 5 ml of fresh DMEM was added. At 48 hours post infection, cells were passaged at a concentration of 1:5 into a new T25 flask. Cells were routinely passaged 1:5 twice weekly for three weeks and supernatants were tested for the presence of virion associated RT activity by a Product Enhanced Reverse Transcriptase (PERT) Assay, as previously described (13), except using a PrecisionPLUS qPCR SYBR Master Mix (Primer Design, Chandler's Ford, UK) and analysed on a QuantStudio 7 Flex Real Time PCR machine (Thermo Fischer Scientific)".

Commented [JH5]: Double-check this interval with AJ

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

Response: We thank the reviewer for noticing this. All of the supplementary figure labels have been updated in the revised manuscript.

Comment 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?

Response: We thank the reviewer for picking this up. The lane preceding HPG in both the reducing and non reducing conditions is a control lane containing expressed supernatant before binding/column purification of the polyhistidine tagged (His_6 tag) HPG Env ectodomain protein. We have modified the figure and legend (Supplementary Figure [redacted]) to include this information [redacted].

****

Title Page

Classification
BIOLOGICAL SCIENCES, Microbiology

Title
Infectious KoRV related retroviruses circulating in Australian bats

Authors and Affiliations
Joshua A. Hayward[1,a,h], Mary Tachedjian[1,b], Claudia Kohl[c], Adam Johnson[a], Megan Dearnley[b], Brianna Jesaveluk[a], Christine Langer[a], Philip D. Solymosi[c], Georg Hille[c], Andreas Nitsche[c], Cecilia A. Sánchez[d,e], Adam Werner[a], Dimitri Kontos[a], Gary Crameri[b], Glenn A. Marsh[b], Michelle L. Baker[b], Pantelis Poumbourios[a,f], Heidi E. Drummer[a,f,i], Edward C. Holmes[g], Lin-Fa Wang[h], Ina Smith[2,b], Gilda Tachedjian[2,3,a,f,i]

[1]J.A.H. and M.T. contributed equally to this work,

[2]I.S. and G.T. contributed equally to this work.

[a]Burnet Institute, Life Sciences, Melbourne, 3004, Australia; [b]Australian Animal Health Laboratory, Health and Biosecurity Business Unit, Commonwealth Scientific and Industrial Reseach Organisation, Geelong, 3220, Australia; [c]Robert Koch Institute, Centre for Biological Threats and Special Pathogens, Berlin, 13353, Germany; [d]University of Georgia, Odum School of Ecology, Athens, GA; [e]Center for the Ecology of Infectious Diseases, University of Georgia, Athens, GA; [f]Monash University, Department of Microbiology, Melbourne, Australia; [g]University of Sydney, School of Life and Environmental Sciences and Sydney Medical School, Sydney, Australia; [h]Duke NUS Medical School, Programme in Emerging Infectious Diseases, Singapore, Singapore; [i]University of Melbourne, Department of Microbiology and Immunology at the Doherty Institute for Infection and Immunity, Melbourne, Australia

Corresponding Author
[3]To whom correspondence may be addressed. Email: s22

Keywords
Bats; Retroviruses; Gammaretrovirus; GALV; KoRV; Pteropid

Abstract

Bats are reservoirs of emerging viruses that are highly pathogenic to other mammals including humans. Despite the diversity and abundance of bat viruses, to date they have not been shown to harbor exogenous retroviruses. Here we report the discovery and characterization of a group of Koala retrovirus-related (KoRV-related) gammaretroviruses in Australian and Asian bats. These include the Hervey pteropid gammaretrovirus (HPG), identified in the scat of the Australian black flying fox (*Pteropus alecto*), which is the first reproduction-competent retrovirus found in bats. HPG is a close relative of KoRV and the Gibbon ape leukemia virus (GALV), with virion morphology and Mn^{2+}-dependent virion associated reverse transcriptase activity typical of a gammaretrovirus. *In vitro,* HPG is capable of infecting bat and human cells, but not mouse cells, and displays a similar pattern of cell tropism as KoRV-A and GALV. Population studies reveal the presence of HPG and KoRV related sequences in several locations across north east Australia as well as serological evidence for HPG in multiple pteropid bat species, while phylogenetic analysis places these bat viruses as the basal group within the KoRV-related retroviruses. Combined, these results reveal bats to be important reservoirs of exogenous KoRV-related gammaretroviruses.

Significance Statement

Bats represent 20% of all mammalian species, and are an important reservoir of viruses that infect humans and other mammals. Retroviruses, such as HIV, are among the most important zoonotic viruses infecting humans, although little is known about their circulation in bat populations. We report the first exogenous retrovirus described in bats, denoted the Hervey pteropid gammaretrovirus (HPG): a reproduction competent retrovirus within north east Australia. Koala populations are currently in severe decline and at risk from koala retrovirus (KoRV), which is closely related to HPG and whose origins remain unclear. The identification of bats as a source of diverse infectious retroviruses related to KoRV implicates bats as a reservoir of KoRV-related viruses that potentially can be transmitted to other mammalian species.

Introduction

Retroviruses are a widespread and diverse group of RNA viruses distinguished by their ability to integrate into the genome of their host cell (1). Several retroviruses cause immunodeficiency [e.g. Human immunodeficiency virus; HIV (2)] and malignancies such as leukemia [e.g. Koala retrovirus; KoRV (3-5)]. When retroviruses integrate into germline cells, they become 'vertically' transmissible from parent to offspring, and are referred to as endogenous retroviruses (ERVs) (1, 6). KoRV, for instance, is an infectious retrovirus currently undergoing endogenization in the koala gene pool (5). Through the course of evolutionary history, ERVs and related retroelements have become ubiquitous across metazoan genomes (6 8): for example, 8% of the human genome is derived from retroviruses (9). ERVs may or may not be capable of producing infectious viral particles.

Bats are reservoirs for many viruses from diverse viral families, and are implicated in the transmission of numerous highly pathogenic viruses to humans and other mammals (10). Previous studies have revealed the presence of ERVs from the genera *Betaretrovirus*, *Gammaretrovirus*, and *Deltaretrovirus* within the genomes of bats (11-14). Analyses of the evolutionary relationships between these bat ERVs and those from other mammals imply that bats have played a key role in the transmission of retroviruses between different mammalian species (15, 16). Indeed, genomic analysis indicates that bats have served as hosts to retroviruses for most of their evolutionary history (11), and evidence of gene expansion and diversification in the antiretroviral APOBEC3 family of immune restriction factors suggests an ongoing relationship between bats and retroviruses (17). At present, however, no infectious, horizontally transmissible exogenous retroviruses (XRVs) have been identified and reported in bats.

KoRV and the Gibbon ape leukemia virus (GALV) are closely related gammaretroviruses (77.5% nucleotide identity). However, the habitats of the hosts of these viruses (koalas in Australia and gibbons in South-East Asia) do not overlap, and are physically separated by the oceanic faunal boundary known as the Wallace line (18). It has been suggested that bats may have played a role in the transmission of gammaretroviruses between gibbons and koalas (19 21). In particular, the habitat of bats such as the black flying fox, *Pteropus alecto*, overlap and connect the habitats of both gibbons and koalas, with bats being capable of traversing the bodies of water that separate the islands of Australia and South East Asia (22). In addition, bat gammaretroviral ERVs are widely distributed across the broader gammaretroviral phylogeny (23), with one recently discovered bat gammaretroviral ERV reportedly falling between KoRV and GALV on phylogenetic trees (14).

To advance our understanding of the role of bats as hosts and potential transmitters of gammaretroviruses closely related to KoRV and GALV (herein referred to as KoRV related viruses), between 2007 and 2014 we collected bat samples (feces, blood, urine, and oral

swabs) from towns and the Daintree rainforest along the east coast of Australia to detect the presence of KoRV-related viruses. From this survey we report the identification and characterization of a novel reproduction-competent bat retrovirus, the Hervey pteropid gammaretrovirus (HPG), from *P. alecto*. In addition, we identified novel gammaretroviral sequences from two species of pteropid bats, *Macroglossus minimus* and *Syconycteris australis,* and two species of Yinpterochiropteran microbats from China, *Hipposideros larvatus* and *Rhinolophus hipposideros.* These gammaretroviral sequences are closely related to KoRV and GALV.

Results

Metagenomic analyses reveal the presence of novel KoRV-related gammaretroviruses in Australian and Asian bats

To identify KoRV related viruses in bats, samples were collected from the east coast of Australia, including feces, oral swabs, blood, and urine. 373 samples were collected from towns in New South Wales and Queensland and 106 from the Daintree Rainforest (Queensland). The species of origin was determined by species-specific cytochrome B gene TaqMan RT PCR. Metagenomic analysis of RNA extracted from the bat samples revealed the presence of KoRV-related viruses in samples collected from the pteropid bat species (subfamily Yinpterochiroptera) *Pteropus alecto* (HPG), *Macroglossus minimus* (Macroglossus minimus gammaretrovirus, MmGRV), and *Syconycteris australis* (Syconycteris australis gammaretrovirus, SaGRV). To broaden our search, we probed the Sequence Read Archive (SRA) for the presence of KoRV-related viruses. This search revealed the presence of two additional viruses in metagenomic RNA extracted from samples obtained from the Asian microbat species (subfamily Yinpterochiroptera) *Hipposideros larvatus* (Hipposideros larvatus gammaretrovirus, HlGRV) and *Rhinolophus hipposideros* (Rhinolophus hipposideros gammaretrovirus, RhGRV). The identified KoRV-related viruses and their origins are summarized in Supplementary Table 1.

The complete genome sequence of HPG, and partial genome sequences of MmGRV, SaGRV, HlGRV, and RhGRV, were assembled and deposited in GenBank (Supplementary Table 2). The source of HPG was a fecal sample collected in 2011 from a single flying fox in Hervey Bay. HPG viral particles in the sample were enriched using a sucrose gradient, total RNA extracted and genomic DNA removed. The complete HPG genome sequence was generated from this total RNA sample by employing a modified single-cell whole transcriptome amplification (WTA) procedure for detecting ultra low-copy viral RNA and a *de novo* sequence assembly pipeline outlined in the supplementary methods. The HPG genome is 8,030 nt in length, similar to KoRV-A and GALV (7,994 nt and 8087 nt, respectively), and contains terminal repeats (R), 5' and 3' unique regions, and open reading frames encoding the canonical gammaretroviral genes *gag*, *pol*, and *env* that do not contain any frameshift mutations or premature stop codons (Figure 1). Other genomic elements essential for retroviral replication and reproduction, including the expected protease, polymerase, and integrase active site motifs, proline tRNA primer binding site, polypurine tract, and polyadenylation signal site, were also present (Figure 1 and Supplementary Figure 1A). The assembled partial genome sequences of MmGRV and SaGRV lacked coverage only at the terminal repeat and unique 5' and 3' regions, while open reading frames encoding *gag*, *pol*, and *env* were intact and free from frameshift mutations or premature stop codons (Supplementary Figure 2). For HlGRV and RhGRV, overall read coverage was low, and in both cases coverage dropped to zero at some locations within each of *gag*, *pol*, and *env* (Supplementary Figure 2).

Results

HPG sequences were not detected in the genomes of pteropid bats

To exclude the possibility that HPG represents a fossilized ERV, we performed a BLAST analysis of the *P. alecto* and *P. vampyrus* genomes. No sequences matching HPG were identified. The closest identified hit against the HPG sequence in this analysis was a 546 nt sequence within the genome of *P. alecto,* aligning to the *pol* gene of HPG, with an e-value of $5.0x10^{-46}$ and a nucleotide identity of 69%. We then performed a HPG specific PCR analysis of the *P. alecto* genome, using genomic DNA extracted from two sources, *P. alecto* tissue from a male bat captured in Brisbane (Australia), and a *P. alecto* kidney cell line (24). This PCR analysis did not generate detectable amplicons, in contrast to amplification of a single copy bat APOBEC3Z3 gene (17) (Supplementary Figure 3). These data suggest that HPG has not integrated into the germline of the *P. alecto* bats tested and is likely to be an XRV currently circulating among Australian bats.

Phylogenetic analysis reveals a close relationship between koala, gibbon, and bat gammaretroviruses

To determine the evolutionary relationships among the retroviruses we identified here (Supplementary Table 1) with known gammaretroviruses (Supplementary Table 2), we performed a phylogenetic analysis of the full retroviral genome. Our maximum likelihood phylogenetic analysis (Figure 2) revealed that HPG, MmGRV, and SaGRV formed a distinct and well-supported clade that is basal to the KoRV and GALV groups. In contrast, the Asian bat derived HlGRV and RhGRV cluster as a sister-group to the GALV clade. This analysis is supported by phylogenetic analyses of the individual *pol* and *env* genes, which reveal the same branching pattern. While analysis of the *gag* gene resulted in a slightly different branching pattern, this is likely as a result of low phylogenetic resolution, as indicated by low bootstrap support for this individual tree (Supplementary Figure). Hence, these data reveal that KoRV-related gammaretroviruses exist within multiple species of Australian and Asian bats, with those from Australia (HPG, FFRV1, MmGRV, SaGRV) phylogenetically distinct from those from Asia. Although the presence of diverse and basal gammaretroviruses in bats suggests they are a key reservoir species and may have transmitted viruses to other mammals, it is striking that those viruses sampled from bats (and other mammals) do not share close common ancestry with KoRV in koalas.

HPG is reproduction-competent in human and bat cells *in vitro*

To assess the biological characteristics of KoRV related bat viruses, we chemically synthesized the proviral genome of HPG (Supplementary Figure 1B). Transfection of human 293T cells with a plasmid construct carrying the HPG provirus resulted in the generation and release of viral particles morphologically similar to ecotropic Moloney murine leukemia virus (M-MLV), as determined by electron microscopy (Figure 3), in contrast to untransfected and mock transfected 293T cells (Supplementary Figure 8). These data are supported by virion associated reverse transcriptase analysis of the samples analyzed by electron microscopy (Supplementary Figure). Measurements of virion diameters indicate that HPG viral particles (mean ± SEM, 98.5 ± 2.5 nm) are smaller than M-MLV (130.8 ± 3.2 nm, p value < 0.001 by

Commented [PHD1]: Is this another Figure X-its not figure 8, on page 15

Mann-Whitney test; Supplementary Figure 4). To support our phylogenetic assessment that HPG is a gammaretrovirus, we performed a virion associated reverse transcriptase (RT) assay using HPG, M MLV, and HIV virions and including either manganese [Mn^{2+}; utilized by gammaretroviruses (25)] or magnesium [Mg^{2+}; utilized by lentiviruses (26)] as the cofactor for the RT DNA polymerase activity (Supplementary Figure 5). These data indicate that HPG RT shows a preference for Mn^{2+} compared to Mg^{2+}, typical of a gammaretrovirus (25). To determine if the HPG virions generated by proviral-plasmid transfection of human 293T cells were reproduction-competent (i.e. capable of establishing a productive infection within the context of a cell culture system *in vitro*), we performed a replication kinetics assay (Figure 4). We found that HPG was capable of entering and establishing a productive infection in human and bat cells, but not mouse cells (Figure 4). In contrast, the ecotropic M MLV infected the mouse cell line, but not human or bat cells (Figure 4). HPG was confirmed to be capable of establishing successive rounds of replication through a secondary infection assay (Supplementary Figure).

HPG displays a similar pattern of cell tropism as GALV and KoRV-A

~~To investigate the cell tropism mediated by the HPG envelope (Env) protein, we performed a viral entry assay in which retroviral particles were pseudotyped with the Env protein of several gammaretroviruses that have distinct tropism for human and mouse cells (Figure 5A). Our data show that HPG displays a similar pattern of cell tropism as GALV and KoRV-A, in that it is capable of entering human cells but not mouse cells. These data suggest that HPG likely utilizes the same cellular receptor, PiT-1 (SLC20A1), as GALV and KoRV-A (19, 27, 28).~~

Field Code Changed

~~To further investigate receptor usage by HPG, we performed a superinfection interference assay (Figure 5B). In this assay, human HeLa cells persistently infected with HPG became strongly resistant to superinfection with a reporter virus pseudotyped with the envelope proteins of KoRV-A, GALV, or HPG (97.8 - 98.6% reduction in infectivity). Infections with viral particles pseudotyped with dualtropic or amphotropic MLV Env were also moderately inhibited (respectively 34.5% and 47.1% reduction in infectivity). Dualtropic MLV uses both PiT-1 & Pit-2 (SLC20A2) cell receptors (29), while amphotropic MLV exclusively uses PiT-2 (30). In contrast, superinfection by particles pseudotyped with the unrelated vesicular stomatitis virus (VSV) envelope G protein was not restricted. These data indicate that HPG utilizes the PiT-1 and PiT-2 cell receptors for cell entry.~~

Commented [PHD2]: Depends how confident you are-I would say "strongly suggest that HPG utilizes ....."

An alignment of the receptor binding (RBD) domain (31) of HPG against other KoRV related viruses reveals numerous differences in the variable regions (VRA and VRB) within the RBD (Supplementary Figure 6). Within this region, the pathologically important CETTG motif within the RBD (32), that is conserved in all other bat KoRV related viruses, contains a threonine to serine mutation in HPG, resulting in a CETSG motif. HPG is more similar to GALV than to KoRV across both the VRA and VRB, where the RBD amino acid identities for HPG compared to GALV and KoRV are 66% and 62%, respectively. However, all of the KoRV related bat gammaretroviruses analyzed contain a large insertion within the VRB of 10 and 16 amino

acids, respectively relative to GALV and KoRV. Amphotropic and dualtropic MLV also contain several insertions within the VRB, increasing the length of their VRB region by 17 and 23 amino acids, relative to GALV and KoRV. These insertions are not present within ecotropic M-MLV, which utilizes the mouse CAT1 (SLC7A1) cell receptor (33, 34). A complementary alignment of the binding motif within mammalian PiT-1 genes further supports this result as the binding sites within *P. alecto* and *P. vampyrus* PiT-1 share the permissive amino acid residues, which are distinct from the non-permissive motif within mouse PiT-1 (Supplementary Figure). Some gammaretroviruses that utilize PiT-1 for cell entry also utilize the related protein, PiT-2, and this has been attributed to subtle differences in the composition and length of amino acid sequences within the VRA and VRB regions of the viral Env protein (35, 36).

To investigate the cell tropism mediated by the HPG envelope (Env) protein, we performed a viral entry assay in which retroviral particles were pseudotyped with the Env protein of several gammaretroviruses that have distinct tropism for human and mouse cells (Figure 5A). Our data show that HPG displays a similar pattern of cell tropism as GALV and KoRV-A, in that it is capable of entering human cells but not mouse cells. These data suggest that HPG likely utilizes the same cellular receptor, PiT 1 (SLC20A1), as GALV and KoRV A (19, 27, 28).

Field Code Changed

To further investigate receptor usage by HPG, we performed a superinfection interference assay (Figure 5B). In this assay, human HeLa cells persistently infected with HPG became strongly resistant to superinfection with a reporter virus pseudotyped with the envelope proteins of KoRV-A, GALV, or HPG (97.8 - 98.6% reduction in infectivity). Infections with viral particles pseudotyped with dualtropic or amphotropic MLV Env were also moderately inhibited (respectively 34.5% and 47.1% reduction in infectivity). Dualtropic MLV uses both PiT-1 & Pit-2 (SLC20A2) cell receptors (29), while amphotropic MLV exclusively uses PiT 2 (30). In contrast, superinfection by particles pseudotyped with the unrelated vesicular stomatitis virus (VSV) envelope G protein was not restricted. These data indicate that HPG utilizes the PiT 1 and PiT-2 cell receptors for cell entry.

Commented [PHD3]: Depends how confident you are I would say "strongly suggest that HPG utilizes ....."

Taken together, these results indicate that HPG may share a similar host range as KoRV-A and GALV, with the caveat that the specific determinants of receptor usage and cell tropism for PiT-1 and PiT-2 are complex (35-37), and further investigation will be required to more accurately delineate the complete host range, ~~and~~ cell tropism and viral entry receptors/cofactors of HPG.

Australian bats have been exposed to HPG and closely related viruses

To assess Australian bats for exposure to HPG or KoRV-related viruses, we tested for the presence of bat antibodies reactive against the HPG Env protein. Bat sera (87 samples) were screened for the presence of antibodies reactive to the HPG Env trimer ectodomain (Glu^{38} - Ser^{603}) and a synthetic peptide of the HPG VRA region of Env in a solid phase enzyme immunoassay. Additional analysis of sera reactive to HPG VRA was conducted against VRA

Results

peptides from KORV, GALV and MLV VRA (Supplementary Figure). ~~To determine whether bat samples that were seropositive for HPG might be cross-reactive against the closely related gammaretroviruses, KoRV-A and GALV, or the more distantly related gammaretrovirus, MLV, we performed a peptide-binding analysis using short peptide sequences derived from the Variable Region A within the Env protein of these viruses.~~ We also tested for the presence of HPG-specific nucleic acid in bat fecal samples.

~~Bat sera (87 samples) were screened for the presence of antibodies reactive to the HPG VRA region of Env in a solid-phase enzyme immunoassay. Where available sample material permitted, analysis was additionally conducted on the full range of VRA peptides and the HPG Env trimer ectodomain (Glu^{38}-Ser^{603}) (Supplementary Figure).~~ A rabbit immune serum raised to the HPG Env trimer was used as a positive control and to determine cross-reactivity to KoRV, GALV and MLV peptides. The immune sera reacted strongly to HPG Env trimer and the HPG VRA peptide sequence but did not show reactivity to KoRV, GALV or MLV peptides (Supplementary table X) nor to an HCV peptides sequence encoding the antigenic region of glycoprotein E2 residues 409-422 (not shown) . In addition, immune serum raised to MLV reacted to the MLV peptides sequence but not to HPG, KORV or GALV VRA peptides sequences (Supplementary table X).

Of the 87 bat samples, 27 showed reactivity to the HPG Env trimer and of these 18 (20.7%) were reactive to the HPG VRA peptide [*P. alecto* (n = 16), *P. conspicillatus* (n = 1) , *Rhinolopus megaphyllus* (n = 1)]. Of the 18 HPG VRA positive sera, 4 showed additional reactivity to KoRV-A and 4 were additionally reactive to KoRV-A and GALV peptides. ~~All of these samples were also reactive to the HPG Env ectodomain.~~ One serum, *P. Alecto* #20, was more strongly reactive towards the GALV VRA peptide than HPG VRA, HPG env or KORV VRA peptide. ~~Nine (50%) of the HPG VRA positive samples were also cross-reactive for KoRV-A and GALV VRA peptides. However, only a single sample (#20, *P. alecto*) was more strongly cross-reactive, against the GALV VRA peptide. Two~~Three samples (#7 and #19 *P. alecto* and #8 P. alecto) were reactive against the KoRV-A and GALV VRA peptide, respectively, but not reactive against the HPG VRA peptide. Only a single bat (#27, *P. alecto*) demonstrated ~~cross reactivity~~weak reactivity to MLV, and this bat was also strongly reactive against the VRA peptides of HPG, KoRV-A, GALV, and HPG Env.

These results reveal that ~~28~~32% of bat samples were seropositive for HPG or KoRV-related protein sequences. Within the species P Alecto, 27% were seropositive to HPG only, and 21% were seropositive to HPG or KoRV related viruses. ~~They also indicate that while some cross-reactivity is observable, reactivity among the tested bat samples is almost entirely strongest against HPG, and cross reactivity generally does not extend to distant gammaretroviral relatives such as MLV.~~

A nucleic acid analysis by reverse transcriptase qPCR (RT-qPCR) was performed on 373 bat fecal samples using both 'broad' primers designed to amplify the HPG related pteropid viruses [HPG, SaGRV, MmGRV, FFRV1 (Figure 2)], and 'specific' primers designed to amplify

Commented [PHD4]: Ive deleted the word cross-reactive as this implies there is cross reactivity in the assay and calls everything into question. I will add something into the discussion to address this...

Commented [PHD5]: Ive deleted the word cross-reactive as this implies there is cross reactivity in the assay and calls everything into question. I will add something into the discussion to address this...

Commented [PHD6]: I counted all HPG env positive plus korv/galv positives=28/87

Commented [PHD7]: I deleted the next sentence. I'll get back to that question in the discussion. Reason is, IDK that you can say this as the HPG env rabbit serum did not cross react and it was by far the strongest of all the sera tested. In additional your PCR validates the EL:ISA results to an extent, although it would be excellent if some of the positive serology results matched the PCR +ves.

Commented [PHD8]:

only HPG (Supplementary Table 5). Notably, the HPG-specific forward primer binds to a site that is not present (has been lost through a deletion event) within the closely related ERV FFRV1 (14) and contains multiple nucleotide differences at the 3' end of the primer, compared with MmGRV and SaGRV. We first performed the qPCR assay in such a way that both DNA and RNA would be amplified. This was followed by a second qPCR assay, performed in the absence of reverse transcriptase, so that only DNA could be amplified, allowing us to discriminate between amplification from retroviral DNA and RNA. Notably this analysis does not discriminate between germline and somatic viral genomic DNA. The results of the first assay reveal that 57/373 samples (15.3%) contained HPG-related nucleic acid (either DNA or RNA) and that 25 of those 57 (6.7% of the total) contained HPG-specific nucleic acid (Supplementary Table 6). The second qPCR assay revealed that all 25 HPG-specific samples were amplified from RNA (Supplementary Table 6), suggesting active infection with HPG. Interestingly, only 13 of the 25 samples were positive for HPG-related RNA. These data indicate that the remaining 12 of the 25 samples were positive for HPG-specific RNA, but not 'broad' HPG related RNA, and were instead positive for 'broad' HPG-related DNA. While we cannot rule out that some or all of these samples may have contained a quantity of HPG related RNA below the limit of detection of this assay, these data suggests that 12 bat samples were actively infected with HPG and were either latently infected with other HPG-related virus(es) or contain endogenous HPG-related sequences. Of the 57 samples positive for HPG related nucleic acid, 32 were positive only for HPG-related DNA, suggesting evidence of endogenization or latent infection with HPG-related viruses. Taken together, these serological and PCR results indicate that HPG and closely related viruses have infected multiple individuals across several species of Australian pteropid bats.

Commented [PHD9]: Are any of these the same as the sera I tested?

Discussion

To determine whether KoRV related viruses are present in Australian bats, we collected samples from bats on the east coast of Australia. Metagenomic analyses of these samples revealed the presence of three KoRV-related viruses, HPG, MmGRV, and SaGRV, from the pteropid bat species *P. alecto*, *M. minimus*, and *S. australis*. Searching the public SRA (38) also revealed two additional KoRV-related viruses, HlGRV and RhGRV, from the Asian microbat species *H. larvatus* and *R. hipposideros*. Phylogenetic analysis of the complete viral genome (Figure 2) revealed that the microbat viruses (HlGRV and RhGRV) clustered within a broad GALV/Woolly monkey virus (WMV) clade, while the pteropid viruses (HPG, FFRV1, MmGRV, SaGRV) form a more divergent clade that is basal to the KoRV and GALV/WMV clades. There is overlap between the habitats of all of the aforementioned bats except *R. hipposideros* (which ranges between Europe and West Asia) (22). Hence, bat communities could in theory provide a route of transmission for KoRV related viruses between Asia and Australia, although the immediate ancestor of KoRV remains uncertain and it is clear that additional animal species need to be sampled. Indeed, there are likely to be other currently unidentified species infected with KoRV related viruses linking the habitats of *R. hipposideros* and Australian bats. The long phylogenetic branch length linking the KoRV clade to its closest known relatives in the GALV/WMV clade indicates that the phylogenetic picture remains incomplete, with additional as yet unknown viruses and host species existing between the KoRV and GALV/WMV lineages of gammaretroviruses.

Other non bat species, particularly rodents, have been suggested as intermediary hosts for the transmission of KoRV-related viruses between Asia and Australia (20, 21). Of particular note is *Melomys burtoni*, an Australian rodent. Short nucleotide sequences representing KoRV related viruses, including the Melomys burtoni retrovirus (MbRV) and the Melomys woolly monkey virus (MelWMV), have been identified in *M. burtoni* (20, 39), both of which cluster closely with the WMV within the GALV clade and hence are no closer to KoRV than the bat viruses identified here [Figure 2; (20, 39)] (sequences of these viruses were omitted from our phylogenetic analysis due to insufficient genome sequence coverage). However, because the habitat of *M. burtoni* does not extend past the Wallace line or overlap with the habitat of gibbons (19, 22), this species is unlikely to be responsible for the direct transmission of KoRV related viruses between Australia and Asia.

KoRV and GALV utilize the PiT-1 receptor for cell entry (19, 27, 28). This receptor is almost ubiquitously expressed throughout the mammalian body at variable levels (40-43), and is highly expressed in many tissues including the colon, breast, testes, bladder, placenta, and brain (40, 41). KoRV and GALV have been detected in numerous tissues and body fluids including blood, sperm, breast milk, feces, and urine (5, 27, 44-49). Given the wide distribution of PiT-1 expression and the detection of KoRV and GALV in body fluids including blood, urine, and feces, it is possible that interspecies transmission might occur along routes

Commented [PHD10]: You should mention conservation of sequence as well

including blood during fighting/predation, and contamination of food sources by feces and urine.

We searched carefully for the presence of HPG in the genomes of *P. alecto* and *P. vampyrus* using molecular analyses, and more broadly for KoRV-related viruses in the SRA and were unable to detect these viral sequences in the genome of any bat species whose genome is currently available. While these data suggest that bat KoRV-related viruses are not endogenous, we cannot rule out the possibility as we have only sampled a small proportion of bats within each species. In this regard, KoRV endogenization in koalas is relatively recent, and accordingly is not represented across the entire koala genepool (44); existing in both endogenous and exogenous forms (5, 27, 50). Thus, given that HPG-specific sequences have been identified across several bat species, either HPG is an exogenous virus or it is undergoing endogenization in real time. A possible example of the latter is FFRV1 (14), which was recently discovered in the brain tissue of a *P. alecto* bat, but which we were not able to identify within the genome of *P. alecto* or other bats. Serological and nucleic acid analyses revealed that numerous individual bats across several species have been exposed to HPG and HPG-related viruses (Supplementary Tables 3, 4, and 6), with 32% of bats tested being seropositive to HPG or KoRV related viruses, and that 6.7% (25/373) of analyzed bat scat samples contained HPG-specific RNA, indicating that these bats are actively infected with HPG, and that more generally, HPG-related viruses are currently circulating among the communities of multiple species of Australian pteropid bats. The close evolutionary relationship between the bat KoRV-related viruses from several species of Australasian pteropid bats, Asian microbats, and the gibbon, koala, and rodent viruses suggest that bat populations in Australia and Asia play an important role in the transmission of KoRV-related viruses between bats and possibly other mammals.

Formatted: Font color: Red

The genome of HPG is typical of gammaretroviruses (Figure 1), and while HPG virions are morphologically similar to M MLV virions, possessing a spherical, electron dense core (Figure 3), measurements of HPG virion diameter reveal that it is smaller than M-MLV virions ($P < 0.001$; Supplementary Figure 4). This may be attributed to the smaller diameter of the viral core ($P < 0.001$). These data may indicate a difference in the quaternary structure of the HPG capsid compared to that of M MLV resulting in a more compact structure.

We generated HPG virions from a synthetic proviral expression construct to assess the reproduction capacity of HPG in cell culture. These HPG virions were capable of infecting human and bat cell lines (Figure 4), but not a mouse cell line, as shown by the production of new virions and their release into the cell culture supernatant over the course of several days. It is important to note while the complete genome of HPG was assembled from RNA extracted from a single bat, and virions generated from this sequence are reproduction competent *in vitro*, the infectious molecular clone was engineered from the consensus sequence of the assembled reads. Accordingly, the HPG molecular clone represents the average of the HPG population contained in the extracted RNA, rather than the exact sequence of a single viral

isolate. Similar consideration should be given to the other bat KoRV-related viruses reported here, which also represent the consensus of assembled sequence data.

An assessment of the cell tropism of HPG revealed that HPG Env-pseudotyped retroviral particles were able to enter human but not mouse cells (Figure 5A). GALV and KoRV-A are similarly restricted from entering mouse NIH 3T3 cells due to their use of the PiT-1 cellular receptor for viral entry (51). This inhibition is attributed to mouse PiT-1 containing differences in the binding site of GALV and KoRV-A (52, 53) which are not present in *P. alecto* or *P. vampyrus* PiT 1 (Supplementary Figure 7).

Infection of cells with a retrovirus can restrict the subsequent superinfection by viruses that use the same receptor by various mechanisms including downregulation of the receptor, and blocking the binding site on the cell receptor, preventing penetration or adsorption of the virus (54, 55). This method has been used to demonstrate the shared use of the PiT-1 receptor between KoRV-A and GALV (56). We undertook a superinfection interference assay which demonstrated that infection with HPG restricts superinfection by a reporter virus pseudotyped with the envelope protein of KoRV-A, GALV, amphotropic MLV, and dualtropic MLV. KoRV-A and GALV utilize the PiT-1 receptor (19, 27, 28), while amphotropic MLV utilizes PiT 2 (57), and dualtropic MLV utilizes both PiT-1 and PiT 2 (29). These results indicate that HPG most probably utilizes the PiT-1 and PiT-2 receptors for cell entry.

Commented [PHD11]: Up to you, but I am always less confident...

Variations in receptor usage can occur between closely related gammaretroviruses. KoRV-B, for example, while closely related to KoRV-A, utilizes the THTR1 receptor (58), which may be the result of a recombination event within the RBD between an ancestral KoRV and an unknown retrovirus (27). This is particularly important to consider in light of the alignment of the RBD of HPG and other bat KoRV related viruses (Supplementary Figure 6), which reveals a large insertion within the hypervariable VRB region. Amphotropic and dualtropic MLV similarly contain a large insertion within the VRB relative to KoRV, GALV, and ecotropic M-MLV. The VRB region of amphotropic MLV is essential for interaction with the PiT-2 cell receptor (59), and the large insertion within the VRB of HPG may be involved in its apparent use of the PiT-2 receptor, demonstrated by the superinfection assay.

Interestingly, HPG contains a modification within the CETTG motif within the RBD (Supplementary Figure 6) that is important for viral pathogenicity (32). Mutations within the CETTG attenuate viral pathogenicity *in vitro,* as is the case for KoRV A which possesses a CETAG motif (60). HPG contains a CETSG motif (Supplementary Figure 6), which is also found in 27% of KoRV-D proviruses and is hypothesized to attenuate syncytia formation related pathogenicity (60). However, other bat KoRV-related viruses analyzed in this study possess the pathogenic CETTG motif. The identification of bats as a source of infectious retroviruses related to KoRV and GALV implicates bats as a reservoir of KoRV-related viruses that can potentially be transmitted between Australia and Asia to other mammalian species.

Materials and Methods

Supplementary figures, tables, and details of the materials and methods used in this study, including all experimental procedures are provided in SI Appendix.

Acknowledgments

We thank the NIH AIDS Research and Reference Reagent program, Division of AIDS, NIAID, NIH for HeLa cells, Damian Purcell for providing NIH 3T3 cells, and Richard Axel for providing HEK 293T cells. We thank Amy Burroughs, Bronwyn Clayton, Huajun Zhang, Jennifer Barr, Johanna Dups, Kate Baker, Peng Zhou, Shawn Todd, Hugh Spencer, and Andreas Kurth for their roles in the collection of bat samples. We thank Vicky Boyd for her assistance with the serological analysis, and Reuben Klein for extracting viral nucleic acids from bat samples. This work was supported by the National Health and Medical Research Council (NHMRC), grant number GNT1121077, awarded to G.T., M.L.B. and G.A.M. and NHMRC Senior Research Fellowship GNT1117748 to G.T. J.A.H. was funded by NHMRC GNT1121077. H.E.D. is supported by the NHMRC grants GNT1041897, GNT1146082. E.C.H. is supported by an ARC Australian Laureate Fellowship FL170100022. L-F.W. is funded by the Singapore National Research Foundation grants NRF2012NRF CRP001 056 and NRF2016NRF-NSFC002-013. We wish to acknowledge Microscopy Australia and NCRIS for funding the CSIRO electron microscopy capability utilized in this study. The authors gratefully acknowledge the contribution to this work of the Victorian Operational Infrastructure Support Program received by the Burnet Institute.

Author's Contributions

M.T. discovered the Hervey pteropid gammaretrovirus and devised the pipeline that enabled amplification and construction of the full-length HPG sequence; J.A.H. identified the Asian bat gammaretroviruses; P.S. identified the pteropid bat gammaretroviruses MmGRV and SaGRV; J.A.H., M.T., C.K., P.S., L F.W., I.S., and G.T. designed the study; M.T., J.A.H., C.K., A.J., M.D., B.J., C.L., P.S., G.H., A.N., C.A.S., A.W., D.K., H.E.D., P.P., G.A.M., E.C.H., and I.S. performed research and analyzed data; J.A.H., M.T., C.K., A.J., M.D., P.S., G.H., A.N., C.A.S., A.W., D.K., H.E.D., P.P., G.A.M., M.L.B., E.C.H., L-F.W., I.S., and G.T. discussed and interpreted the results; J.A.H. wrote the first draft of the paper; J.A.H., M.T., C.K, A.J., M.D., C.A.S., H.E.D., P.P., M.L.B., E.C.H., L-F.W., I.S., and G.T. critically edited the manuscript.

References

1. Vogt VM (1997) Retroviral virions and genomes. *Retroviruses*, eds Coffin JM, Hughes SH, & Varmus HE (Cold Spring Harbor Laboratory Press, New York, USA).
2. Gao F, *et al.* (1999) Origin of HIV 1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* 397(6718):436.
3. Tarlinton R, Meers J, Hanger J, & Young P (2005) Real-time reverse transcriptase PCR for the endogenous koala retrovirus reveals an association between plasma viral load and neoplastic disease in koalas. *J Gen Virol* 86(3):783-787.
4. Ito N, *et al.* (2019) Clinical and pathological characteristics of acute myelogenous leukemia in a female koala with diabetes mellitus. *J Vet Med Sci*:19-0006.
5. Tarlinton RE, Meers J, & Young PR (2006) Retroviral invasion of the koala genome. *Nature* 442(7098):79.
6. Stoye JP (2012) Studies of endogenous retroviruses reveal a continuing evolutionary saga. *Nat Rev Micro* 10(6):395 406.
7. Herniou E, *et al.* (1998) Retroviral diversity and distribution in vertebrates. *J Virol* 72(7):5955 5966.
8. Boeke JD & Symer DE (2010) An everlasting war dance between retrotransposons and their metazoan hosts. *Retroviruses*, eds Kurth R & Bannert N (Caister Academic Press, Berlin), pp 1 34.
9. Griffiths DJ (2001) Endogenous retroviruses in the human genome sequence. *Genome Biol* 2(6):reviews1017. 1011.
10. Hayman DT (2016) Bats as viral reservoirs. *Annu Rev Virol* 3:77 99.
11. Hayward JA, *et al.* (2013) Identification of diverse full-length endogenous betaretroviruses in megabats and microbats. *Retrovirology* 10(1):35.
12. Cui J, *et al.* (2012) Identification of diverse groups of endogenous gammaretroviruses in mega and microbats. *J Gen Virol* 93(9):2037 2045.
13. Farkašová H, *et al.* (2017) Discovery of an endogenous deltaretrovirus in the genome of long-fingered bats (chiroptera: Miniopteridae). *Proc Natl Acad Sci U S A* 114(12):3145 3150.
14. McMichael L, *et al.* (2019) A novel australian flying-fox retrovirus shares an evolutionary ancestor with koala, gibbon and Melomys gamma-retroviruses. *Virus Genes* 55(3):421 424.
15. Cui J, Tachedjian G, & Wang L F (2015) Bats and rodents shape mammalian retroviral phylogeny. *Sci Rep* 5:16561.
16. Hayward A, Grabherr M, & Jern P (2013) Broad-scale phylogenomics provides insights into retrovirus–host evolution. *Proc Natl Acad Sci U S A* 110(50):20146-20151.
17. Hayward JA, *et al.* (2018) Differential evolution of antiretroviral restriction factors in pteropid bats as revealed by APOBEC3 gene complexity. *Mol Biol Evol* 35(7):1626-1637.
18. Riddle BR & Hafner DJ (2010) Integrating pattern with process at biogeographic boundaries: The legacy of wallace. *Ecography* 33(2):321 325.
19. Denner J (2016) Transspecies transmission of gammaretroviruses and the origin of the gibbon ape leukaemia virus (galv) and the koala retrovirus (KoRV). *Viruses* 8(12):336.
20. Simmons G, Clarke D, McKee J, Young P, & Meers J (2014) Discovery of a novel retrovirus sequence in an australian native rodent (Melomys burtoni): A putative link between gibbon ape leukemia virus and koala retrovirus. *PLoS One* 9(9):e106954.

References

21. Denner J & Young PR (2013) Koala retroviruses: Characterization and impact on the life of koalas. *Retrovirology* 10(1):108.
22. IUCN (2016) The IUCN red list of threatened species. Version 2016 1.
23. Cui J, *et al.* (2012) Discovery of retroviral homologs in bats: Implications for the origin of mammalian gammaretroviruses. *J Virol* 86(8):4288 4293.
24. Crameri G, *et al.* (2009) Establishment, immortalisation and characterisation of pteropid bat cell lines. *PLoS One* 4(12):e8266.
25. Yaniv A, *et al.* (1979) Biochemical characterization of the type c retrovirus associated with lymphoproliferative disease of turkeys. *J Virol* 30(1):351-357.
26. Billamboz M, *et al.* (2011) Magnesium chelating 2 hydroxyisoquinoline 1, 3 (2 h, 4 h) diones, as inhibitors of HIV 1 integrase and/or the HIV-1 reverse transcriptase ribonuclease h domain: Discovery of a novel selective inhibitor of the ribonuclease h function. *J Med Chem* 54(6):1812 1824.
27. Xu W, *et al.* (2013) An exogenous retrovirus isolated from koalas with malignant neoplasias in a us zoo. *Proc Natl Acad Sci U S A* 110(28):11547 11552.
28. O'Hara B, *et al.* (1990) Characterization of a human gene conferring sensitivity to infection by gibbon ape leukemia virus. *Cell Growth Differ* 1(3):119 127.
29. Miller AD & Chen F (1996) Retrovirus packaging cells based on 10a1 murine leukemia virus for production of vectors that use multiple receptors for cell entry. *J Virol* 70(8):5564 5571.
30. Feldman SA, Farrell KB, Murthy RK, Russ JL, & Eiden MV (2004) Identification of an extracellular domain within the human PiT2 receptor that is required for amphotropic murine leukemia virus binding. *J Virol* 78(2):595.
31. Fass D, *et al.* (1997) Structure of a murine leukemia virus receptor-binding glycoprotein at 2.0 angstrom resolution. *Science* 277(5332):1662 1666.
32. Oliveira NM, Satija H, Kouwenhoven IA, & Eiden MV (2007) Changes in viral protein function that accompany retroviral endogenization. *Proc Natl Acad Sci U S A* 104(44):17506-17511.
33. Albritton LM, Tseng L, Scadden D, & Cunningham JM (1989) A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane spanning protein and confers susceptibility to virus infection. *Cell* 57(4):659-666.
34. Yoshimoto T, Yoshimoto E, & Meruelo D (1993) Identification of amino acid residues critical for infection with ecotropic murine leukemia retrovirus. *J Virol* 67(3):1310-1314.
35. Boomer S, Eiden M, Burns CC, & Overbaugh J (1997) Three distinct envelope domains, variably present in subgroup b feline leukemia virus recombinants, mediate pit1 and pit2 receptor recognition. *J Virol* 71(11):8116 8123.
36. Sugai J, *et al.* (2001) Identification of envelope determinants of feline leukemia virus subgroup b that permit infection and gene transfer to cells expressing human pit1 or pit2. *J Virol* 75(15):6841 6849.
37. Han J Y, Zhao Y, Anderson WF, & Cannon PM (1998) Role of variable regions a and b in receptor binding domain of amphotropic murine leukemia virus envelope protein. *J Virol* 72(11):9101-9108.
38. Wu Z, *et al.* (2012) Virome analysis for identification of novel mammalian viruses in bat species from chinese provinces. *J Virol* 86(20):10999-11012.

References

39. Alfano N, *et al.* (2016) An endogenous gibbon ape leukemia virus (GALV) identified in a rodent (Melomys burtoni subsp.) from Wallacea (Indonesia). *J Virol*:JVI. 00723-00716.
40. Petryszak R, *et al.* (2015) Expression atlas update—an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res* 44(D1):D746 D752.
41. Bastian F, *et al.* (2008) Bgee: Integrating and comparing heterogeneous transcriptome data among species.), pp 124 131.
42. Johann S, Gibbons J, & O'hara B (1992) Glvr1, a receptor for gibbon ape leukemia virus, is homologous to a phosphate permease of neurospora crassa and is expressed at high levels in the brain and thymus. *J Virol* 66(3):1635 1640.
43. Kavanaugh MP, *et al.* (1994) Cell surface receptors for gibbon ape leukemia virus and amphotropic murine retrovirus are inducible sodium dependent phosphate symporters. *Proc Natl Acad Sci U S A* 91(15):7071 7075.
44. Simmons G, *et al.* (2012) Prevalence of koala retrovirus in geographically diverse populations in a ustralia. *Aust Vet J* 90(10):404 409.
45. Waugh CA, *et al.* (2017) Infection with koala retrovirus subgroup b (KoRV-b), but not KoRV-a, is associated with chlamydial disease in free ranging koalas (phascolarctos cinereus). *Sci Rep* 7(1):134.
46. Morris KM, *et al.* (2016) Characterisation of the immune compounds in koala milk using a combined transcriptomic and proteomic approach. *Sci Rep* 6:35011.
47. Wedrowicz F, Saxton T, Mosse J, Wright W, & Hogan FE (2016) A non-invasive tool for assessing pathogen prevalence in koala (phascolarctos cinereus) populations: Detection of chlamydia pecorum and koala retrovirus (KoRV) DNA in genetic material sourced from scats. *Conservation Genetics Resources* 8(4):511-521.
48. Kawakami TG, Sun L, & McDOWELL TS (1977) Infectious primate type-c virus shed by healthy gibbons. *Nature* 268(5619):448.
49. Kawakami TG, Sun L, & McDowell TS (1978) Natural transmission of gibbon leukemia virus. *J Natl Cancer Inst* 61(4):1113-1115.
50. Miyazawa T, Shojima T, Yoshikawa R, & Ohata T (2011) Isolation of koala retroviruses from koalas in japan. *J Vet Med Sci* 73(1):65 70.
51. Blair DG, *et al.* (1990) Characterization of a human gene conferring sensitivity to infection by gibbon ape leukemia virus'. *Cell Growth Differ* 1(3):119 127.
52. Schneiderman RD, Farrell KB, Wilson CA, & Eiden MV (1996) The japanese feral mouse pit1 and pit2 homologs lack an acidic residue at position 550 but still function as gibbon ape leukemia virus receptors: Implications for virus binding motif. *J Virol* 70(10):6982 6986.
53. 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.
54. Steck FT & Rubin H (1966) The mechanism of interference between an avion leukosis virus and rous sarcoma virus: Ii. Early steps of infection by rsv of cells under conditions of interference. *Virology* 29(4):642-653.
55. Nethe M, Berkhout B, & Van der Kuyl AC (2005) Retroviral superinfection resistance. *Retrovirology* 2(1):52.
56. Oliveira NM, Farrell KB, & Eiden MV (2006) In vitro characterization of a koala retrovirus. *J Virol* 80(6):3104-3107.

References

57. Feldman SA, Farrell KB, Murthy RK, Russ JL, & Eiden MV (2004) Identification of an extracellular domain within the human PiT2 receptor that is required for amphotropic murine leukemia virus binding. *J Virol* 78(2):595-602.
58. Shojima T, *et al.* (2013) Identification of a novel subgroup of koala retrovirus from koalas in japanese zoos. *J Virol* 87(17):9943-9948.
59. Tailor CS & Kabat D (1997) Variable regions a and b in the envelope glycoproteins of feline leukemia virus subgroup b and amphotropic murine leukemia virus interact with discrete receptor domains. *J Virol* 71(12):9383-9391.
60. Quigley BL, *et al.* (2019) Changes in endogenous and exogenous koala retrovirus (KoRV) subtype expression over time reflects koala health outcomes. *J Virol*:JVI. 00849-00819.
61. Gelderblom HR (1997) Fine structure of HIV and SIV.
62. Lefkowitz EJ, *et al.* (2017) Virus taxonomy: The database of the international committee on taxonomy of viruses (ICTV). *Nucleic Acids Res* 46(D1):D708-D717.
63. Bacharach E, Gonsky J, Lim D, & Goff SP (2000) Deletion of a short, untranslated region adjacent to the polypurine tract in moloney murine leukemia virus leads to formation of aberrant 5' plus-strand DNA ends in vivo. *J Virol* 74(10):4755 4764.

Figures

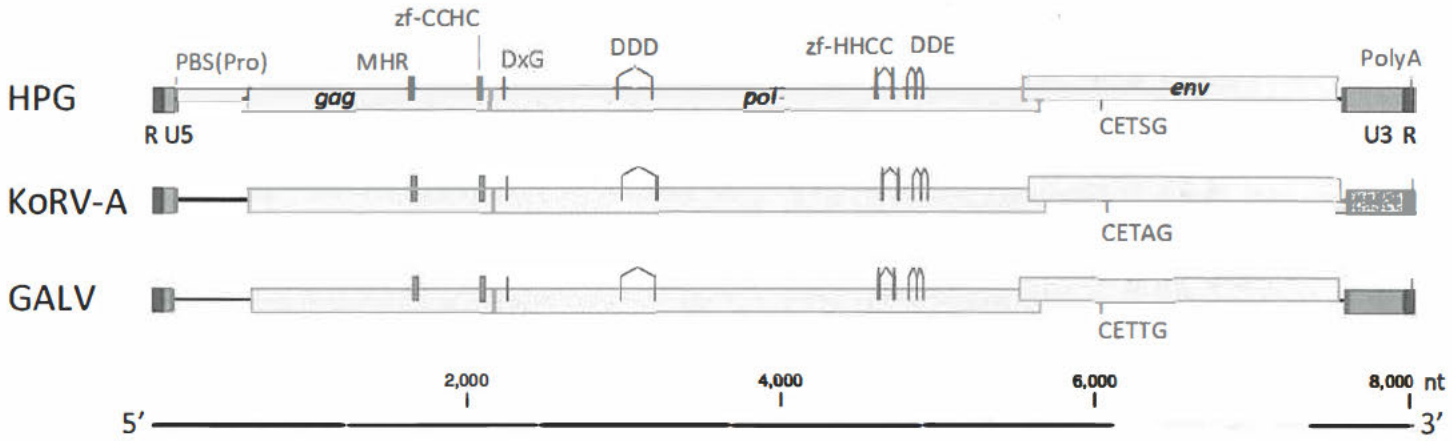


Figure 1. *The genome of the Hervey pteropid gammaretrovirus (HPG) contains conserved functional motifs and is analogous to KoRV-A and GALV.* R, terminal repeat sequence; U5/U3, unique 5'/3' region; PBS(Pro), proline tRNA primer binding site; *gag*, group-specific antigen; MHR, major homology region; zf, zinc finger; DxG, protease active site motif; DDD, reverse transcriptase active site motif; *pol*, polymerase; DDE, integrase active site motif; *env*, envelope; CET(S/A/T)G, pathogenicity motif; PolyA, polyadenylation signal.

Figures

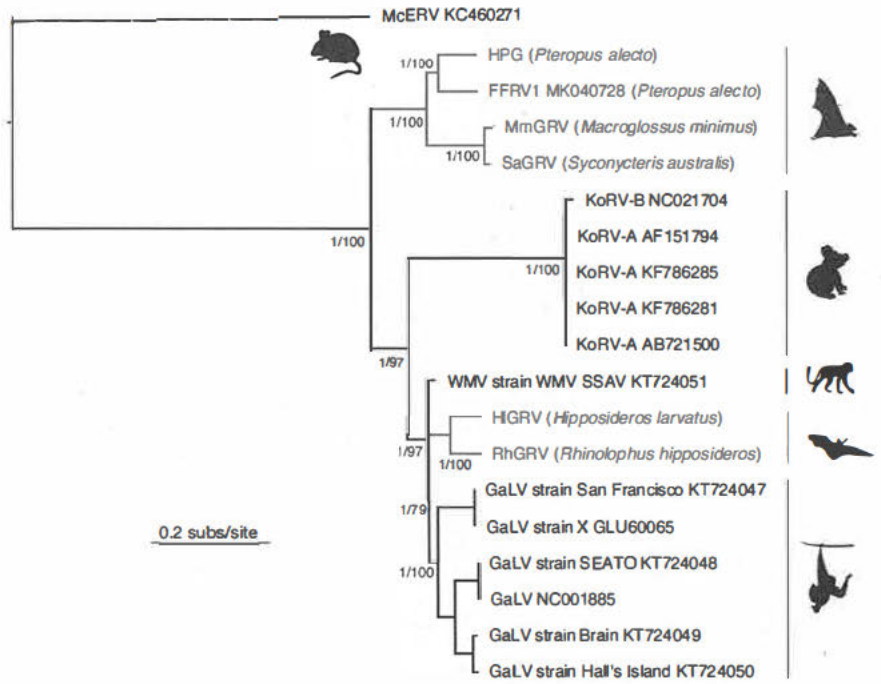


Figure 2. *Evolutionary relationships among KoRV related viruses.* Maximum likelihood phylogeny of the complete (nucleotide) sequence genome of 19 gammaretroviruses. All branches are scaled according to the number of nucleotide substitutions per site, and branches representing bat retroviruses are shown in red. Support for key nodes on the phylogeny are shown in the form SH-like branch support/bootstrap support. Silhouettes represent the host species; top left, mice; left (in descending order), pteropid bats, koalas, woolly monkeys, microbats, gibbons. The tree was rooted using the McERV (Mus caroli endogenous retrovirus) KC460271 sequence. HPG, Hervey pteropid gammaretrovirus; FFRV1, flying fox retrovirus; MmGRV, Macroglossus minimus gammaretrovirus; SaGRV, Syconycteris australis gammaretrovirus; KoRV, Koala retrovirus; WMV, Woolly monkey virus; HIGRV, Hipposideros larvatus gammaretrovirus; RhGRV, Rhinolophus hipposideros gammaretrovirus; GaLV, Gibbon ape leukemia virus.

Figures

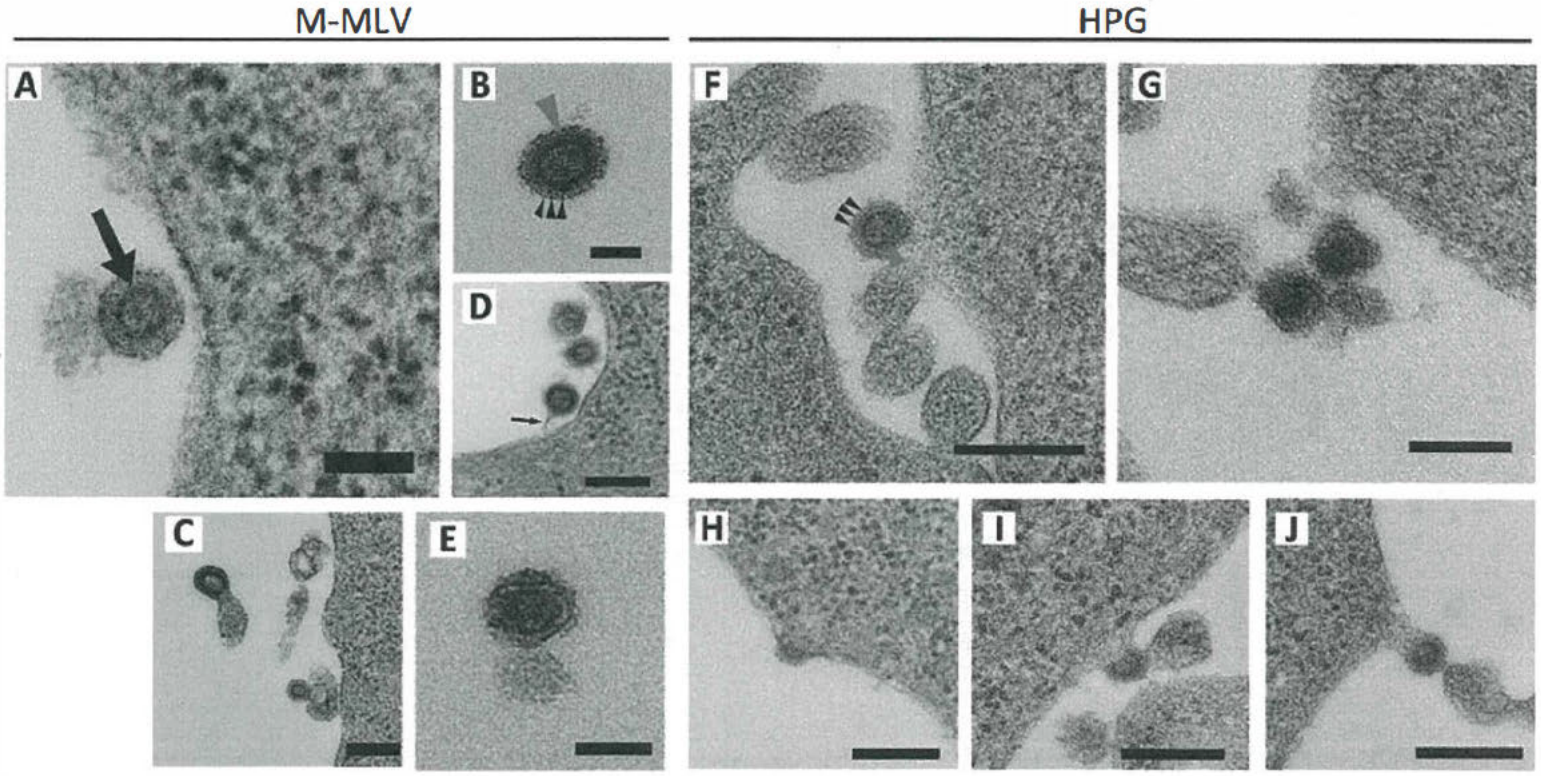


Figure 3. *Electron micrographs (EM) of Moloney murine leukemia virus (M-MLV) and Hervey pteropid gammaretrovirus (HPG).*

Figures

(A) An extracellular, roughly spherical, enveloped virus-like particle with a concentric icosahedral core (arrow). The cores have variable electron translucence, from lucent to dense, indicating variable stages of particle maturation. (B) An immature extracellular virus-like particle with tooth-like appearance of the viral envelope (red arrow) surrounding the double-layered shell of the core. Distinct banding can also be seen in the envelope of the particle (black arrow heads). (C) Virus-like particle exiting the cell demonstrating a Type C budding profile, characteristic of viruses belonging to the genus *Gammaretrovirus* (61, 62). (D) Evidence of virus assembly and budding from the plasma membrane of the cell, including the presence of a tether-like structure connecting the cell membrane to the newly budded virus (black arrow). (E) A mature virus-like particle with an electron dense core encapsulated in an envelope. HPG: (F) Immature virus-like particle exhibiting tooth-like appearance of the viral envelope (black arrows) surrounding the double-layered shell of the core (red arrow). (G) A mature virus-like particle with an electron dense core encapsulated in an envelope. (H-J) Evidence of virus-like particle assembly, and budding from the plasma membrane of the cell. Budding begins with electron dense material forming under the membrane (H), which progresses until the nascent virus-like particle pushes out from the membrane and is pinched off to form a free particle. Scale bars represent (A) 50nm (B) 100nm (C) 250nm (D) 200nm (E) 100nm (F) 200nm (G) 200nm (H-J) 250nm. Negative transfection controls were untransfected cells and cells mock transfected with the empty vector pcDNA3.1. These controls were not observed to contain or produce viral particles (Supplementary Figure).

Figures

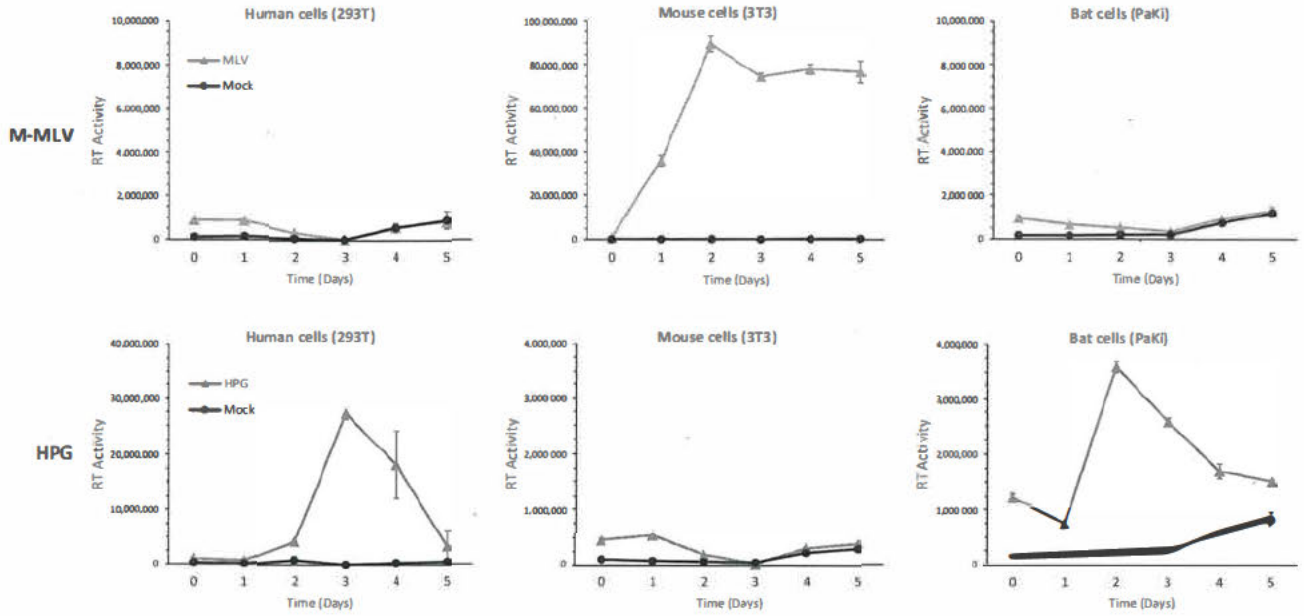


Figure 4. *Replication kinetics of Moloney murine leukemia virus (M-MLV) and Hervey pteropid gammaretrovirus (HPG) in human, mouse, and bat cells.* M-MLV and HPG virions were generated by transfection of human 293T cells with pNCS (63) and pCC1-HPG retroviral expression plasmids, respectively, and used to infect human 293T, mouse 3T3, and bat PaKi cell lines. Culture supernatants were collected daily for five days and assessed for the presence of virus by measuring virion reverse transcriptase (RT) activity. Error bars represent the standard error of the mean (SE, n = 6).

Figures

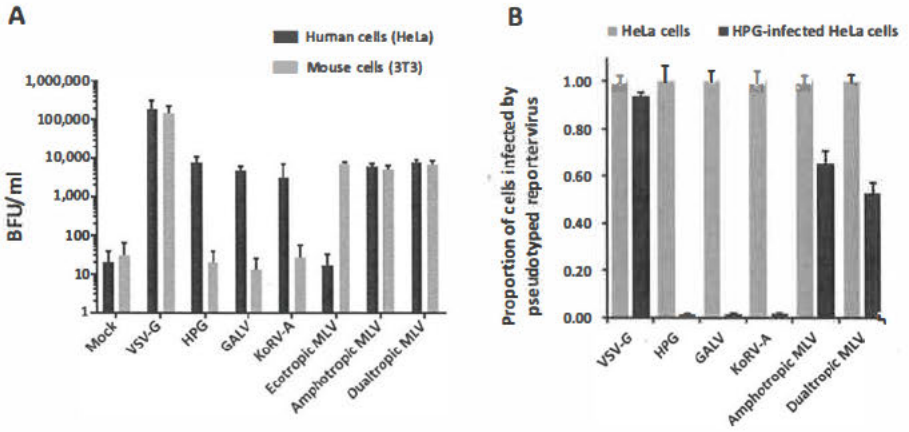


Figure 5. *Human and mouse cell tropism of pseudotyped gammaretroviral virus like particles and HPG induced resistance to cross-infection.* (A) Gammaretrovirus Envelope- or VSV-G pseudotyped retroviral particles, containing the lacZ reporter gene, were generated using the Retro-X packaging system. The infectivity of pseudotyped viral particles was determined in human HeLa cells and mouse 3T3 cells. Infected cells were quantified by counting blue cell forming units (BFU) following incubation with X gal (5-bromo-4-chloro-3-indolyl-β-D galactopyranoside). Uninfected cells were used as a control (Mock). (B) Human HeLa cells were persistently infected with HPG and then challenged with infection by a reporter virus pseudotyped with gammaretrovirus Envelope- or VSV-G-pseudotyped retroviral particles. Infected cells were quantified as for (A), and infection in persistently HPG-infected cells is expressed as a proportion of the amount of uninfected HeLa cells infected by the same reporter virus. VSV-G, Vesicular stomatitis virus G protein; HPG, Hervey pteropid gammaretrovirus; GALV, Gibbon ape leukemia virus; KoRV-A, Koala retrovirus A; MLV, Murine leukemia virus.

SI Appendix

Contents

Supplementary Methods

Ethics

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

Sample collection along east coast of Australia and Daintree Rainforest

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

Metagenomic analysis of viral nucleic acid in bat samples

For samples collected in the Daintree rainforest, total RNA was extracted with the QIAamp viral RNA Mini Kit (QIAGEN) and DNA was digested using the TURBO DNA free kit (Thermo Fisher Scientific), all according to the manufacturer's protocol. Other samples were processed as follows: Briefly, PBS homogenized bat feces was enriched for viral particles using a discontinuous sucrose gradient (2, 3). Total RNA was extracted with the QIAamp viral RNA Mini Kit (QIAGEN) except carrier RNA (poly A) was omitted from Buffer AVL and genomic DNA was removed with DNase I digestion prior to RNA extraction as previously described (2). Random RT PCR amplification and double stranded cDNA was prepared as previously described (2) except K8N random primers were replaced with (5' GTTTCCCAGTAGGTCTCNNN NNNNN 3') for cDNA synthesis and 5' A*G*C*A*C TGTAGGTTTCCCAGTAGGTCTC 3' for double stranded cDNA amplification (4). Sequencing libraries were generated using Illumina Nextera XT library construction, sequencing was performed on the Illumina MiSeq platform, and bioinformatics including FASTQ paired end read quality control and *de novo* assembly was performed as described previously (3). KoRV related viral contigs were identified by BLASTn and BLASTx analysis using the assembled contigs as query sequences against the NCBI nucleotide collection database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using default parameters.

Amplification and assembly of the Hervey pteropid gammaretrovirus (HPG) genome

Partial HPG sequences were initially identified in the metagenomic analysis of a *P. alecto* scat sample obtained in Hervey bay, in 2011, using the method described above in "Metagenomic analysis of viral nucleic acid in bat samples", employing random RT PCR amplification. To generate the complete HPG

genome sequence we used a modified single cell whole transcriptome amplification (WTA) procedure for detecting ultra low copy viral RNA, and *de novo* sequence assembly pipeline.

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

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

Dual indexed libraries were prepared according to the Illumina Nextera XT DNA Library Prep Kit (Illumina). Library concentration was determined with the Qubit HS dsDNA assay (Invitrogen). Library quality and distribution was determined by loading 3 ng of sample on an Agilent Technology 2100 Bioanalyzer using the Agilent High Sensitivity DNA assay. Libraries were normalized, denatured then diluted to a final concentration of 10 pM with HT1 buffer (Illumina) and spiked with 1% PhiX control library (Illumina). Libraries were sequenced on the Illumina MiSeq platform, using the MiSeq Reagent v2 kit (300 cycles), generating 150 bp paired end reads.

Illumina FASTQ paired end reads (8,162,956) were imported into CLC Genomics Workbench v10.1.1 using default Illumina import parameters, and then trimmed for size, quality, and ambiguous bases using default parameters except for the following: Quality Limit = 0.01, Ambiguous limit = 2 and Minimum number of nucleotides in reads = 30. Host reads (*Pteropus alecto* draft genome assembly GenBank assembly accession GCA_000325575.1 and mitochondrion Genbank accession NC_023122) were removed by read mapping using default settings on the CLC Genomics Grid Worker v7.0.1 except "Length" and "Similarity" fractions were both set to 0.9.

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

Identification and assembly of KoRV related viruses in publicly available databases

To identify KoRV related gammaretroviruses in public databases, data from the Sequence Read Archive (SRA) derived from bat RNA and DNA were subjected to SRA BLAST analysis (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 pteropid

gammaretrovirus (HPG) was used as the query sequence. The algorithm parameters set were to: Program = blastn, Max target sequences = 1000, Expect threshold < $1x10^{-10}$, word size = 11, match score = 2, mismatch cost = 3, gap costs = existence 5 extension 2, no filtering or masking. SRA that contained reads aligning to the query sequences were from the Chinese microbats *Rhinolophus hipposideros* (Genbank: SRX1059482 & SRX1059481) and *Hipposideros larvartus* (Genbank: SRX1059446). Sequencing reads aligning to the query sequences were downloaded and assembled into the partially complete genomes of RhGRV and HIGRV as follows: Reads were downloaded and assembled using the CLC Genomics Workbench 11.0 (QIAGEN, Aarhus, Denmark) "Assemble Sequences" tool into a contiguous consensus sequence using the following parameters: Minimum aligned read length = 20, alignment stringency = high, conflicts = Vote (A, C, G, T). Assembled contigs were subsequently used as a new query in an otherwise identical BLASTn search against the same SRA. This process was iteratively repeated until all contigs could be extended out until they overlapped with each other or reached a region of zero read coverage. The extended and overlapping contigs were assembled by alignment against the reference/query HPG genome sequence in CLC Genomics Workbench.

Annotation of Retroviral Genomes

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

HPG specific analysis of Pteropid genomes

To determine whether HPG could be identified as an endogenous retrovirus within the genomes of Pteropid bats, we performed *in vitro* and *in silico* analyses. For the *in silico* analysis, we performed a BLAST analysis using CLC Genomics Workbench, of the genomes of *Pteropus alecto* (Genbank: PRJNA232518) and *P. vampyrus* (Genbank: PRJNA275879) using the HPG genome as the query sequence, with the algorithm parameters: Expect threshold = $1x10^{-10}$; word size = 11; Low complexity regions filtered.

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

HPG positive controls were derived from two regions within HPG [1.34 & 1.55 kb in length (Supplementary Figure 3), which were identified in the NGS metagenomics analysis described above in "Metagenomic analysis of viral nucleic acid in bat samples". The two sequences were amplified by PCR with the QIAGEN HotStar HiFidelity polymerase according to the manufacture's protocol from random RT PCR amplified bat scat sample used for the initial metagenomics NGS using primers designed to amplify each sequence (Supplementary Table 5). The two amplicons were cloned into the pCR4 TOPO (Invitrogen) vector using the Zero Blunt TOPO PCR cloning kit (Invitrogen) according to the manufacturer's protocol. Primer sensitivity tests were conducted to determine the template copy

number required for a band to be present in this analysis using the 1.3 & 1.5 kb sequence primers (Supplementary Table 5) based on plasmid controls. Primers specific for the 1.3 & 1.5 kb HPG sequences were found to be sensitive to $1.4x10^8$ and $9.0x10^3$ copies, respectively.

Phylogenetic analysis

To determine the evolutionary relationships among KoRV related gammaretroviruses we performed phylogenetic analyses using aligned complete genome nucleotide sequences (Supplementary Table 2) and individual gene sequences. Accordingly, a multiple sequence alignment of 19 complete genomes was performed using a combination of MAFFT (8) and MUSCLE algorithms (9). Following alignment, regions of ambiguous and uncertain alignment were removed using Gblocks (10). For the complete genomes, this resulted in final alignment of 6,925 nt that was used to infer evolutionary relationships. Subsets of this alignment covering the *gag*, *pol*, and *env* gene regions were used for the individual gene analyses. Phylogenetic trees of these data were estimated using the maximum likelihood (ML) method available in the PhyML program (11), assuming a GTR model of nucleotide substitution with a proportion of invariant sites (I) and a gamma distribution of among site rate variation (Γ). To determine the robustness of each node a bootstrap resampling analysis (1,000 replications) was performed using the same nucleotide substitution model. For the complete genome tree (Figure 2), a Shimodaira Hasegawa (SH) test was conducted, providing additional nodal support. The *Mus caroli* ERV, McERV (Supplementary Table 2), sequence was used as an outgroup to root the tree.

Generation of HPG proviral sequence and synthesis of HPG proviral expression construct

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

Cell cultures

Human embryonic kidney (HEK) 293T cells (kindly provided by Richard Axel, Columbia University), human epithelial cervical adenocarcinoma (HeLa) cells (NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH), NIH Swiss mouse embryo (NIH/3T3) fibroblast ce ls (American Type Culture Collection), a *P. alecto* kidney (PaKi) cell line (7), and HEK cells that express M MLV Gag and Pol polyproteins (GP2 293 cells; Takara Bio) were utilized. 293T and HeLa cells were authenticated using the Promega GenePrint 10 system performed by the Australian Genome Research Facility (AGRF). The short tandem repeat (STR) profile was used to search the ATCC STR database https://www.atcc.org/en/STR_Database.aspx and the DSMZ German Collection of Microorganisms and Cell cultures database https://www.dsmz.de/services/human and animal cell lines/online str analysis. All cell cultures were maintained at 37°C, 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific) supplemented with heat inactivated fetal bovine serum (100mL/l; Invitrogen), glutamine (292 mg/mL; Invitrogen), and the antibiotics penicillin (100 units/mL; Invitrogen) and streptomycin (100 units/mL; Invitrogen), with the exception of the PaKi cells for which DMEM was substituted with DMEM/F 12 1:1 medium (Thermo Fisher Scientific).

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

HPG and M MLV viral particles were generated by transfection of 293T cells with the pCC1 HPG and pNCS (Addgene: 17362) plasmids, respectively. 293T cells were transfected at 50% confluency with 20 µg or pCC1 HPG or 10 µg of pNCS, using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. Untransfected cells and cells transfected with 20 µg of either the empty plasmid pCR2.1 (Thermo Fisher Scientific) or the empty plasmid pcDNA3.1 (Invitrogen), were used as controls. Transfected cells were incubated at 37°C, 5% CO_2 for 48 h, and then virion containing supernatants were collected and clarified by centrifugation at 200 x g for 5 min. Virus production was determined by quantifying virion associated RT activity, as previously described (12).

Generation of HeLa cells persistently infected with HPG

HPG virion containing supernatants were generated as previously described in, "Transfection of 293T cells for generation of HPG & M MLV viral particles", and used to infect HeLa cells. Cells were seeded at a density of $7x10^5$ cells per T25 tissue culture flask (BD Biosciences, Bedford MA). Once cells reached 50% confluency, media was replaced with a mix of 4 ml DMEM, 1 ml HPG virion containing supernatant and DEAE Dextran (Sigma Aldrich) at a final concentration of 10 µg/mL. Cells were incubated for 16 h at which point the supernatant was removed, cells were washed twice in PBS and 5 ml of fresh DMEM was added. At 48 hours post infection, cells were passaged at a concentration of 1:5 into a new T25 flask. Cells were routinely passaged 1:5 twice weekly for three weeks and supernatants were tested for the presence of virion associated RT activity by a Product Enhanced Reverse Transcriptase (PERT) Assay, as previously described (13), except using a PrecisionPLUS qPCR SYBR Master Mix (Primer Design, Chandler's Ford, UK) and analysed on a QuantStudio 7 Flex Real-Time PCR machine (Thermo Fischer Scientific).

Electron Microscopy and viral particle morphology

HPG, M MLV, and pcDNA3.1 transfected cell cultures were generated as described above, in "Transfection of 293T cells for generation of HPG & M MLV viral particles", using Lipofectamine 2000 (Invitrogen). Untransfected cells were used as a control. For thin section electron microscopy (EM), cells were pelleted and immersed in 2.5% glutaraldehyde in 0.1 M Sorenson's phosphate buffer pH 7.2, (300 mOsmol/kg) for 1 h at room temperature. After washing with Sorensen's phosphate buffer the cells were fixed with 1% osmium tetroxide for 1 h and dehydrated in a graded ethanol series at room temperature. Samples were embedded in Spurr's resin (ProSciTech, Australia) according to the manufacturer's protocol. Ultrathin sections were obtained using a Leica ultracut UCT Microtome and stained with saturated uranyl acetate in 50% ethanol and lead citrate. All prepared grids were examined using a Philips CM120 or JEOL JEM 1400 transmission electron microscope at 120kV. Electron micrographs representing negative control untransfected and mock transfected cells did not reveal the presence of viral like particles (Supplementary Figure 8).

Reverse transcriptase divalent cation preference

To evaluate the divalent cation preference of HPG reverse transcriptase (RT), we performed a virion associated RT assay using the gammaretroviral RT co factor, manganese (Mn^{2+}), and the lentiviral RT co factor, magnesium (Mg^{2+}). We compared HPG RT activity to HIV 1 RT activity in the presence of each co factor. 293T cells were co transfected with different quantities of HPG (pCC1 HPG; 0.04 0.22 pmol), M MLV (pNCS; 0.04 0.22 pmol), or HIV 1 (pNL4 3; 0.01 – 0.04 pmol), with the total mass of

transfected plasmid DNA equalized by the addition of the empty expression vector pCR2.1 (Thermo Fisher Scientific). Cells transfected only with pCR2.1 (0.04 – 0.22 pmol) were used as controls. Transfections were performed using Lipofectamine 2000 according to the manufacturer's protocol. Transfected cell cultures were incubated at 37°C, 5% CO_2 for 48 h, and then virion containing supernatants were collected and clarified by centrifugation at 200 x g for 5 min. To assess magnesium co factor usage by the viral reverse transcriptase, we performed a virion associated RT activity, as previously described (12). To assess manganese co factor usage, the virion associate RT activity assay was modified by the replacement of magnesium with 0.1M manganese.

Replication kinetics assay

To determine the cell tropism of HPG compared to ecotropic M MLV, HPG and M MLV transfected cell cultures were generated as described in the previous section "Transfection of 293T cells for generation of HPG & M MLV viral particles". Virion containing supernatants were normalized by virion associated RT activity, using manganese as the cofactor, as described in (12) except with the above mentioned modifications. Virion containing supernatants were used to infect human HeLa, mouse 3T3, and bat PaKi cells. Untransfected 293T cell culture supernatant was used as a control. Cells were seeded in a 96 well plate (Sarstedt, Nümbrecht, Germany) at a density of 15,000 cells/well. When cells reached ~50% confluency, the media was replaced with 225 µL of normalized HPG or M MLV virion containing supernatant with the addition of DEAE Dextran (Sigma Aldrich) at a final concentration of 10 µg/mL. Cells were incubated for 6 h, and then the supernatant was removed, cells were washed twice with phosphate buffered saline (PBS), and 250 µL of DMEM medium was added. To assess cell culture supernatants for the release of viral particles, 20 µL samples were collected from the supernatant of each well at 24 h intervals for 5 days. To measure the presence of HPG or M MLV virions in the cell culture supernatant samples, we performed a virion associated RT activity assay.

Secondary infection assay

To confirm that HPG was capable of establishing successive rounds of infection, the HPG provirus was transfected into 293T cell culture as described in "Transfection of 293T cells for generation of HPG & M MLV viral particles". To establish a primary infection from HPG virions, 293T cells were infected as described in "Replication kinetics assay". To establish a secondary infection, clarified supernatant containing HPG viral particles generated from the primary infection was collected and used to establish a second infection in 293T cells as described in "Replication kinetics assay". 5 µL samples were collected at inoculation, 6 h following inoculation and washing (t = 0), and at 48 h following washing (t = 48). Collected samples were analyzed for the presence of virion associated reverse transcriptase activity by PERT assay, as described in "Generation of HeLa cells persistently infected with HPG". Values derived from the PERT assay represent arbitrary units of RT activity in comparison to a dilution series (10 fold dilution series down to $1.0x10^{-7}$) of HPG virions which were generated in house, as described in "Transfection of 293T cells for generation of HPG & M MLV viral particles.

Generation of gammaretroviral *env* gene expression constructs for pseudotyping

Cloning constructs pUC57 GALV env and pUC57 KoRV A env encoding the Envelope proteins of GALV (Genbank: KT724048) and KoRV A (Genbank: NC039228), respectively, were chemically synthesized (GenScript). The Envelope sequences were enzymatically cut from the cloning plasmids using *Bam*HI and *Xho*I enzymes (New England Biolabs) and ligated into the mammalian expression vector pcDNA3.1

(Invitrogen), using T4 DNA ligase (New England Biolabs) following the manufacturer's protocols, generating the expression vectors pcD GALV env & pcD KoRV A env. To generate an expression plasmid for the HPG Envelope protein, the HPG *env* gene was amplified from the pCC1 HPG plasmid using primers (HPG env F and HPG env R; Supplementary Table 5) designed to anneal upstream of the cytoplasmic accumulation element (14) and downstream of *env* stop codon. To facilitate directional cloning, *Eco*RI and *Xba*I restriction sequences were incorporated into the forward and reverse primers, respectively. The HPG *env* gene was amplified using the Phusion High Fidelity PCR Kit (New England Biolabs) according to the manufacturer's instructions and using 50 ng of pCCI HBPG template and 0.5 µM of each forward and reverse primer in a 20 µL reaction. The HPG *env* amplicon was ligated into the pcDNA3.1 vector using T4 DNA ligase (New England Biolabs) following the manufacturer's protocols, using the restriction enzymes *Eco*RI and *Xba*I (New England Biolabs), generating the expression plasmid pcD HPG env. The sequences of all expression plasmids were confirmed by Sanger sequencing. Expression plasmids for other Envelope proteins including VSV G (pVSV G), ecotropic MLV (pEco), 4070A amphotropic MLV (pAmpho), and 10A1 amphotropic MLV ('dualtropic' MLV, p10A1) were obtained from the Retro X Universal Packaging System (Takara Bio).

Generation of pseudotyped retroviral particles for host cell tropism and superinfection interference assays

To determine the tropism of HPG in comparison to M MLV, pseudotyped viral particles were produced using the Retro X Universal Packaging System (Takara Bio). To generate viral particles pseudotyped with each gammaretroviral Envelope protein, GP2 293 cells that express M MLV Gag and Pol were transfected with the reporter vector, pQCLIN, and the Envelope expression vector, pVSV G, pEco, pAmpho, p10A1, pcD HPG env, pcd KoRV A env, or pcD GALV env, to generate viral particles pseudotyped with Env derived from VSV G, Ecotropic MLV, Amphotropic MLV, HPG, KoRV_A, and GALV, respectively. T75 tissue culture flasks (Nunc/Thermo Fisher Scientific) were seeded with $2.1x10^6$ GP2 293 cells. Cells were incubated for 16 h and then transfected with 5 µg of pQCLIN and 10 µg of the Env expression plasmid using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol. Transfected cells were incubated for 48 h, and then viral particles were collected from clarified supernatants and concentrated using the Lenti X Concentrator (Takara Bio).

Mouse 3T3 cells, human HeLa cells, or HeLa cells persistently infected with HPG were seeded in 96 well plates (Nunc) at a density of 20,000 cells/well, and incubated for 16 h. Following incubation, equal volumes of pseudotyped viral particles containing DEAE Dextran (Sigma Aldrich) at a final concentration of 10 µg/mL was added to the cells. Cells were incubated for 48 h and then viral cell entry was determined by the presence of blue cell forming units as previously described (12).

RT qPCR for detecting the presence of gammaretroviral nucleic acids in bat samples

Nucleic acids from 50 µL PBS resuspended bat scat samples were extracted on a KingFisher Flex Purification System (Thermo Fisher Scientific) using the Applied Biosystems MagMAX 96 Viral RNA Isolation Kit (Thermo Fisher Scientific) according to manufacturer's protocol except final purified nucleic acids were eluted in 50 µL nuclease free water instead of 90 µL elution buffer. No DNase I digestion was performed. 'Broad' primers (HPG rel F and HPG rel R) were designed to bind to HPG, FFRV1, MmGRV, and SaGRV, in the region upstream of the *gag* gene. Primers specific for HPG (HPG gag F and HPG gag R) were designed to bind to within the *gag* gene. Sequences for all primers are provided in (Supplementary Table 5). The presence of viral nucleic acids was determined by reverse transcriptase quantitative PCR (RT qPCR) using the Power SYBR Green RNA to CT 1 Step Kit (Thermo

Fisher Scientific). Bat samples were classified as positive or negative based on their fluorescence signal compared against a standard curve generated using $1x10^0 - 1x10^7$ copies of the HPG proviral plasmid. The cut off for determining a positive result was a cycle threshold of 36, which correlated to $1x10^1$ copies of the HPG provirus. Reaction mixtures contained 4.5 µL of purified RNA, 200 nM of each primer, and RT enzyme mix. Reactions were performed in either 384 or 96 well plates on the QuantStudio 7 Flex qPCR machine (Thermo Fisher Scientific). To determine if nucleic acid amplification was from RNA or DNA, an identical reaction was performed where the RT enzyme mix was excluded, to prevent amplification from RNA. Cycling conditions were as follows: 1x cycle of 48°C for 30 min, then 95°C for 10 min, and 40x cycles of 95°C for 15 s then 60°C for 1 min.

Generation of HPG Envelope protein for serological assays

A codon optimized (*Homo sapiens*) synthetic gene, encoding the predicted HPG retrovirus Env ectodomain (Glu^{38} Ser^{603}), was chemically synthesized (GeneArt, Regensburg, Germany). The synthetic gene incorporated an in frame 5' *Nhe*I site, a C terminal His_6 tag followed by a termination codon and 3' *Xba*I site. In addition, the putative SU TM cleavage site, Arg^{473}LeuLysArg, was ablated by substitution with Ser^{473}LeuGlnSer. The synthetic gene was ligated downstream of the tissue plasminogen activator leader sequence in the pcDNA3 based vector, pcE2^{661}myc (15) to give pcHPG Env^{603}. For expression, 293 F cells were transfected with pcHPG Env^{603} using 293fectin (Thermo Fisher Scientific). At 24 h post transfection, 0.5% (w/v) lupin peptone and 0.02% (w/v) pluronic F 68 were added to cells. Proteins were harvested following 3 5 days of incubation by centrifugation at 1,500 x *g* for 5 min, followed by filtration through a 0.45 µM filter before storage at 4°C. The Envelope protein was purified using Talon metal affinity resin (Takara) and 250 mM imidazole/PBS as the elution buffer. The protein was exchanged into PBS and concentrated using an Amicon centrifugal filter device (Merck Millipore, Burlington, USA). SDS PAGE in the presence and absence of β mercaptoethanol revealed a single diffuse band with a molecular weight range of ~80 90 kDa (Supplementary Figure 8), consistent with the molecular weight predicted from the amino acid sequence (62,805 Da) with 6 N linked glycans (~ 18 kDa).

Generation of anti HPG Envelope sera

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

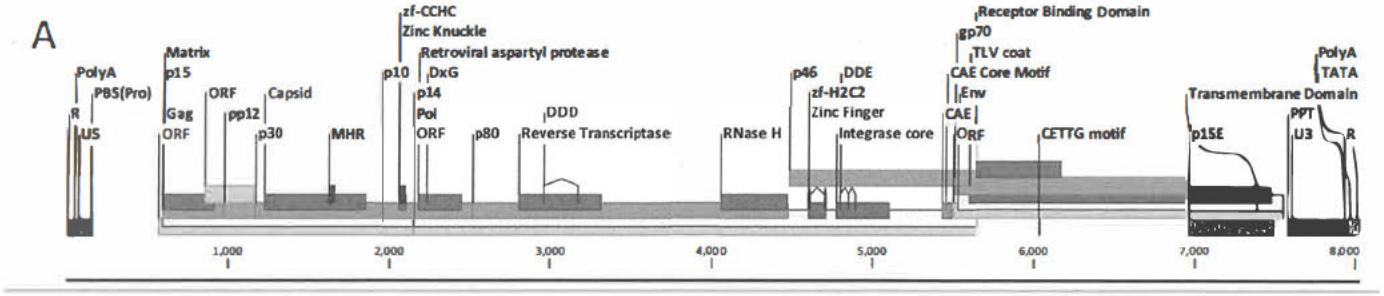
Serological assay for the presence of anti HPG antibodies in bats

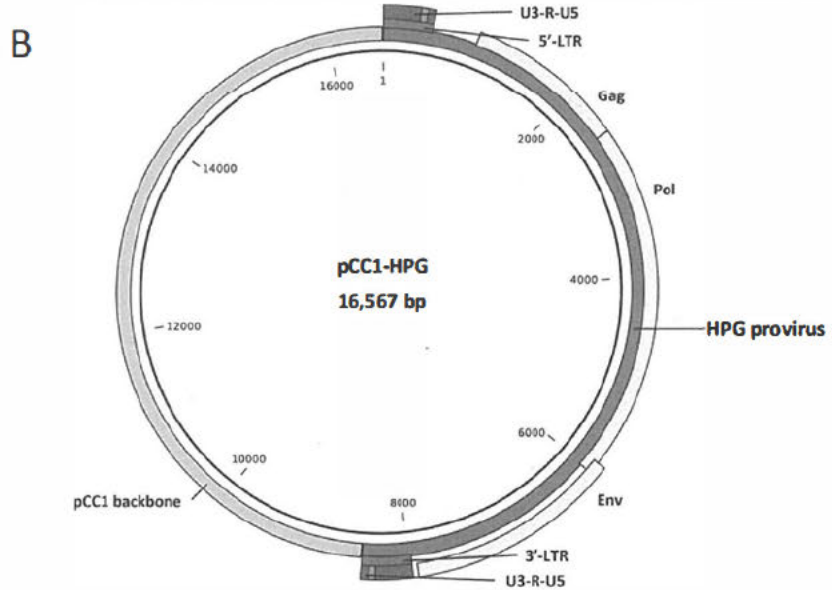
Bat sera were screened for the presence of antibodies reactive to the HPG Env trimer and the VRA region of Env using synthetic peptides in a solid phase enzyme immunoassay. N terminal biotinylated synthetic peptide encoding the HPG VRA region (LETWDIPDSDVSASTRVRPADSD, Genscript, USA) was added to Avidin coated plates (Nunc, Maxisorb) at 5 µg/ml followed by the addition of serially diluted bat serum in PBS containing 2.5 mg/ml bovine serum albumin and Tween 20 (0.05%). Bound antibodies were detected with horseradish peroxidase labelled Protein A/G (Thermo Scientific, Rockford) followed by 3,3',5,5' Tetramethylbenzidine (TMB) substrate (Sigma, USA). Sera that

displayed above ~~10x~~ 5x background levels of binding were further screened for reactivity to biotinylated synthetic peptides of the equivalent regions of KoRV A (LESWDIPELTASASQQARPPDSN), GALV (LESWDIPGTDVSSSKRVRPPDSD), and MLV (PSYWGLEYQSPFSSPPGPPCCS) in the same way.

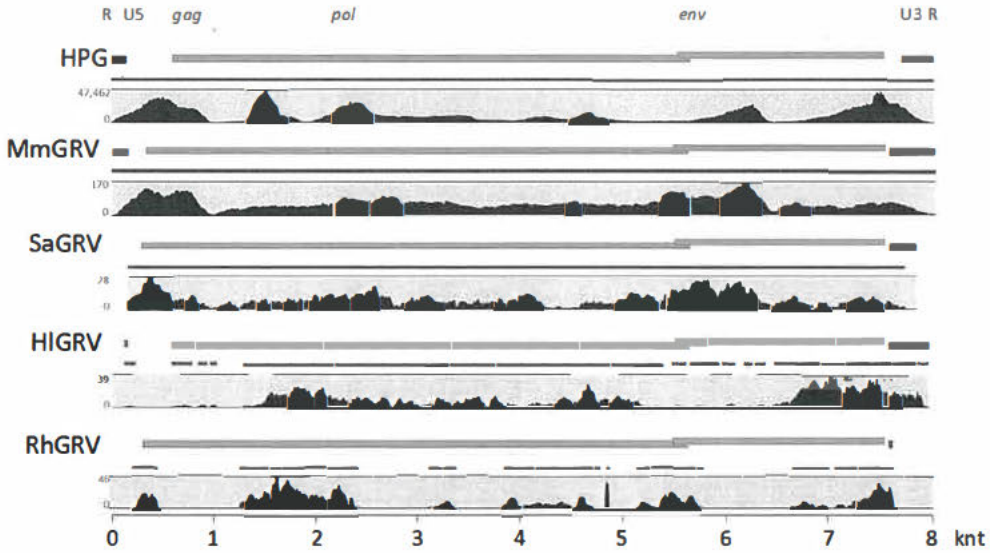
Commented [JH1]: Need a line regarding a-MLV sera from Heidi/Andy

Supplementary Figures

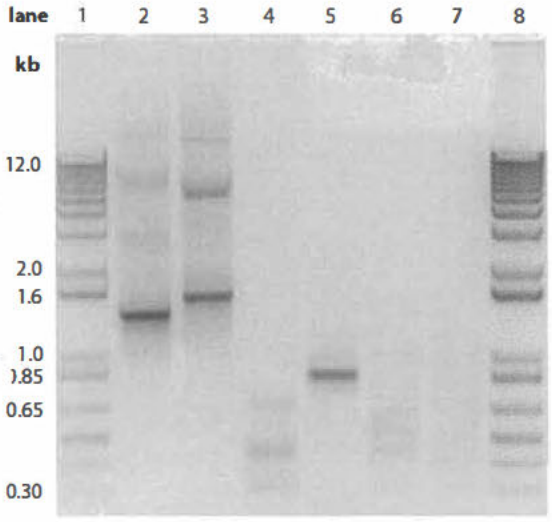




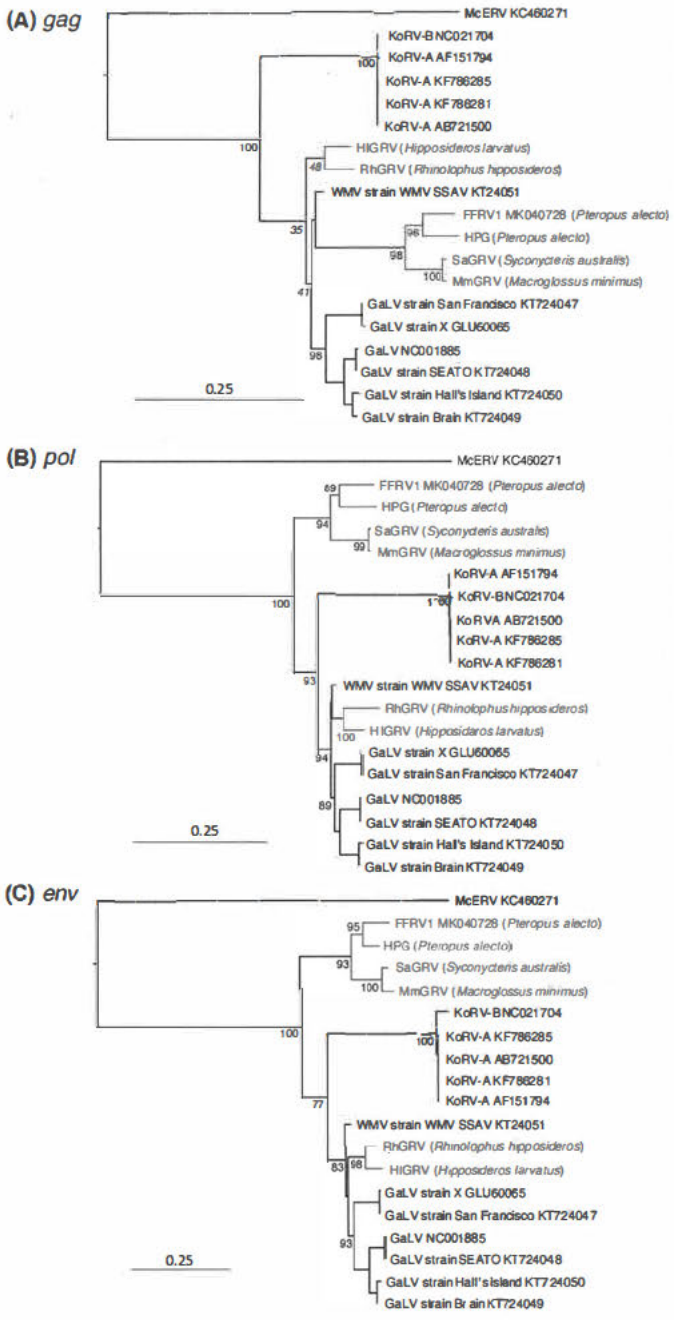
Supplementary Figure 1. *Schematic diagrams of the Hervey pteropid gammaretrovirus (HPG) expression construct and genome.* (A) Annotated HPG genome. Colors represent open reading frames (ORF; grey), nucleotide motifs (dark purple), polyprotein coding sequences (yellow), putative mature proteins (orange), 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.



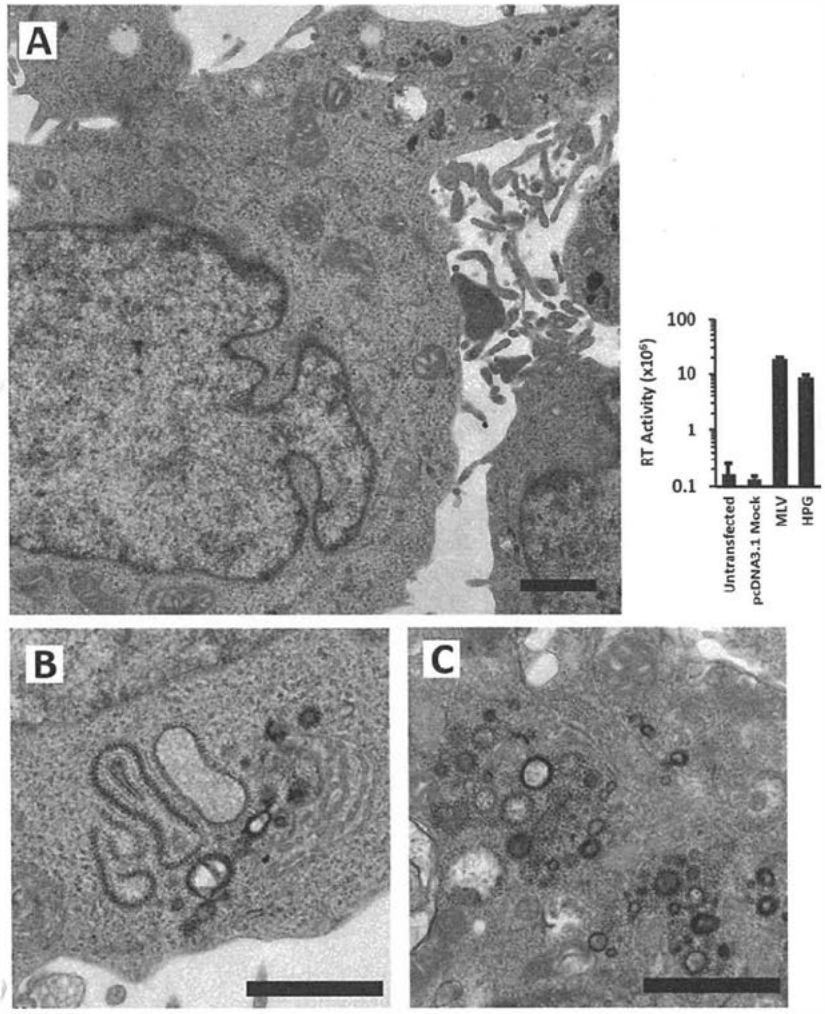
Supplementary Figure 2. *Sequencing read maps for KoRV-related retroviruses in Australian bats.* Colored bars indicate the genomic regions. Red: R, repeated regions at both ends of RNA genome; Dark blue: U5/U3, unique 5' & 3' regions; Light blue: retroviral genes *gag*, *pol*, and *env*. Black graphs 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; HlGRV, Hipposideros larvatus gammaretrovirus; RhGRV, Rhinolophus hipposideros gammaretrovirus.



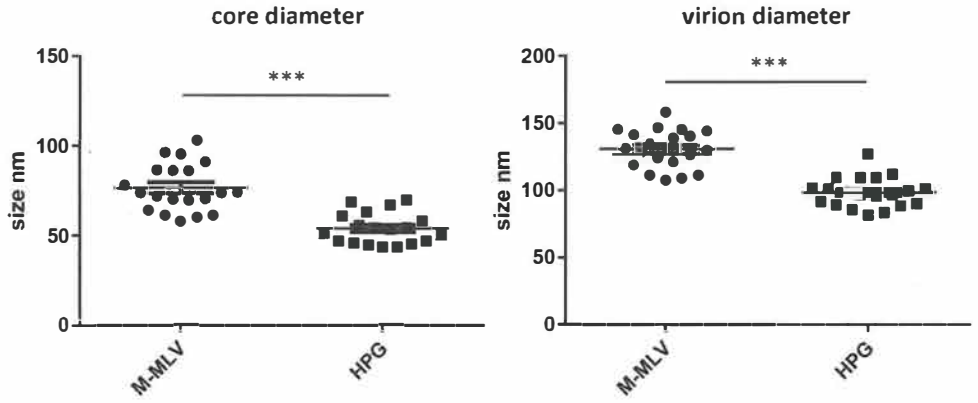
Supplementary Figure 3. *Assessment of the presence of endogenous Hervey pteropid gammaretrovirus (HPG) in the Pteropus alecto genome.* PCR primers (Supplementary Table 5) were used to amplify HPG DNA to generate HPG amplicons of 1.34 kb or 1.55 kb and were visualised on a 1% (w/v) agarose gel. Lane 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.



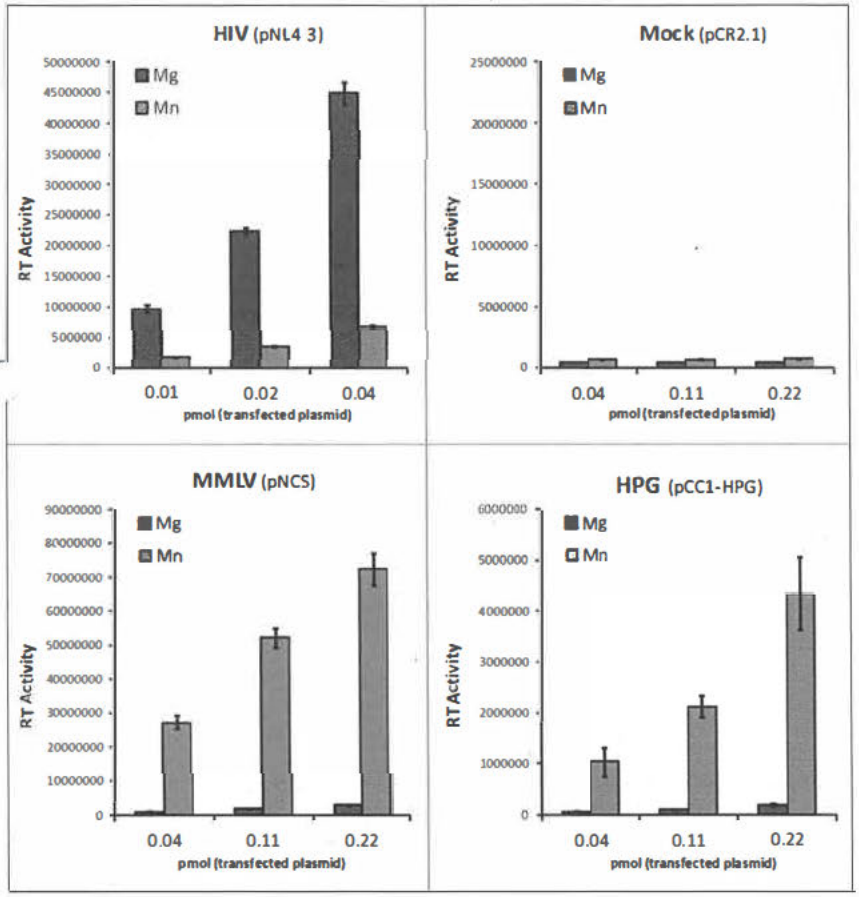
Supplementary Figure 1. Evolutionary relationships among KoRV related viral genes. Maximum likelihood phylogeny of the (A) *gag*, (B) *pol*, and (C) *env* genes of 19 gammaretroviruses. All branches are scaled according to the number of nucleotide substitutions per site, and branches representing bat retroviruses are shown in red. The tree was rooted using the McERV (Mus caroli endogenous retrovirus) KC460271 sequence. HPG, Hervey pteropid gammaretrovirus; FFRV1, flying fox retrovirus; MmGRV, Macroglossus minimus gammaretrovirus; SaGRV, Syconycteris australis gammaretrovirus; KoRV, Koala retrovirus; WMV, Woolly monkey virus; HIGRV, Hipposideros larvatus gammaretrovirus; RhGRV, Rhinolophus hipposideros gammaretrovirus; GaLV, Gibbon ape leukemia virus. Values at the nodes represent boostrap support.



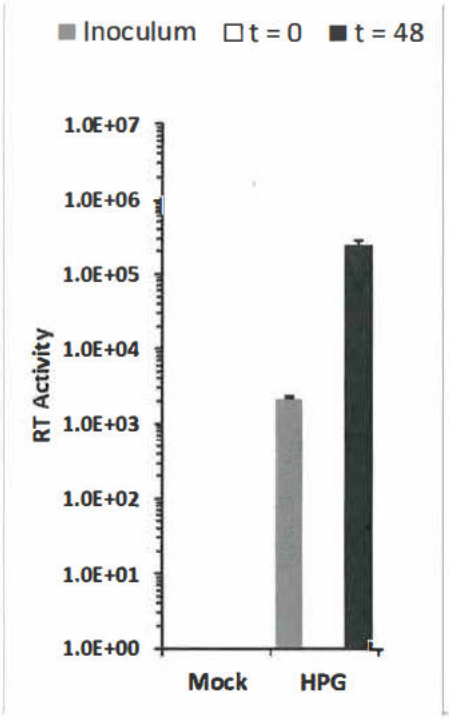
Supplementary Figure [redacted]. *Electron micrographs (EM) of control untransfected cells and cells mock transfected with the empty vector pcDNA3.1, and virion associated reverse transcriptase (RT) assay.* The graph displays the result of a virion associated RT assay, supporting the result that cells transfected with proviral M MLV and HPG expression plasmids generate retroviral particles, while untransfected and mock pcDNA3.1 transfected cells do not. 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 and 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 µm.



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.



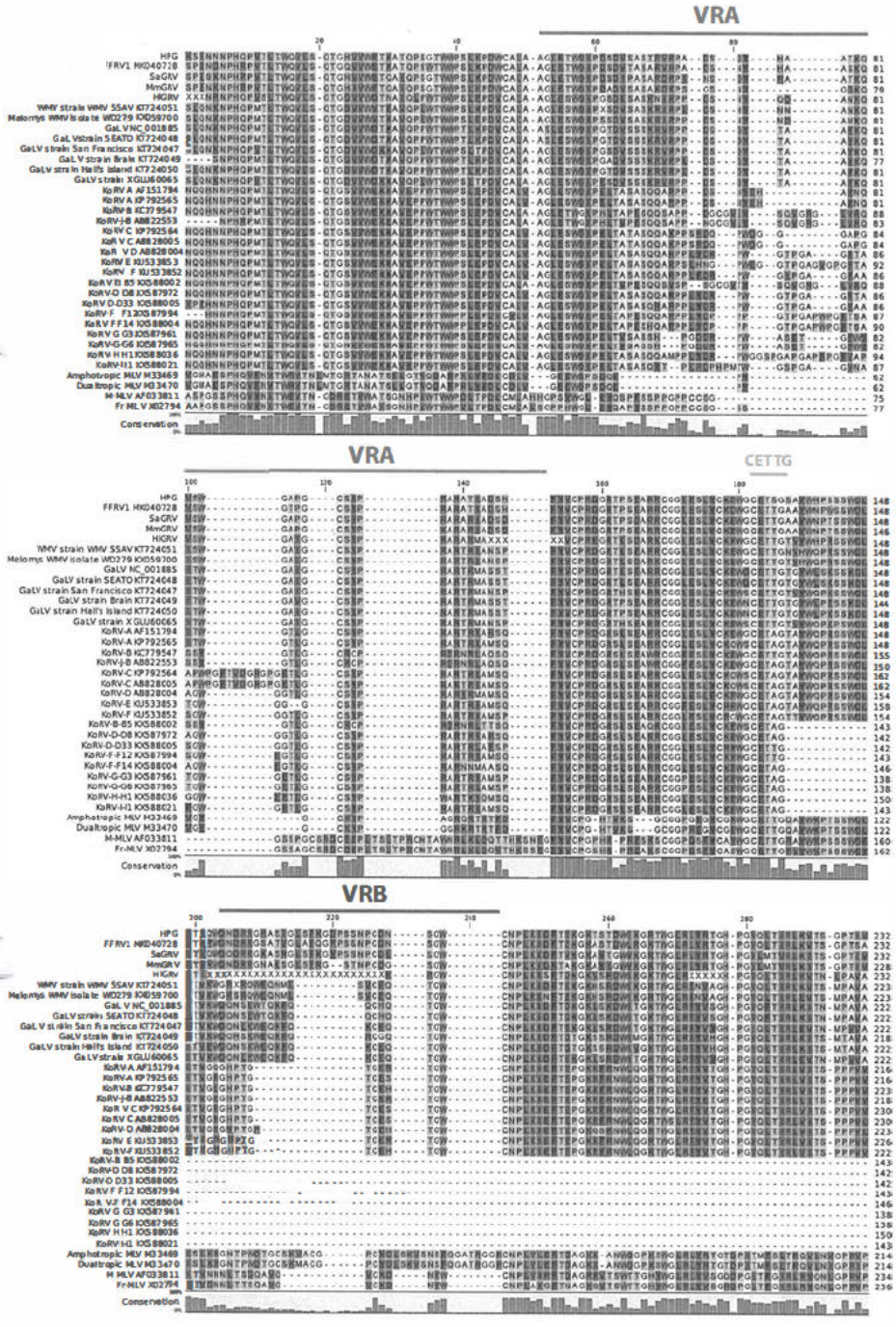
Supplementary Figure 5. *Divalent cation preferences of Human immunodeficiency virus (HIV), Moloney murine leukemia virus (MMLV), and Hervey pteropid Gammaretrovirus (HPG) virion associated reverse transcriptase.* Human or murine cell lines were transfected with infectious molecular clones of HPG, MMLV, 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).



Commented [JH2]: This figure will be updated to show representative bars for mock and HPG-t=0 as 'below limit of detection'.

Supplementary Figure X. *Secondary infection assay of HPG in 293T cells.* 293T cells were infected with clarified supernatant collected from 293T cells infected by HPG virions, establishing that HPG is capable of generating successive rounds of infectivity in human cells. Untransfected 293T cell culture supernatant (Mock) was used as a control. HPG and Mock supernatants were used in a product enhanced reverse transcriptase (PERT) assay to determine virion associated RT activity. Y axis values represent arbitrary units of RT activity in comparison to a dilution series of HPG virions. Error bars represent the standard error of the mean (SE, n = 3).

Commented [PHD3]: At T=0, there is no detectable RT activity-do you mean this is below the threshold of the assay? It looks a bit strange as you can't see the bar for t=0. A suggestion is to extend the y-axis to say 0.1 (if this is the ToD), show the bar and have a dotted line at 0.1 to say this is the threshold. The mock would also then have bars at/below the ToD for all time points.


VRA
VRA
CETTG
VRB
Conservation

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

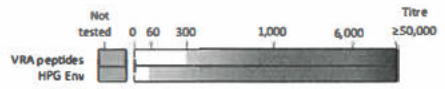
The alignment was generated using MUSCLE (9). The non KoRV related murine gammaretroviruses amphotropic MLV, dualtropic MLV, M MLV and Fr MLV are included for comparison. Highly variable regions A (VRA) and B (VRB) indicated by the green and red lines, respectively. The CETTG motif is denoted with an 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.

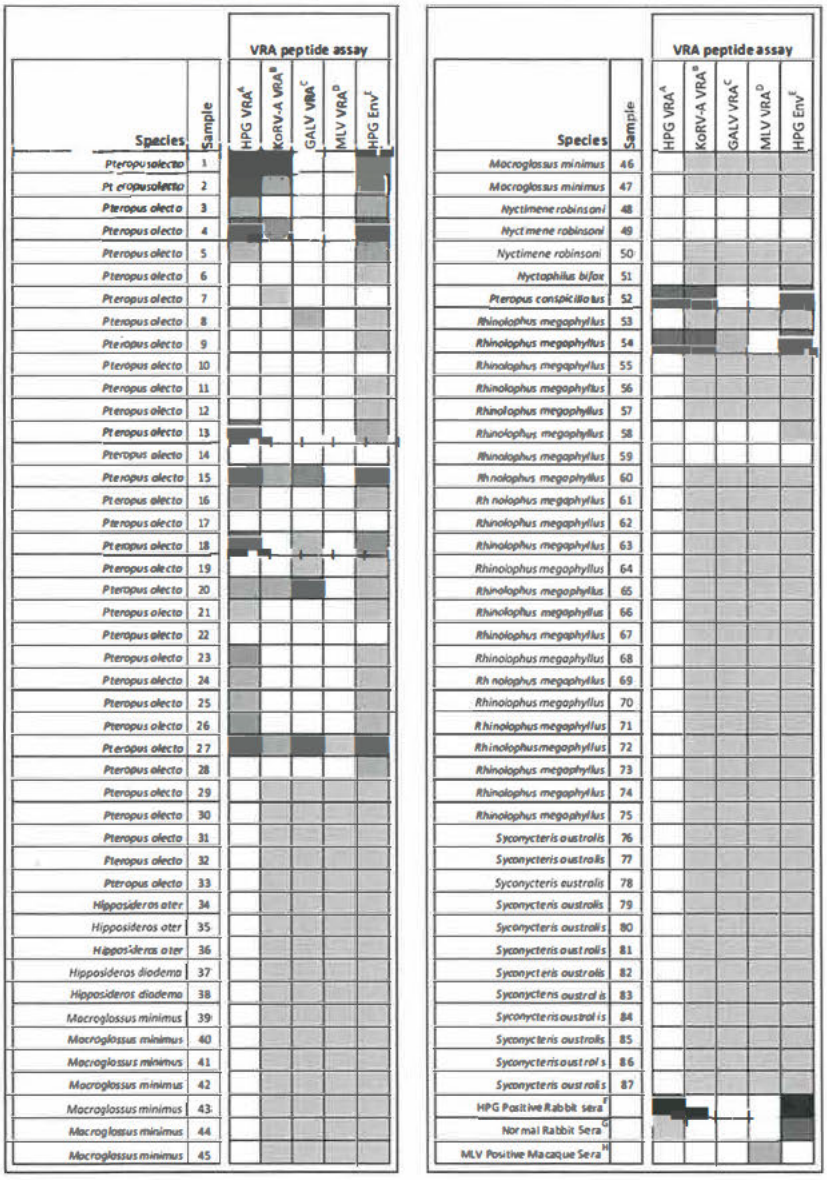
Commented [JH4]: Figure has been updated to include amphotropic and dualtropic MLV sequences

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 (16).

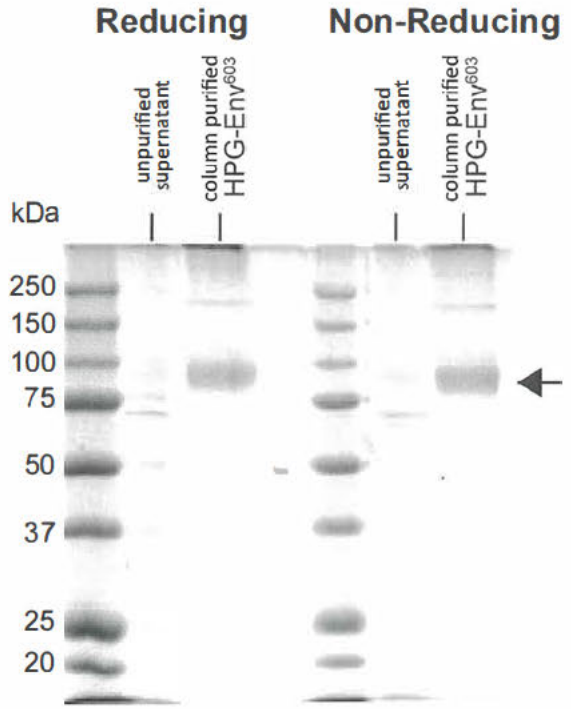
VRA peptide sequences

HPG: LETWDIPDSDVSASTRVRPADSD
KoRV-A: LESWDIPGTDVSSSKRVRPPDSD
GALV: LESWDIPELTASASQQARPPDSN
MLV: PSYWGLEYQSPFSSPPGPPCCS





Supplementary Figure X. *Serological analysis of bat samples for reactivity to KoRV related protein sequences.* Bat sera were screened for the presence of antibodies reactive to the variable region A (VRA) of HPG, KoRV, GALV, MLV, and the ectodomain of HPG envelope protein. The heatmap depicts the results; white indicates a negative result, colors other than white indicate a positive result, with magnitude indicated by the color scale. [A]HPG VRA peptide; [B]Koala retrovirus A VRA peptide; [C]Gibbon ape leukemia virus VRA peptide; [D]Murine leukemia virus VRA peptide; [E]HPG trimeric Envelope ectodomain protein; [F]Positive sera, rabbit anti HPG Env; [G]Negative sera, rabbit prebleed serum; [H]Positive sera, macaque anti MLV sera. Titre calculated as the reciprocal of the dilution of the serum needed to generate a signal five times above background.



Supplementary Figure 9. *Hervey pteropid gammaretrovirus recombinant envelope protein expressed from the pcHPG Env^{603} construct.* Lanes depict expressed supernatant before and after HiTrap His column purification. SDS PAGE in the presence (reducing) and absence (non reducing) of β 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.

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

HPG, Hervey pteropid gammaretrovirus; -GRV = gammaretrovirus

Supplementary Table 2. Accession numbers

Abbreviation	Name	Accession
FFRV1	Flying fox retrovirus isolate FFRV1	MK040728
GALV	Gibbon ape leukemia virus	NC_001885
GALV	Gibbon ape leukemia virus strain Brain	KT724049
GALV	Gibbon ape leukemia virus strain Hall's Island	KT724050
GALV	Gibbon ape leukemia virus strain San Francisco	KT724047
GALV	Gibbon ape leukemia virus strain SEATO	KT724048
GALV	Gibbon ape leukemia virus strain X	GLU60065
HlGRV	Hipposideros larvatus gammaretrovirus	MN413613
HPG	Hervey pteropid gammaretrovirus	MN413610
KoRV-A	Koala retrovirus	AF151794
KoRV-A	Koala retrovirus clone KV522	AB721500
KoRV-A	Koala retrovirus isolate Pci-maex1738	KF786281
KoRV-A	Koala retrovirus isolate Pci-SN265	KF786285
KoRV-B	Koala retrovirus isolate Br2-1CETTG	NC_021704
McERV	Mus caroli endogenous virus	KC460271
MmGRV	Macroglossus minimus gammaretrovirus	MN413611
RhGRV	Rhinolophus hipposideros gammaretrovirus	MN413614
SaGRV	Syconycteris australis gammaretrovirus	MN413612
WMV	Woolly monkey virus strain WMV SSAV	KT724051

Supplementary Table 3. Summary results of HPG seroprevalence

Bat species	Samples tested	α-HPG-VRA Ig-positive (Titre ≥ 300)	
Hipposideros ater	3	0	0.0%
Hipposideros diadema	2	0	0.0%
Macroglossus minimus	9	0	0.0%
Nyctimene robinsoni	3	0	0.0%
Nyctophilus bifax	1	0	0.0%
Pteropus alecto	33	16	48.5%
Pteropus conspicillatus	1	1	100.0%
Rhinolophus megaphyllus	23	1	4.3%
Syconycteris australis	12	0	0.0%
Total	87	18	20.7%

HPG, Hervey pteropid gammaretrovirus; VRA, Variable region A

Commented [JH5]: Table has been updated to represent VRA data

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
***Pteropus alecto* APOBEC3Z3 gene**	A3Z3F	(2300..2317)	CAGCTCCGAGTCAAAAAG
	A3Z3R	(3104..3123)	AGCGGATCTTGTTGATAAAG
HPG *pol* 1.34 kb sequence	HPG-pol-F1	(4184..4204)	GAACTCATCGCCTTGACTCAG
	HPG-pol-R1	(5521..5500)	AGCAATACCGTCGACCTTTACC
HPG *pol* 1.55 kb sequence	HPG-pol-F2	(2967..2988)	TCTTCTGCCTCAAACTGCATCC
	HPG-pol R2	(4511..4491)	CTGTGGTTTCAGCCAGTACTC
HPG *env* gene	HPG-env F	(5354..5375)	GGAAGAATTCAAAGAGGTATACAGACCTGG
	HPG env R	(7998..8020)	GCATTCTAGAAGAGGTTTATTAGGTACACGGG
HPG *gag* 'specific'	HPG gag F	(512..532)	AACTCGCTACCGCTTTCCATT
	HPG-gag-R	(683..664)	CTTCCACGGACAGGTTGTGA
HPG-related leader 'broad'	HPG-rel-F	(192..212)	CCATCGACGGGAGGTAAGC
	HPG-rel-R	(389..373)	CTGATCCTGGGGCGTCC

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[1] (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[2]		Positive for HPG-specific RNA[2]	
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%
Goonah	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%
Total samples	373	57	15.3%	25	6.7%	13	3.5%	25	6.7%

[1]57 samples that were positive in the 1st assay for HPG related nudeic 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.

[2]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. Vermeire J, *et al.* (2012) Quantification of reverse transcriptase activity by real time PCR as a fast and accurate method for titration of HIV, lenti and retroviral vectors. *PLoS One* 7(12).

14. 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.

15. Drummer HE, Boo I, Maerz AL, & Poumbourios 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.

16. 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.