bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Unique Evolution of Antiviral Tetherin in Bats

Joshua A. Hayward[a,b], Mary Tachedjian[c], Adam Johnson[a], Aaron T. Irving[d,e,f], Tamsin B. Gordon[a,b], Jie Cui[g,h], Alexis Nicolas[c], Ina Smith[c], Vicky Boyd[c], Glenn A. Marsh[c], Michelle L. Baker[c], Lin-Fa Wang[f,i,j], Gilda Tachedjian[a,b,j]

[a]Health Security Program, Life Sciences Discipline, Burnet Institute, Life Sciences, Melbourne, VIC 3004, Australia; [b]Department of Microbiology, Monash University, Clayton, VIC 3168, Australia; [c]CSIRO, Australian Centre for Disease Preparedness, Health and Biosecurity Business Unit, Geelong, VIC 3220, Australia; [d]Zhejiang University-University of Edinburgh Institute, Zhejiang University School of Medicine, Zhejiang University International Campus, China; [e]Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China; [f]Programme in Emerging Infectious Diseases, Duke-NUS Medical School, Singapore, Singapore; [g]CAS Key Laboratory of Molecular Virology & Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences, China; [h]Center for Biosafety Mega-Science, Chinese Academy of Sciences, China; [i]Singhealth Duke-NUS Global Health Institute, Singapore; [j]Department of Microbiology and Immunology, The University of Melbourne, at The Peter Doherty Institute for Infection and Immunity, Melbourne, VIC 3000, Australia.

Corresponding Author

Gilda Tachedjian, 85 Commercial Rd, Melbourne, 3004, +61-3-9282-2256, gilda.tachedjian@burnet.edu.au

Keywords

Bats; Tetherin; Restriction Factors; Innate Immunity; Antiviral; Virus

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Bats are recognised as important reservoirs of viruses deadly to other mammals, including humans. These infections are typically nonpathogenic in bats raising questions about host response differences that might exist between bats and other mammals. Tetherin is a restriction factor which inhibits the release of a diverse range of viruses from host cells, including retroviruses, coronaviruses, filoviruses, and paramyxoviruses, some of which are deadly to humans and transmitted by bats. Here we characterise the tetherin genes from 27 species of bats, revealing that they have evolved under strong selective pressure, and that fruit bats and vesper bats express unique structural variants of the tetherin protein. Tetherin was widely and variably expressed across fruit bat tissue-types and upregulated in spleen tissue when stimulated with Toll-like receptor agonists. The expression of two computationally predicted splice isoforms of fruit bat tetherin was verified. We identified an additional third unique splice isoform which includes a C-terminal region that is not homologous to known mammalian tetherin variants but was functionally capable of restricting the release of filoviral particles. We also report that vesper bats possess and express at least five tetherin genes, including structural variants, a greater number than any other mammal reported to date. These findings support the hypothesis of differential antiviral gene evolution in bats relative to other mammals.

Introduction

Bats are reservoirs of viruses that are highly pathogenic to other mammals, including humans. Viruses such as Hendra, Nipah, Ebola, Marburg, severe acute respiratory syndrome coronaviruses (SARS-CoV-1 and likely also SARS-CoV-2), and the Sosuga virus have crossed species barriers from bats into humans (1-5). Laboratory studies demonstrate that specific bat species can be infected with Ebola, Marburg, SARS-CoV-1, Hendra, and Nipah viruses without showing clinical signs of disease (6-10). Recently, this has led to efforts to elucidate if there are differences observed between the antiviral strategies of bats and other mammals (11-19). These studies have identified that genes related to immunity in bats are under a significant degree of positive selection, in addition to differences in the copy number and diversity of innate immune genes of bats relative to other mammals (12-15, 17, 20).

Tetherin (BST-2, CD317) is a mammalian restriction factor that inhibits the release of diverse enveloped viral particles from the cells in which they are produced (21-23). hese viruses include retroviruses, coronaviruses, filoviruses, and paramyxoviruses, some of which are transmissible by bats (24-27). Tetherin is a membrane-associated glycoprotein and its activity is determined by its structure, rather than its primary amino acid sequence (28, 29). In addition, tetherin possesses a secondary role

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

in immune signalling events by triggering the NF-κB signalling pathway, leading to stimulation of the antiviral interferon response (30-33).

In humans, tetherin is expressed across most cell types, including its BST-2 namesake bone marrow stromal cells, and its expression is upregulated by stimulation with type I interferons (34-36). Tetherin is a dimeric dual-anchor type II membrane protein that contains one protein anchor, a transmembrane domain near its N-terminus, an extracellular coiled-coil domain, and a glycophosphatidylinositol (GPI) lipid anchor, which is attached to its C-terminus as a post-translational modification (37-39). Tetherin contains a number of conserved cytosine and asparagine motifs within the extracellular domain, with respective roles in dimerisation and glycosylation, and a dual-tyrosine motif (Y·x·Y) in its cytoplasmic region, which has a role in viral particle endocytosis and immune signalling cascades (31, 40). Tetherin is located in lipid rafts at the plasma membrane where many viral particles bud during acquisition of their host membrane-derived viral envelopes (39). During the viral budding process one anchor remains embedded in the nascent viral envelope while the other remains attached to the plasma membrane, tethering the virion to the cell and preventing its release into the extracellular environment (28, 38).

Tetherin is common to all mammals, and orthologs which share structural and functional, but not sequence similarity, exist in other vertebrates (29). Most mammals carry only a single tetherin gene; however gene duplication has been observed in sheep, cattle, opossums and wallabies (29, 41). Furthermore, human tetherin is expressed in two alternative isoforms (30), the long (l-tetherin) and short (s-tetherin) isoforms, which differ through the truncation of 12 amino acid (AA) residues at the N-terminus of l-tetherin. Computational analysis of the genomes of megabats from the *Pteropus* genus predict two additional isoforms, X1 and X2, with an internal, rather than N-terminal difference in amino acid sequences, although whether these isoforms are expressed is unknown (12). The predicted isoform X1 is homologous to human l-tetherin, and the X2 isoform is a splice variant which contains a 7 AA exclusion within the extracellular coiled-coil domain relative to isoform X1. A recent analysis of tetherin in the fruit bats *Hypsignathus monstrosus* and *Epomops buettikoferi* revealed that these species expressed a homolog of *Pteropus* isoform X1 and that it is a functional restriction factor capable of inhibiting the release of the bat-hosted Nipah virus, as well as virus-like particles (VLPs) produced from the Ebola virus and the human immunodeficiency virus (23).

We hypothesised that given the role of bats as hosts of pathogenic enveloped viruses, and the differences observed in other antiviral genes of their innate immune system (14), that diversity might also exist in the tetherin genes of bats. Here we report that fruit bats possess a unique structural isoform which has differential activity against viral VLPs, restricting filoviral but not

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

retroviral VLPs, in addition to the two computational predictions, tetherin isoforms X1 and X2, which were generated through the automated NCBI annotation pipeline from the published *P. alecto* genome [NCBI: PRJNA171993] (12). We report that microbats of the suborder Yangochiroptera, genus *Myotis*, possess at least five tetherin genes, two of which contain large structural differences in the extracellular coiled-coil domain. An analysis of the tetherin genes of 27 species of bats revealed that bat tetherin has been subjected to strong positive selection, indicating that these flying mammals have experienced evolutionary pressure on this antiviral gene. Collectively our findings indicate that bats have undergone tetherin gene expansion and diversification relative to other mammals, supporting the hypothesis of differential antiviral gene evolution in bats.

Results

Bats possess structural homologs of human tetherin

To advance our understanding of bat tetherin homologs, we mined all available publicly accessible sequence read archives representing 26 species of bats through a BLASTn analysis (Table 1). The species analysed represent 8 of the 20 families within the order Chiroptera: three families (Hipposideridae, Pteropodidae, and Rhinolophidae) from within the suborder Yinpterochiroptera, and five families (Emballonuridae, Miniopteridae, Molossidae, Vespertilionidae, and Phyllostomidae) from within the suborder Yangochiroptera. This analysis revealed mRNA sequences representing homologs of human tetherin from 26 bat species, in addition to the homolog obtained from *P. alecto* cDNA, enabling the generation of consensus bat tetherin structural domains and amino acid sequence (Figure 1A & 1C) from a multiple sequence alignment (SI Data 1). The primary sequence lengths range from 177 to 221 AA, compared to 180 AA for human l-tetherin, and of these only 23 residues were conserved in 100% of the bat tetherin sequences analysed (Figure 1B). The assembled tetherin sequence for all bats was a homolog of human l-tetherin, except for the Black-bearded Tomb bat (*Taphozous melanopogon*), which was predicted to possess an equivalent of the human s-tetherin. The predicted *T. melanopogon* tetherin nucleotide sequence contains two key mutations upstream of its short isoform start site, an ATG to CTG change in the long isoform start codon, and a TAG stop codon in between the two start codon sites, either of which would be sufficient to prevent the production of the long isoform of tetherin. No nucleotide conflicts exist between the 17 reads mapping to this region of the *T. melanopogon* tetherin sequence, indicating that these mutations are unlikely to represent sequencing or consensus alignment errors.

The dual-tyrosine Y·x·Y motif, which is critical for mediating viral-particle endocytosis and is involved in immune signalling (31, 40), is variable among bat species and exists in different combinations of

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Y|C·x·Y|H (Figure 1A and 1C). All bat species possess at least one tyrosine residue within this motif (Figure 1C). Conservation of the protein domain organization and key structural motifs of bat tetherin, despite significant amino acid sequence diversity, supports the present understanding of tetherin as a protein whose functions are mediated through structural rather than sequence-mediated interactions (28).

Bat tetherin genes are under significant positive selection

To assess the evolutionary pressure applied during the speciation of bats, a selection test was performed to analyse bat tetherin genes spanning to ~65 million years ago (mya) of Chiropteran history. The analysis revealed that overall, bat tetherin genes have been subjected to positive selection, with an average ratio of non-synonymous (dN) to synonymous (dS) mutations, dN/dS, of 1.125 over the Chiropteran phylogeny, with numerous specific positions subjected to a large degree of positive selection, with dN/dS values of 2.626 - 2.668 (Table 2). These sites are predominantly located in the transmembrane domain and the regions of the cytoplasmic and extracellular domains immediately adjacent to the transmembrane domain (Figure 1A). This analysis suggests the presence of evolutionary pressure exerted on tetherin, by viral antagonists or countermeasures that target residues within tetherin, following the speciation of the Chiropteran order. The apparent ubiquity of tetherin among bats, in concert with significant sequence diversity and evidence of strong positive selection pressures operating on each species, supports the notion that tetherin plays a major role in the antiviral repertoire of bats.

Fruit bats possess unique structural isoforms of the tetherin protein generated through alternative splicing of a single tetherin gene

The expression of the fruit bat tetherin gene was initially assessed by a BLASTn search within the transcriptome of an Australian fruit bat, the Black flying fox, *Pteropus alecto*, using the human tetherin amino acid sequence as the search query. We identified 30 contigs matching the query (lowest E-value = 9.21×10^{-23}). A single *P. alecto* homolog of tetherin, herein described as isoform A, homologous to human l-tetherin and the predicted X1 isoform (GenBank: XM_006904279), was identified. *P. alecto* tetherin isoform A has low primary amino acid sequence conservation (37%) compared to human l-tetherin (Figure 2A) and other mammalian (28 - 47%) tetherin sequences homologous to human l-tetherin (Figure 2B). Despite the low amino acid sequence identity, the predicted secondary structures and protein domains are largely conserved between bats and other mammals (Figure 2B). All tetherin proteins containing the cytoplasmic (CD), transmembrane (TM), extracellular domains (ED), and the

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

post-translationally cleaved GPI-anchor signal peptide (GSP) with the exception of sheep tetherin B, which does not contain the GSP domain.

To verify the expression of bat tetherin isoform A, a cDNA library was prepared from *P. alecto* spleen tissue, comprising several immune cell types, and assessed for the presence of transcripts matching that identified in the *P. alecto* transcriptome. Primers were designed to flank the open reading frame of tetherin isoform A, and amplicons were generated by PCR. Amplicons were then cloned into plasmid vectors for DNA sequencing. This analysis revealed two additional isoforms of tetherin, isoforms B & C. Isoform B was homologous to the computationally predicted isoform X2 (GenBank: XM_015587122). Both isoforms B and C were predicted to contain structural differences relative to isoform A. They both possess a 7 AA exclusion in the middle of the extracellular domain while isoform C was additionally predicted to contain an alternative C-terminus, without the GSP domain (Figure 2B). The absence of a GSP domain suggests that it is unlikely that parallel homodimers of isoform C would possess the capacity to tether viral particles to the cell.

Mapping the tetherin cDNA sequences against the *P. alecto* genome identified a single tetherin gene located on Scaffold KB030270 (Figure 3A), between the MVB12A (multivesicular body subunit 12A) and PLVAP (plasmalemma vesicle associated protein) genes, similar to the human tetherin gene (Viewable in the Ensembl genome browser: http://www.ensembl.org/Homo_sapiens/Location/View?db=core;g=ENSG00000130303;r=19:17402886-17405683). *P. alecto* Isoform B and C mRNA are generated from the use of an alternative splice acceptor site for the first intron, resulting in the 7 AA exclusion, while the distinct C-terminus of isoform C is a result of alternative mRNA splicing that incorporates exon 5 while excluding exon 4 used in isoform A and B (Figure 3B).

Vesper bats possess multiple tetherin genes that contain unique structural variations

To further investigate the diversity of tetherin in non-fruit bat species, we amplified the tetherin nucleotide sequences for the vesper bat species *Myotis ricketti* and *Miniopterus schreibersii* using primers designed on the basis of the transcriptome sequences (Table 3A). These primers were used for PCR of cDNA samples generated from *M. ricketti* spleen tissue and a *M. schreibersii* kidney cell line to amplify the tetherin coding domain sequences (Table 3A). Primers for *M. ricketti* were also used to amplify the tetherin coding region from cDNA generated from a *Myotis macropus* kidney cell line as it was reasoned that since these species belong to the same genus, primers designed for one might be capable of amplifying tetherin from the other. Following PCR, amplified DNA from *M. ricketti, M. macropus*, and *M. schreibersii*, were gel purified, cloned and sequenced.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

These microbats were found to express various homologs of tetherin that include encoding of unique structural variations (Figure 4A and 4B). The tetherin homolog of human l-tetherin, predicted for *M. schreibersii,* was detected (Figure 4A). Five unique tetherin variants, differing in their encoded amino acid sequences (tetherin A – E), were identified for *M. macropus*, two of which, tetherin A and B, were also detected in *M. ricketti* (Figure 4A). These tetherin variants encode three distinct homologs (tetherin A, C, D) of the human l-tetherin, sharing the same protein domains but differing in their amino acid sequence (Figure 4B). Pairwise comparisons (Figure 5) reveal that the tetherin variants of *M. macropus* share between 84.2 – 92.0% amino acid sequence identity with each other, along with 44.6 – 53.5% identity with *P. alecto* tetherin isoform A, and 37.5 – 43.5% identity with human l-tetherin. The *M. ricketti* tetherin A and B share 100% amino acid sequence identity with their *M. macropus* counterparts. The tetherin of *M. schreibersii* shares between 34.6 and 49.7% sequence identity across all pairs.

An additional splice variant (isoform B) of *M. macropus* tetherin C was also identified. *M. macropus* tetherin C isoform B is predicted to have a C-terminal truncation and possesses only the cytoplasmic domain, transmembrane domain, and less than half of the extracellular domain (Figure 4B). Accordingly, it is predicted to lack the GPI signal peptide required for the post-translational addition of a GPI-anchor, also found for *P. alecto* tetherin isoform C (Figure 2B). The absence of a GPI anchor indicates that it is unlikely that homodimers of *M. macropus* tetherin C isoform B would possess the capacity to tether viral particles to the cell. *M. macropus* tetherin B and E were found to be structurally unique, possessing a large deletion of ~60 amino acids in the extracellular coiled-coil domain (Figure 4B), compared to the other tetherin homologs, which has been shown to be critical for viral particle restriction (28, 42). This deletion results in the exclusion of the conserved disulphide bond-forming cysteine residues, and the conserved asparagine-linked glycosylation sites, indicating that this form of tetherin is unlikely to form dimers or be glycosylated in the manner of human tetherin (28, 43). They are however predicted to possess all other tetherin domains including the transmembrane domain and GPI signal peptide, indicating that they may still be capable of tethering viral particles. *M. macropus* tetherin B and E share 92.1% amino acid sequence identity indicating they are the product of a relatively recent gene duplication (Figure 5).

While there is no available genome for *M. macropus*, the most closely related publicly accessible genome assemblies are for the relatively distantly related *Myotis lucifugus* and *M. davidii,* which diverged from *M. macropus* approximately 19 and 14 mya, respectively (44). For comparison, humans and chimpanzees diverged approximately 6.5 mya (45). A BLASTn analysis of the *M. macropus* tetherin nucleotide sequences against the *M. davidii* and *M. lucifugus* genomes revealed first that both species were too distantly related to *M. macropus* to match each *M. macropus* gene product specific to *M.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

davidii or *M. lucifugus* gene regions. However, the *M. lucifugus* genome assembly included a single gene scaffold (GL430608) which contained matches to all query *M. macropus* tetherin sequences, comprising seven potential *M. lucifugus* tetherin genes (Figure 6A) and indicating the presence of a tetherin gene locus within this scaffold. Further analysis of the mapped exons, as described below, enabled confirmation of the *M. lucifugus* tetherin regions. The *M. lucifugus* tetherin gene locus was located between the MVB12A and PLVAP genes, as observed for the single *P. alecto* tetherin gene (Figures 3 and 6).

The genome assembly of *M. davidii* also contained significant matches to *M. macropus* tetherin sequences, however these were spread across seven different gene scaffolds against which no *M. macropus* tetherin coding domain sequence could be matched in its entirety. For this reason, further genomic analysis of *Myotis* tetherin genes was confined to the *M. lucifugus* genome.

The *M. lucifugus* tetherin gene locus (Figure 6A) contains seven potential genes that are defined by the presence of two or more exons mapping to the *M. macropus* tetherin cDNA. Three exons across tetherin gene regions 1, 2, and 6, possess all of the coding exons required to express homologs to *M. macropus* tetherin A, C, and D (Figure 6B). Furthermore, *M. lucifugus* tetherin gene regions 2, and 3 possess all of the coding exons required to produce homologs to *M. macropus* tetherin B and E (Figure 6C). These data indicate that microbats of the genus *Myotis* possess a greater number and diversity of tetherin genes than any mammal reported to date.

These findings expand upon a recent report by Hölzer and colleagues (18), that identified the expression of three tetherin paralogs, upregulated by interferon treatment, in the vesper bat, *M. daubentonii*. Mapping of these paralogs against the genome of *M. lucifugus* revealed four tetherin genes, currently listed as novel gene predictions within the Ensembl database: ENSMLUG00000023691, ENSMLUG00000029243, ENSMLUG00000026989, and ENSMLUG00000023562 (18). Herein these genes are labelled Tetherin 1, 2, 6, and 7, respectively (Figure 6A).

Tetherin expression pattern in *P. alecto* tissue samples

Tetherin is widely and variably expressed across human tissues and is upregulated in response to interferon (35, 36). To assess the relative expression of bat tetherin across different tissue types, tissue samples were obtained from three individual *P. alecto* bats: a male, a pregnant female, and a juvenile male. Using primers that amplify all three isoforms of *P. alecto* tetherin, expression was analysed by qPCR with Ct values normalised against the 18S rRNA housekeeping gene (Figure 7). The two male bats exhibited similar tetherin expression patterns, with greatest levels in the thymus. A thymus tissue sample was not obtained from the female, although a similar pattern to the male bats was observed

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

among most other tissues. The female bat had higher expression in the lung compared to the male bats, possibly reflecting an active response to a respiratory infection, an artefact of hormonal regulation, or natural heterogeneity between bats.

Previous studies in vesper bats, have demonstrated that tetherin is upregulated in response to type I interferon alpha stimulation (18). To determine if pathogen-mediated stimulation can upregulate tetherin expression in fruit bats we analysed the transcriptome of *P. alecto* spleen tissue from animals treated *in vivo* with the Toll-like receptor (TLR) agonists, lipopolysaccharide (LPS) or polyinosinic:polycytidylic acid (PIC). In *P. alecto*, LPS treatment increases the percentage of B-cells in the spleen (46), and PIC treatment was observed to increase the expression of the interferon-stimulated gene, *ISG54* (47). We found that *P. alecto* tetherin was significantly upregulated by LPS (P = 0.028) and PIC (P = 0.004) (Figure 8).

We next assessed if there was a difference in the expression of the alternative tetherin isoforms A, B and C. This analysis revealed that tetherin is predominantly expressed as isoform B (14/15 samples; Figure 9A & 9B) while stimulation with LPS significantly increased the expression of Isoform A (P = 0.048), but not Isoform B or C. Stimulation with PIC increased the expression of isoform A and B compared to unstimulated tissue; however this effect was variable among individual bat spleen samples and did not reach significance. Expression of Isoform C was only observed in a single PIC treated spleen sample. These data show that tetherin isoforms are expressed in tissue and can be upregulated by stimulation with TLR-agonists.

P. alecto tetherin protein expression and cellular localisation

To characterise bat tetherin, we analysed tetherin from two Australian bats, the fruit bat *P. alecto* and the vesper bat *M. macropus*. The coding sequences of the *P. alecto* tetherin isoforms A, B, and C, and *M. macropus* tetherin A and B, were cloned into mammalian expression plasmids. While numerous vesper bat tetherins were identified in this study, *M. macropus* tetherin A and tetherin B were selected for further analysis as tetherin A is a homolog of human l-tetherin and tetherin B was the most structurally unique, possessing a large deletion within the extracellular coiled-coil domain compared to tetherin A (Figure 4B).

All tetherin coding sequences were modified with the addition of a haemagglutinin (HA) tag sequence (N-SG**YPYDVPDYA**GS-C) to enable antibody-based detection of protein expression. In all cases the HA tag was inserted at the position immediately following the end of the coiled-coil region of the extracellular domain (Figure 1A) which is the equivalent location to that previously utilised in the tagging of human tetherin (28).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

To evaluate tetherin protein expression, mammalian HEK293T cells, which do not express tetherin in the absence of type I interferon stimulation (21), were transfected with bat tetherin expression constructs. Cell lysates were extracted using a method for enhanced extraction of GPI-anchored proteins (48) and tetherin expression was observed by non-reducing SDS-PAGE and Western blot analysis (Figure 10).

P. alecto tetherin isoforms A and B were detected primarily as broad bands at the expected positions ~56-75 kDa while isoform C was present as both dimers and predominantly, monomers (~32-35 kDa) (Figure 10A). The presence of broad/multiple bands in close proximity to each other likely reflects variable levels of tetherin glycosylation as previously reported for human tetherin (28). *M. macropus* tetherin A was also detected as a dimer at the expected size of ~55 kDa while tetherin B was only detected as a monomer with a size of ~26 kDa (Figure 10B). Bat tetherin that are homologous to human l-tetherin dimerised as expected, while the structurally unique tetherin proteins behaved differently. In this regard, *P. alecto* tetherin isoform C formed both dimers and monomers, while *M. macropus* tetherin B was detected exclusively as a monomer. These data show that HA-tagged tetherin can be expressed and that these proteins differ in their ability to dimerise.

Human tetherin localises to the plasma membrane at sites of virus budding, in addition to membranes of the trans-Golgi network and recycling compartments (21, 49). To determine localisation of *P. alecto* tetherin isoforms, HEK293T cells were transfected with constructs expressing HA-tagged tetherin that was visualised by immunofluorescence. *P. alecto* tetherin isoform A localised to the plasma membrane, displaying a similar cellular localisation pattern as human tetherin (Figure 11A and 11C), while isoform C was predominantly localised in the cytoplasm (Figure 11E). *P. alecto* tetherin isoform B demonstrated a fluorescence pattern with features shared by both isoforms A and C (Figure 11D). These data show that bat tetherin isoforms can be expressed with isoforms A and B, but not C predominately localising at the plasma membrane.

P. alecto and *M. macropus* tetherin proteins display distinct ability to restrict HIVΔVpu virus-like particles

Tetherin from humans and other mammals including cats, sheep, and other bat species (*Hypsignathus monstrosus* and *Epomops buettikoferi*) restrict the release of HIV-1 VLPs from cells (23, 50, 51). To examine if tetherin from *P. alecto* and *M. macropus* function to similarly block viral particle release, we assessed their ability to restrict HIV-1 VLPs. HEK293T cells were co-transfected with constructs expressing bat tetherin and HIVΔVpu VLPs, that do not express Vpu, an antagonist of human tetherin (21). *P. alecto* tetherin isoforms A and B inhibited the release of HIVΔVpu VLPs in contrast to tetherin isoform C which failed to block VLP release (Figure 12A). *M. macropus* tetherin A also restricted

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

HIVΔVpu VLP release from HEK293T cells, however tetherin B was unable to inhibit HIVΔVpu VLP egress (Figure 12B). These data indicate that, with the exception of *P. alecto* tetherin isoform C and *M. macropus* tetherin B, tagged bat tetherin proteins are functionally capable of restricting the release of HIV-1 particles from mammalian cells in the absence of Vpu.

The structurally unique *P. alecto* tetherin isoform C restricts the release of filoviral virus-like particles

We next investigated whether *P. alecto* tetherin isomers are able to restrict filovirus VLPs, including isoform C, which contains a unique C-terminal domain compared to isomers A and B (Figure 2B). In contrast to experiments with HIV-1 VLP, tetherin isoform C was able to restrict the release of VLPs composed of Ebola and Marburg virus VP40 matrix proteins (P = 0.008 for both; Figure 13A & 13B). Ebola VLPs were restricted by isoforms A, B, and C to similar extents (Figure 13A), while Marburg VLPs were restricted by isoform C to a lesser extent than isoforms A and B (Figure 13B). These data demonstrate that all *P. alecto* isomers were able to restrict filoviral VLPs including isomer C which lacks the GSP domain.

Discussion

Bats are increasingly being recognised as hosts of viruses with zoonotic potential, driving efforts to better understand these host-virus relationships and the evolutionary features of bats that differentiate them from other mammals (14, 17, 52, 53). It has been hypothesised that the evolutionary adaptation to flight, including changes to the DNA damage response, increased metabolic rates, and higher body temperatures, have influenced the immune system of bats in such a way as to make them ideal hosts for viruses (12, 54, 55). To determine the differences between the innate antiviral defences of bats relative to other mammals, we analysed the genes and expressed transcripts of tetherin from diverse bat genera and species within the order Chiroptera. We found that in all but one species, bats possess genes that express transcripts encoding a tetherin protein homologous to the long isoform of human tetherin (l-tetherin). In addition, we found that *P. alecto* expresses three isoforms from a single tetherin gene, and that vesper bats (genus *Myotis*) encode five, and possibly as many as seven, distinct tetherin genes.

The one bat species that lacked the human l-tetherin homolog, *T. melanopogon*, possessed an equivalent of the short isoform of human tetherin (s-tetherin), an observation also reported for cats (56). In a study of feline tetherin, short isoform-exclusivity was found to improve viral restriction and decrease sensitivity to tetherin antagonism by HIV-1 Vpu when compared to an engineered long isoform version of feline tetherin (56). Conversely, the feline immunodeficiency virus (FIV), which is

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

adapted to the feline host, is not restricted by either long or short isoforms of feline tetherin (56). In humans, l-tetherin acts as a virus sensor which induces NF-κB signalling in contrast to s-tetherin, which lacks the dual tyrosine motif required for eliciting this innate immune response (30, 31). The findings presented here suggest that if bat tetherin proteins are confirmed to similarly mediate cytoplasmic domain-mediated immune signalling, then *T. melanopogon*, which only encodes a homolog of s-tetherin, would be predicted not to mediate this effect.

Bat tetherin amino acid sequences were found to be highly variable. The predicted protein lengths range from 177 to 221 AA (human l-tetherin is 180 AA) and of these, only 23 amino acid residues were conserved in 100% of the 27 bat tetherin sequences analysed (Figure 1B). Among these are the structurally important cysteine and asparagine residues which are responsible for tetherin dimerisation and glycosylation, respectively (57, 58). The dual-tyrosine motif, responsible for mediating viral particle endocytosis and immune signalling (31, 40), was found to exist as a variable Y|C·x·Y|H motif across the bat tetherin variants analysed. All bats maintained at least one of the two tyrosine residues. This observation is significant because mutational studies of human tetherin have demonstrated that the dual tyrosines provide redundancy for both the endocytic and signalling activities, which are maintained as long as either tyrosine is present (31, 40).

To understand the evolutionary selective pressures on tetherin genes across bat species, a selection test was performed that revealed that bat tetherin genes are under strong positive selection with amino acid positions in and around the transmembrane domain being the region under the strongest selective pressure (Table 2 and Figure 1A). This is in agreement with previous reports that primate tetherin possess multiple sites under positive selection in this same region (28). The driver of positive selection in the case of primate tetherin is antagonism by viral counter-measures including HIV-1 Vpu (59). Venkatesh et al. (60) demonstrated that the configuration that tetherin dimers adopt during viral particle retention primarily consists of the GPI-anchor being embedded in the viral envelope and the transmembrane domain remaining attached to the cellular membrane. If this paradigm holds true for bat tetherin, then it follows that the tetherin-cell membrane interface is the major site of tetherin antagonism in bats and it would be reasonable to speculate that the drivers of this selection are viral antagonists analogous in the mode of interaction, if not structure or function, to lentiviral tetherin antagonists.

We amplified tetherin from spleen-derived cDNA of an Australian fruit bat, *P. alecto*, and confirmed the expression of the two computationally predicted splice variants (isoforms A [X1] and B [X2]), and additionally identified the expression of a third isoform of tetherin, isoform C. Mapped against the *P. alecto* genome, all three isoforms were derived from the alternative splicing of a single tetherin gene

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Figure 3B). *P. alecto* tetherin isoform B possesses a 7 AA exclusion within the extracellular coiled- coil domain relative to isoform A, while isoform C is predicted to harbour the same 7 AA exclusion and an alternative C-terminus, that lacks the GSP domain predicted to be present in tetherin isoforms A and B which is necessary for the post-translational addition of the GPI-anchor. This is important because studies of human tetherin have demonstrated that the presence of a GPI-anchor is essential for restricting the release of viral particles (28, 42). Sheep and cows possess a duplication of the tetherin gene (41, 51). In sheep, the duplicate tetherin, named tetherin B, similarly does not encode a GSP, and studies of sheep tetherin B function reveal that while it is capable of limited restriction of VLPs, it is significantly less potent than sheep tetherin A (51). The mechanism of its function is unknown and its C-terminal amino acid sequence is entirely dissimilar from that of *P. alecto* tetherin isoform C. One possible function of additional isoforms of tetherin is to expand the viral target range of bats by undefined mechanisms that are active in the absence of a GPI anchor.

Intriguingly, one of the non-transmembrane domain-associated sites under strong positive selection across tetherin from bats is located in the middle of the extracellular domain. This is the site that distinguishes the *P. alecto* tetherin isoforms A and B, with the 7 AA exclusion at this location in isoform B. This suggests that the expression of *P. alecto* tetherin isoform B, and possibly also *M. macropus* tetherin B and E, which differ from tetherin A through a large 60 AA deletion in the same region, is to express a tetherin variant lacking sequence motif that appears to be a possible site of viral antagonism. This hypothesis could be supported by future experiments aimed at identifying viral antagonists of *P. alecto* tetherin isoform A that are ineffective against isoform B.

We found that the microbat *M. macropus* expresses five different tetherin genes, tetherin A, B, C, D and E, of which three (tetherin A, C, and D), encode homologs of human l-tetherin, while *M. macropus* tetherin B and E possess a large 60 AA deletion within the extracellular coiled-coil domain (Figure 4). Mapping of these genes against the genome assembly of the relatively distantly related *M. lucifugus* indicates that vesper bats of the genus *Myotis* may possess as many as seven tetherin genes (Figure 6). The presence of multiple tetherin genes suggests that these bats have expanded and diversified their use of these antiviral restriction factors relative to other mammals, supporting the hypothesis of differential antiviral gene evolution in bats.

Tetherin was expressed widely and variably across the tissues of *P. alecto*, which is consistent with previous observations of human tetherin (36). We observed the highest levels of *P. alecto* tetherin expression in the thymus, possibly reflecting the role of the thymus in expressing a large proportion of the proteome due to its role in central tolerance. High levels of expression were also observed in lung tissue, particularly in the lung of an individual female bat. This may reflect a frontline defensive role of tetherin in the lung against enveloped viruses that are respiratory pathogens.

Discussion

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Tetherin was upregulated in spleen tissue treated with TLR agonists LPS and PIC. Interestingly, we observed that the expression of isoforms A, B, and C was highly variable among individual spleen samples. In several spleens, only isoform B was expressed, while in one spleen treated with PIC the vast majority of expression was attributed to isoform A. The remainder expressed variable levels of each isoform A and B, with the majority of expression trending toward isoform B. Surprisingly, isoform C was only expressed, at a low level, in a single spleen sample that had been treated with PIC. The biological importance of this observation is not presently known. Ongoing assessments of bat tetherin expression should determine if this strong bias in alternative isoform expression is present in other tissues, such as the thymus and lung.

The expression of the *P. alecto* tetherin isoforms A, B, and C, and *M. macropus* tetherin A and B, reveals differences in their relative capacities to form homodimers under our assay conditions (Figure 10). This observation is notable because the ability of tetherin to form cysteine-linked dimers is required for the restriction of HIV-1 viral particles (58) which we chose as the VLPs against which bat tetherin proteins were functionally validated for inhibiting viral particle release. In contrast to restriction of HIV-1, tetherin dimerisation is not required for the restriction of arenaviruses and filoviruses (61), suggesting that the need for dimerisation is virus-dependent.

The *P. alecto* tetherin isoforms A and B, and *M. macropus* tetherin A, that were most similar to human l-tetherin, were all found to predominantly form dimers. This was determined by non-reducing SDS-PAGE and Western blot analysis where bands were observed almost exclusively at the expected size of tetherin dimers (Figure 10). In contrast, *P. alecto* tetherin isoform C was present as both dimeric and monomeric forms (Figure 10A). The presence of monomeric forms is surprising because *P. alecto* tetherin isoform C contains all of the conserved cysteine residues and the same extracellular coiled-coil domain as isoforms A and B. The distinct C-terminal region in isoform C is the result of an alternative splicing event that also causes the loss of the final three amino acid residues of the coiled-coil region possessed by isoforms A and B. Additionally, the inserted HA tag is located in the position immediately following the coiled-coil domain. It is possible that the absence of the three terminal amino acid residues in the coiled-coil domain of isoform C relative to isoform A and B accounts for the difference in relative extents of dimerisation. Alternatively, the observed extent of dimerisation of isoform C may be an artefact of the method used to process the samples for analysis and may not be reflective of the extent of dimerisation *in situ*. It is also important to note that we cannot currently rule out the possibility that the insertion of the HA tag in between the coiled-coil domain and the alternative C-terminal sequence might be responsible for the observed reduction in dimerisation. *M. macropus* tetherin B was observed exclusively as a monomer (Figure 10B). The monomeric exclusivity of *M. macropus* tetherin B can be explained by the loss of all conserved cysteine residues necessary

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

for dimerisation within the 60 AA deletion in the extracellular coiled-coil domain of tetherin B (Figure 4).

Given the differing extent of homodimer formation for each bat tetherin and reports that dimerisation is important for the restriction of some viruses and not required for others (58, 61), it is likely that these differences might affect the extent by which each tetherin is capable of restricting various VLPs. Tetherin function was validated through an assessment of the capacity of *P. alecto* tetherin isoforms A, B, and C, and *M. macropus* tetherins A and B to restrict the release of HIVΔVpu VLPs. While primate lentiviruses are not known to natively infect bats, bats have recently been discovered to host extant gammaretroviruses and deltaretroviruses, and bat genomes contain diverse endogenous retroviral sequences (27, 62-65). *P. alecto* tetherin isoforms A and B inhibited the release of HIVΔVpu VLPs, in contrast to isoform C which lacked restriction activity (Figure 7A). *M. macropus* tetherin A restricted HIVΔVpu VLP release, while tetherin B did not (Figure 7B). These findings are consistent with the expected effects of tetherin dimerisation on the restriction of HIVΔVpu VLPs (58).

Previous reports on the necessity of the tetherin GPI anchor for inhibition of viral particle release (28, 42) indicates that the lack of a GPI anchor on isoform C would likely result in an inability to inhibit viral egress from the host cell. However, sheep tetherin B, which does not possess a GPI-signal peptide (Figure 2B), is capable of limited restriction of betaretroviral VLPs through an unknown mechanism (51). Because of the unique sequence of the C-terminus and lack of GSP in *P. alecto* tetherin isoform C, we extended our evaluation of its restrictive capacity to include filoviral VLPs derived from Ebola and Marburg virus VP40 matrix proteins. Surprisingly, *P. alecto* tetherin isoform C was capable of restricting the release of Ebola VLPs to an extent similar to isoforms A and B, although it was marginally less restrictive of Marburg VLPs compared to isoforms A and B (Figure 13).

How isoform C could be capable of restricting the release of VLPs without a C-terminal GPI anchor is not presently known, although there are at least two possible explanations. The first is that the alternative C-terminal sequence of isoform C is involved in envelope or cellular membrane binding through an unknown mechanism. The second is that isoform C forms dimers in a manner distinct from that of isoforms A and B. One possibility is that it can form an antiparallel homodimer configuration in such a way that the N-terminus of one monomer aligns with the C-terminus of another monomer, forming a dimer with a transmembrane domain at each end of the protein. Another possibility is the formation of standard parallel dimers of tetherin followed by the formation of an antiparallel tetramer, similarly resulting in a quaternary structure possessing a transmembrane domain at either end. This possibility is supported by structural analyses demonstrating that human tetherin is capable of forming tetramers composed of antiparallel pairs of parallel dimers (57, 66).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Conclusions

Tetherin variants of the fruit bat *P. alecto* and the vesper bat *M. macropus* are functionally capable of restricting the release of VLPs in a mammalian cell culture system. These bats have evolved to express unique forms of tetherin that do not conform to the standard tetherin protein structure observed in tetherin proteins of other mammals. A structurally unique tetherin of *P. alecto*, isoform C, was not observed to restrict the release of HIVΔVpu VLPs but was capable of restricting filoviral VLPs. These findings raise questions regarding the possible antiviral range, mechanism of action, and *in vivo* function of structurally diverse forms of tetherin. Bat tetherin genes are under strong positive selection, indicating an ongoing process of evolution involving tetherin targets and viral antagonists. This evolutionary pressure has resulted in the expansion and diversification of tetherin genes in vesper bats of the genus *Myotis*, and the emergence of unique splice variants of tetherin in the fruit bat *P. alecto*, supporting the hypothesis of differential and unique antiviral gene evolution in bats.

Methods and Materials

Sequence read archive (SRA) BLAST analysis

To predict the tetherin sequences of other bat species, the *P. alecto* tetherin isoform A coding domain nucleotide sequence (549 nt) was used as the search query for a separate BLASTn analysis of each of the publicly accessible sequence read archives (SRA) of bat transcriptomes (Table 1). This approach allowed the individual assembly of a tetherin homolog from each bat species for which a transcriptome sequence read archive was available. The BLASTn analyses were performed using the online SRA Nucleotide BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The program selection was optimised for 'somewhat similar sequences' (BLASTn) and the selected algorithm parameters were as follows: Max target sequences = 1000; Expect threshold = 1.0×10^{-10}; Word size = 7; Max matches in a query range = 21; Match/Mismatch scores = 2,-3; Gap costs of existence 5 and extension 2; and no filtering or masking was selected. Matching reads were downloaded and assembled using the Assemble Sequences tool in CLC Genomics Workbench 8.0 (CLC; Qiagen, Germany). The consensus sequence of the assembly was then used as the search query for a second SRA BLAST analysis with the same parameters as the first search with the following exceptions: Program selection was optimised for 'highly similar sequences' (megablast); Expect threshold = 1.0×10^{-20}; Word size = 28; Match/Mismatch scores = 1,-2; and linear gap costs of existence and extension. Matching reads were downloaded and assembled in the same manner,

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and the assembled consensus sequence was used in a third SRA BLAST using the same parameters as the second. This process was iteratively repeated for the assembled consensus sequence of each bat tetherin until it extended through the tetherin coding domain in both directions into the 5' and 3' untranslated regions, respectively demarcated by the locations of the start methionine and stop codon.

Evolutionary selection test

To determine if evolutionary selective pressures were being applied to bat tetherin, a selection test was performed. To detect the positively selected sites among bat tetherin sequences, a maximum likelihood (ML) phylogeny was generated with CODEML implemented in PAML4 software (67). The input tree was generated based on a pre-existing bat species tree (68) as well as cytb phylogeny derived from MEGA6 (69) under the GTR+I+G nucleotide substitution model. The coding sequence alignments were fit to the NSsites models allowing (M8; positive selection model) or disallowing (M8a; null model) positive selection. Models were compared using a chi-squared test (degrees of freedom = 2) on twice the difference of likelihood values to derive P-values. The codon frequency model F3x4 was used. In cases where a significant difference ($P < 0.01$) between M8a versus M8 was detected, the Bayes Empirical Bayes (BEB) analysis was used to identify codons with ω (dN/dS) > 1, reporting values with posterior probability > 0.99 as high significance or 0.95 - 0.99 as moderate significance.

Identification of protein domains, structures, and motifs

To confirm the predicted tetherin nucleotide sequences and identify functional domains within bat tetherin protein sequences, translated coding domains were compared against known tetherin amino acid sequences including the human l-tetherin and the *P. alecto* tetherin isoform A (Figure 2). For all tetherin amino acid sequences, secondary structures were determined using the Predict Secondary Structure tool in CLC. The presence of cytoplasmic, transmembrane, and extracellular domains were determined using a hidden Markov model in TMHMM 2.0 (70). GPI signal sequences were determined using the online tool PredGPI (71) (http://gpcr2.biocomp.unibo.it/gpipe/index.htm). Pairwise comparisons between all tetherin sequences was conducted using the Create Pairwise Comparison tool in CLC. Multiple sequence alignments (MSA) of predicted and known tetherin amino acid and nucleotide sequences were performed using MUSCLE v3.8 (72) with the following parameters: Maximum iterations = 16, Find diagonals = No.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Transcriptome and contig analysis

Approval for the use of bat tissue was granted by the Australian Centre for Disease Preparedness (ACDP) (formerly the Australian Animal Health Laboratory, AAHL) Animal Ethics committee (Protocol AEC1281). The *P. alecto* transcriptome is accessible through the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/Traces/sra/) [SRA: SRP008674].

To identify the homologs of tetherin in *P. alecto*, the human l-tetherin protein sequence [UniProt: Q10589-1] was used as a query in a tBLASTn analysis of the transcriptome of *P. alecto*. To identify assembled sequences (contigs) representing mRNA transcripts of interest, local tBLASTn analyses of the transcriptome of *P. alecto* were conducted with CLC using the following parameters: BLOSUM62 matrix, word size = 3, E-values < $1x10^{-12}$, gap costs of existence 11, extension 1, with no filtering of regions of low complexity.

cDNA analysis

To amplify tetherin nucleotide sequences from bat cDNA, polymerase chain reaction (PCR) assays were performed using cDNA generated from *P. alecto* and *M. ricketti* spleen tissue, and *M. macropus* and *M. schreibersii* kidney cell lines, using various combinations of forward and reverse primers (Table 3A). Primers were designed using the tetherin predictions identified in the contig and SRA analyses. Primers were designed to bind to the 5' and 3' untranslated regions of bat cDNA such that the full CDS could be amplified. Bat capture, tissue collection and RNA extraction was conducted as previously reported (73) with the exception that RNAlater (Ambion, USA) preserved spleen tissue from four male adult bats was pooled before tissue homogenisation and followed with total RNA extraction with the Qiagen RNeasy Mini kit with on-column extraction of genomic DNA with DNase I. Total RNA was reverse transcribed into cDNA with the Qiagen Omniscript reverse transcriptase according to the manufacturer's protocol with the exception that the reaction contained 100 ng/µl total RNA, 1 µM oligo-dT18 (Qiagen) and 10 µM random hexamers (Promega).

All PCR amplification assays were performed using the Roche FastStart High Fidelity PCR system (Cat # 04738292001) with an annealing temperature gradient of 54°C to 64°C in 2°C increments. Each reaction was made up to a total of 20 µl, containing 1 unit of polymerase, 2 ng of total cDNA, and 8 pmol of each primer. All other parameters for PCR amplification were performed according to the manufacturer's protocol.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Amplicons from PCR reactions were analysed by agarose gel electrophoresis (74). Individual DNA bands were physically excised from the gel. DNA was purified from the gel fragments using the Wizard SV Gel and PCR Clean Up kit (Promega, Fitchburg, USA) according to the manufacturer's protocol.

To further analyse the PCR amplified DNA, each DNA fragment was blunt-end ligated into the pCR2.1-TOPO-TA or pCR-Blunt-II-TOPO plasmid vector (Invitrogen, Waltham, USA). Ligation was performed using the TOPO TA or Zero Blunt TOPO PCR Cloning Kit (Invitrogen) according to the manufacturer's instructions and plasmids were transformed into Top10 *E. coli* using a standard heat-shock method (74). Plasmids were purified from these cultures using a Wizard Plus SV Miniprep DNA Purification kit (Promega). All inserted sequences were confirmed by Sanger sequencing using M13 forward and reverse primers (M13F and M13R; Table 3B).

Genome Mapping

To identify tetherin genes within the genomes of *P. alecto* and *M. macropus*, local BLASTn analyses were performed using CLC. For these analyses the *P. alecto* and *M. macropus* tetherin sequences were used as query sequences against the *P. alecto* genome [NCBI: PRJNA171993] and *M. lucifugus* genome [NCBI: GCF_000147115]. BLASTn was performed using the following parameters: Word size = 11, E-values < $1x10^{-3}$, gap costs of existence 5, extension 2, with no filtering of regions of low complexity. Genes were delineated as beginning and ending at the first and last nucleotide of the contig query sequence. Exons were delineated as consisting of the nucleotide regions within the gene mapping to the tetherin cDNA sequences, bordered by the canonical 5'-AG and 3'-GT dinucleotide motifs (75). The 5' and 3' untranslated regions were defined as the regions upstream of the start methionine and downstream of the stop codon of the CDS, respectively. Contig sequences were mapped against gene scaffolds through a local BLASTn analysis using CLC Genomics Workbench with default settings. The nucleotide sequence preceding the first exon of each gene was assessed for the presence of promoter motifs.

Neighbouring genes were identified by BLAT (BLAST-like alignment tool) analysis (76) of the bordering nucleotide sequences upstream and downstream of tetherin genes using the Ensembl genome database BLAT tool (http://asia.ensembl.org/Multi/Tools/Blast?db=core).

Generation of tagged tetherin constructs for expression in mammalian cells

To enable detection of tetherin expression, the *P. alecto* tetherin isoforms A, B, and C, and *M. macropus* tetherin A and B were genetically modified with the insertion of nucleotide sequences encoding the

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

haemagglutinin (HA) antibody-binding epitope. Tetherin sequences were modified through a 2-step PCR process. PCR reactions were performed using the Roche FastStart HighFidelity PCR kit according to the manufacturer's recommendations.

To express tetherin proteins in a mammalian cell culture system the tagged tetherin inserts were sub-cloned from the pCR2.1-TOPO (*P. alecto* tetherin isoforms A, B, and C) and pCR-Blunt-II-TOPO (*M. macropus* tetherin A and B) vectors into pcDNA3.1 mammalian expression vectors. The HA-tagged tetherin constructs contained terminal enzyme restriction sites. *P. alecto* tetherin constructs contained *Xho*I and *Xba*I sites at their 5' and 3' ends, respectively. *M. macropus* tetherin constructs contained an *Eco*RI site at each end. These sites, which are also present in the pcDNA3.1 vector, were used for digestion-ligation transfer of the HA-tagged tetherin sequences into pcDNA3.1. The ligation reaction products were transformed into Top10 *E. coli* and plasmid clones were purified from *E. coli* colonies using the method described under 'cDNA analysis'. All inserted sequences were verified by Sanger sequencing using the pcDNA3.1 vector sequencing primers, T7F forward and BGHR reverse (Table 3B).

Expression, extraction, and detection of tetherin in a mammalian cell culture system

To determine if the tetherin plasmids can express tetherin protein in a mammalian cell culture system, adherent human embryonic kidney HEK293T cells (kindly provided by Richard Axel, Columbia University) were transfected with each tetherin construct with protein expression determined by Western blot analysis.

HEK293T cells were maintained in Dulbecco's modified Eagle medium (DMEM-10, Thermo Fisher) enriched with heat-inactivated foetal calf serum (100ml/L; Invitrogen), glutamine (292 µg/ml; Invitrogen), and the antibiotics penicillin (100 units/ml; Invitrogen) and streptomycin (100 units/ml; Invitrogen) (DMEM-10). Cells were incubated at 37°C with 5% CO_2.

The expression was performed in 6-well plates. Each well was seeded with 3.0×10^5 cells/well in 2 ml of DMEM-10. Cells were transfected when the monolayer had reached 50-60% confluency. The tetherin constructs analysed are listed in Table 4. Each plasmid construct was transfected into cells in duplicate wells at 2 µg/well using the transfection reagent Lipofectamine 2000 (Thermo Fisher) according to the manufacturer's protocol. Tetherin was extracted using a previously published GPI-anchored protein extraction protocol (48).

Protein samples were analysed by size-based separation through SDS-PAGE and visualisation using Western blot analysis (74). For the Western blot analysis, the primary antibody solution contained a

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1/1000 dilution of a monoclonal rabbit anti-HA antibody (C29F4, Cell Signaling Technology, Danvers, USA) in TBS containing 0.1% Tween-20, and the secondary antibody solution contained a 1/10,000 dilution of a polyclonal goat anti-rabbit IRD800 fluorophore-conjugate secondary antibody (Rockland, USA). The primary antibody solution was incubated overnight at 4°C, and the secondary antibody solution was incubated at room temperature for 1 h. To visualise fluorescent antibody-bound tetherin proteins, membranes were scanned using the Odyssey Imaging System (LI-COR Biosciences, Lincoln, USA) at wavelengths of 600 and 800 nm, using the default software settings.

Fluorescence microscopy

To visualise tetherin localisation within cells, 500 ng of plasmids encoding HA-tagged human l-tetherin and *P. alecto* tetherin isoforms A, B, and C were transfected into HEK293T cells seeded on glass coverslips as described above. At 48 h post transfection, cells were fixed with 4% paraformaldehyde (Sigma) in PBS for 10 min at room temperature, and then permeabilised in 0.2% Triton X-100 in PBS for 5 min at room temperature. Tetherin localisation in cells was detected by staining cells with anti-HA-tag rabbit monoclonal IgG (Thermo Fisher) diluted in 0.2% Triton X-100, 3% BSA in PBS for 1 h at room temperature. Cells were subsequently stained with anti-rabbit AlexaFlour 488 secondary IgG (Thermo Fisher) diluted in 0.2% Triton X-100, 3% BSA in PBS for 20 min at room temperature. Nuclei are blue stained with Hoechst 33342 (Thermo Fisher) according to the manufacturer's instructions. Coverslips were mounted onto glass slides with ProLong Gold Antifade Mountant (Thermo Fisher), then imaged on a Nikon AR1 confocal microscope and analysed with ImageJ (NIH) software.

qPCR analysis of tetherin expression across multiple bat tissues

To assess tetherin expression across various *P. alecto* tissues, tetherin mRNA levels were measured by qPCR analysis. The primers and probes were designed using the program Primer Express (Perkin–Elmer, Applied Biosystems, USA). The tetherin primers amplify all three known isoforms of *P. alecto* tetherin. *P. alecto* bats were trapped in Queensland, Australia, and transported alive by air to the ACDP in Victoria, where they were euthanised for dissection using methods approved by the ACDP animal ethics committee (AEC1389). Tissues were stored at −80 ◦C in RNAlater (Ambion). Total RNA was extracted from frozen *P. alecto* tissues using a Precellys 24 tissue homogeniser (Bertin Technologies, France) and an RNeasy mini kit (Qiagen) with on-column DNase-I treatment (Qiagen) to remove traces of genomic DNA.

Total RNA was subjected to real-time PCR using SuperScript III One-Step RT-PCR System (Invitrogen) with Platinum *Taq* DNA Polymerase. Reactions were performed on 100 ng of template RNA with 200 nM of

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

each primer and 150 mM of the Taqman probe in an Applied Biosystems 7500 Fast Real-Time qPCR instrument. Cycling parameters consisted of 50°C, 5 min for the reverse transcription of RNA to cDNA followed by 95°C for 2 min. The cDNA was amplified by PCR for 40 cycles, each consisting of 95°C for 3 s and 60°C, 30 s.

Using the Livak/ΔΔ Ct method (77), tetherin Ct values were normalised against the expression/Ct values of the 18S rRNA housekeeping gene and reported as fold-difference, calibrated against the Ct values for wing tissue.

Transcriptome analysis of isoform expression under immune-stimulating treatments

To compare the expression of alternative isoforms of *P. alecto* tetherin and the impact of treatment with immune-stimulating compounds, bats were treated with PBS, LPS (Invivogen, #tlrl-pb5lps) or poly(I:C) (Invivogen, #vac-pic) as published previously (46). Briefly, 5 h post-intraperitoneal injection, bats were anaesthetised, culled and organs were processed for RNA, DNA, protein and cell suspensions as described previously (78). RNA libraries were prepared using RiboZero Plus rRNA-depletion kits (Illumina, USA) and cDNA was generated using a mix of oligo-dT/random hexamer primers, prior to sequencing in 2x150PE on the Illumina HiSeq platform.

Sequencing read libraries were quality controlled by analysis using FastQC (79). Illumina sequence adapters were removed and reads were trimmed or discarded on the basis of quality using the Trim Sequences tool in CLC. Overlapping paired reads were merged using the Merge Overlapping Pairs tool in CLC. Using the RNA-Seq Analysis tool in CLC, sequence reads were mapped against the *P. alecto* gene scaffold containing the tetherin gene (GenBank accession: KB030270.1), which was manually annotated with the tetherin gene and mRNA sequences for isoforms A, B, and C. The following parameters were used for the RNA-Seq analysis: mismatch = 3, insertion = 3, deletion = 3, length fraction = 0.6, similarity fraction = 0.95; default parameters were used otherwise. Isoform expression levels were normalised for post-trimming read library size and reported as counts-per-million reads (CPM). Non-parametric one-tailed Mann-Whitney tests were performed to calculate statistical significance between treatments. The *P. alecto* tetherin gene scaffold, annotation files, and read maps are provided in SI Data 2.

Functional validation of tetherin activity

To assess the effect of bat tetherin on the release of virus-like particles (VLPs), HEK293T cells were co-transfected with each tagged bat tetherin construct and either a HIVΔVpu plasmid construct, pCRV1-NLgagpolΔVpu, which expresses HIV NL4.3 Gag-polΔVpu protein [NCBI: P12497.4]; an EBOV VP40

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

plasmid, pGFPEVP40, expressing the Ebola Zaire virus VP40 Matrix protein [NCBI: NP_066245]; or a MARV VP40 plasmid, pGFPEMVP40, expressing the Marburg virus VP40 Matrix protein [NCBI: YP_001531155]. EBOV and MARV VP40 proteins were encoded and expressed as VP40-eGFP fusion proteins. VLP plasmid constructs were kindly provided by Paul Bieniasz (Aaron Diamond AIDS Research Center, Rockefeller University) (80).

Transfections were performed as described above, with the following modifications: 12-well plates were seeded with $1.5x10^5$ cells/well in a total volume of 1 ml of DMEM-10, DNA-Lipofectamine mixtures for each well contained 1 µl of Lipofectamine, and a total of 400 ng of DNA comprised of 200 ng of VLP plasmid DNA, and 0 – 200 ng of tetherin-encoding plasmid DNA, made up to a total mass of 200 ng with insert-free pcDNA3.1 plasmid DNA. Each sample was prepared in duplicate wells. Cells were incubated at 37°C with 5% CO_2 for a total of 48 h. After 24 h, 1 ml of DMEM-10 was added to each well. Following incubation cell lysates and supernatants were collected for further analysis.

Cell culture supernatants were collected and clarified. Clarified supernatants were layered over 2 ml of a 25% (w/v) sucrose solution in 13.2 ml Thinwall Polypropylene SW41 Ti ultracentrifuge tubes (Beckman-Coulter, Brea, USA) and made up to a total volume of 10 ml with PBS without calcium chloride or magnesium chloride (PBS[-]). Samples were then centrifuged at 130,000 x *g* for 2 h using an Optima L-100 XP ultracentrifuge (Beckmann-Coulter). VLP pellets were lysed in 60 µl of NP-40 lysis buffer (Sigma-Aldrich), supplemented with 10 µg/ml of aprotinin, leupeptin, and pepstatin A protease inhibitors (Sigma-Aldrich).

The cells from each well were resuspended in 500 µl of PBS[-] and transferred to 1.5 ml tubes and centrifuged at 200 x *g* for 5 min. Cell pellets were lysed in 100 µl of NP-40 lysis buffer. Following lysis samples were centrifuged at 20,000 x *g* for 5 min, supernatants were transferred to a new 1.5 ml tubes.

Cell and viral lysates were analysed by SDS-PAGE and Western blot analysis as described above, with the following exceptions. For HIV VLPs, the primary antibody binding was directed against the Gag protein using a solution containing a 1/4000 dilution of a mouse anti-p24 antibody (NIH AIDS Reagent Repository, Germantown, USA), 0.1% Tween-20, in PBS[-]. For Ebola and Marburg VLPs, the primary antibody binding was directed against eGFP using a solution containing a 1/4000 dilution of a mouse anti-GFP 4B10 antibody (Cell Signaling Technology, USA). For all VLPs, the secondary antibody was a 1/10,000 diluted polyclonal goat anti-mouse Alexa-Fluor 680 fluorophore-conjugate fluorescent secondary antibody (Thermo Fisher).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

VLPs released into the supernatant for each sample were quantified using a densitometric analysis of the signal strength of the fluorescent antibodies bound to VLP proteins. Densitometry was performed using the Image Studio Lite v5.2 software suite (LI-COR). Variation in viral lysate signal strength across samples was normalised using the signal strength of the cell lysates. The non-parametric Wilcoxon Rank Sum test was performed to calculate the statistical significance of the restriction of Ebola and Marburg VLPs in the 200 ng tetherin isoform C treatment groups.

Data availability

The data underlying this article are available in this article and its online supplementary material. Novel tetherin sequences have been deposited in GenBank, at https://www.ncbi.nlm.nih.gov/genbank/ with accession numbers listed in supplementary table 1.

Acknowledgments

We thank Paul Bieniasz for providing the pCRV1-NLgagpolΔVpu plasmid construct. We thank Richard Axel for providing HEK293T cells. 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 and a Monash University Vice Chancellor's Honours-PhD scholarship. L-F.W. is funded by the Singapore National Research Foundation grants NRF2012NRF-CRP001-056 and NRF2016NRF-NSFC002-013. The authors gratefully acknowledge the contribution to this work of the Victorian Operational Infrastructure Support Program received by the Burnet Institute.

Author's Contributions

J.A.H., M.T., L-F.W., and G.T. designed the study; J.A.H., J.C., A.T.I, A.N., I.S., T.B.G. and A.J., performed research and analysed data; J.A.H., M.T., A.J., A.N., V.B., G.A.M., M.L.B., L-F.W., and G.T. discussed and interpreted the results; J.A.H. wrote the first draft of the paper; J.A.H., and G.T. critically edited the manuscript.

References

1. C. H. Calisher, J. E. Childs, H. E. Field, K. V. Holmes, T. Schountz, Bats: Important reservoir hosts of emerging viruses. *Clin. Microbiol. Rev.* **19**, 531-545 (2006).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

2. I. Smith, L.-F. Wang, Bats and their virome: an important source of emerging viruses capable of infecting humans. *Curr. Opin. Virol.* **3**, 84-91 (2013).

3. B. R. Amman *et al.*, A recently discovered pathogenic paramyxovirus, Sosuga virus, is present in *Rousettus aegyptiacus* fruit bats at multiple locations in Uganda. *J. Wildl. Dis.* **51**, 774-779 (2015).

4. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 1-4 (2020).

5. M. Letko, S. N. Seifert, K. J. Olival, R. K. Plowright, V. J. Munster, Bat-borne virus diversity, spillover and emergence. *Nat. Rev. Microbiol.*, 1-11 (2020).

6. R. Swanepoel, Experimental inoculation of plants and animals with Ebola virus. *Emerging Infect. Dis.* **2**, 321 (1996).

7. M. M. Williamson *et al.*, Transmission studies of Hendra virus (equine morbilli-virus) in fruit bats, horses and cats. *Aust. Vet. J.* **76**, 813-818 (1998).

8. D. J. Middleton *et al.*, Experimental Nipah virus infection in pteropid bats (*Pteropus poliocephalus*). *J. Comp. Pathol.* **136**, 266-272 (2007).

9. S. Watanabe *et al.*, Bat coronaviruses and experimental infection of bats, the Philippines. *Emerging Infect. Dis.* **16**, 1217-1223 (2010).

10. M. E. Jones *et al.*, Experimental inoculation of Egyptian rousette bats (*Rousettus aegyptiacus*) with viruses of the Ebolavirus and Marburgvirus genera. *Viruses* **7**, 3420-3442 (2015).

11. M. L. Baker, T. Schountz, L. F. Wang, Antiviral immune responses of bats: A review. *Zoonoses Public Hlth.* 10.1111/j.1863-2378.2012.01528.x, 1-13 (2013).

12. G. Zhang *et al.*, Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science* **339**, 456-460 (2013).

13. P. Zhou *et al.*, Contraction of the type I IFN locus and unusual constitutive expression of IFN-α in bats. *Proc. Natl. Acad. Sci. USA*, 201518240 (2016).

14. J. A. Hayward *et al.*, Differential evolution of antiretroviral restriction factors in pteropid bats as revealed by *APOBEC3* gene complexity. *Mol. Biol. Evol.* **35**, 1626-1637 (2018).

15. S. S. Pavlovich *et al.*, The Egyptian rousette genome reveals unexpected features of bat antiviral immunity. *Cell* **173**, 1098-1110. e1018 (2018).

16. J. N. Mandl, C. Schneider, D. S. Schneider, M. L. Baker, Going to bat(s) for studies of disease tolerance. *Front. Immunol.* **9**, 2112 (2018).

17. J. A. Hawkins *et al.*, A metaanalysis of bat phylogenetics and positive selection based on genomes and transcriptomes from 18 species. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 11351-11360 (2019).

18. M. Hölzer *et al.*, Virus-and interferon alpha-induced transcriptomes of cells from the microbat *Myotis daubentonii*. *iScience* **19**, 647-661 (2019).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

19. J. H. Morrison *et al.*, A potent postentry restriction to primate lentiviruses in a yinpterochiropteran bat. *Mbio* **11** (2020).

20. D. Jebb *et al.*, Six reference-quality genomes reveal evolution of bat adaptations. *Nature* **583**, 578-584 (2020).

21. S. J. D. Neil, T. Zang, P. D. Bieniasz, Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature* **451**, 425-430 (2008).

22. T. Sakuma, T. Noda, S. Urata, Y. Kawaoka, J. Yasuda, Inhibition of Lassa and Marburg virus production by tetherin. *J. Virol.* **83**, 2382-2385 (2009).

23. M. Hoffmann *et al.*, Tetherin inhibits Nipah virus but not Ebola virus replication in fruit bat cells. *J. Virol.* **93** (2019).

24. J. W. Wynne, L.-F. Wang, Bats and viruses: Friend or foe? *PLoS Pathog.* **9**, e1003651 (2013).

25. S. M. Wang, K. J. Huang, C. T. Wang, Severe acute respiratory syndrome coronavirus spike protein counteracts BST2-mediated restriction of virus-like particle release. *J. Med. Virol.* **91**, 1743-1750 (2019).

26. L.-F. Wang, D. E. Anderson, Viruses in bats and potential spillover to animals and humans. *Curr. Opin. Virol.* **34**, 79-89 (2019).

27. J. A. Hayward *et al.*, Infectious KoRV-related retroviruses circulating in Australian bats. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 9529-9536 (2020).

28. D. Perez-Caballero *et al.*, Tetherin inhibits HIV-1 release by directly tethering virions to cells. *Cell* **139**, 499-511 (2009).

29. D. Blanco-Melo, S. Venkatesh, P. D. Bieniasz, Origins and evolution of tetherin, an orphan antiviral gene. *Cell Host Microbe* **20**, 189-201 (2016).

30. L. J. Cocka, P. Bates, Identification of alternatively translated tetherin isoforms with differing antiviral and signaling activities. *PLoS Pathog.* **8**, e1002931 (2012).

31. Rui P. Galão, A. Le Tortorec, S. Pickering, T. Kueck, Stuart J. D. Neil, Innate sensing of HIV-1 assembly by Tetherin induces NFκB-dependent proinflammatory responses. *Cell Host Microbe* **12**, 633-644 (2012).

32. A. Tokarev *et al.*, Stimulation of NF-κB activity by the HIV restriction factor BST2. *J. Virol.* **87**, 2046-2057 (2013).

33. A. Matsuda *et al.*, Large-scale identification and characterization of human genes that activate NF-κB and MAPK signaling pathways. *Oncogene* **22**, 3307-3318 (2003).

34. J. Ishikawa *et al.*, Molecular cloning and chromosomal mapping of a bone marrow stromal cell surface gene, BST2, that may be involved in pre-B-cell growth. *Genomics* **26**, 527-534 (1995).

35. R. A. Liberatore, P. D. Bieniasz, Tetherin is a key effector of the antiretroviral activity of type I interferon in vitro and in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 18097-18101 (2011).

36. E. Erikson *et al.*, In vivo expression profile of the antiviral restriction factor and tumor-targeting antigen CD317/BST-2/HM1. 24/tetherin in humans. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 13688-13693 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

37. H. Yang *et al.*, Structural insight into the mechanisms of enveloped virus tethering by tetherin. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 18428-18432 (2010).

38. A. Hinz *et al.*, Structural basis of HIV-1 tethering to membranes by the BST-2/tetherin ectodomain. *Cell Host Microbe* **7**, 314-323 (2010).

39. S. Kupzig *et al.*, Bst-2/HM1. 24 is a raft-associated apical membrane protein with an unusual topology. *Traffic* **4**, 694-709 (2003).

40. R. Rollason, V. Korolchuk, C. Hamilton, P. Schu, G. Banting, Clathrin-mediated endocytosis of a lipid-raft-associated protein is mediated through a dual tyrosine motif. *J. Cell Sci.* **120**, 3850-3858 (2007).

41. E. Takeda *et al.*, Identification and functional analysis of three isoforms of bovine BST-2. *PLoS One* **7**, e41483 (2012).

42. Y. Iwabu *et al.*, HIV-1 accessory protein Vpu internalizes cell-surface BST-2/tetherin through transmembrane interactions leading to lysosomes. *J. Biol. Chem.* **284**, 35060-35072 (2009).

43. A. Fukuma, M. Abe, Y. Morikawa, T. Miyazawa, J. Yasuda, Cloning and characterization of the antiviral activity of feline Tetherin/BST-2. *PloS one* **6** (2011).

44. M. Ruedi *et al.*, Molecular phylogenetic reconstructions identify East Asia as the cradle for the evolution of the cosmopolitan genus *Myotis* (Mammalia, Chiroptera). *Mol. Phylogen. Evol.* **69**, 437-449 (2013).

45. M. J. Benton, P. C. J. Donoghue, Paleontological evidence to date the tree of life. *Mol. Biol. Evol.* **24**, 26-53 (2007).

46. P. Periasamy *et al.*, Studies on B cells in the fruit-eating black flying fox (*Pteropus alecto*). *Front. Immunol.* **10** (2019).

47. L. Mok *et al.*, Proteomic analysis of *Pteropus alecto* kidney cells in response to the viral mimic, Poly I:C. *Proteome Sci.* **13**, 1-11 (2015).

48. T. L. Doering, P. T. Englund, G. W. Hart, Detection of glycophospholipid anchors on proteins. *Curr. Protoc. Protein. Sci.*, 12.15. 11-12.15. 14 (2001).

49. M. Dubé *et al.*, Suppression of Tetherin-restricting activity upon human immunodeficiency virus type 1 particle release correlates with localization of Vpu in the trans-Golgi network. *J. Virol.* **83**, 4574-4590 (2009).

50. I. Dietrich *et al.*, Feline tetherin efficiently restricts release of feline immunodeficiency virus but not spreading of infection. *J. Virol.* **85**, 5840-5852 (2011).

51. F. Arnaud *et al.*, Interplay between ovine bone marrow stromal cell antigen 2/tetherin and endogenous retroviruses. *J. Virol.* **84**, 4415-4425 (2010).

52. L. F. Wang, D. E. Anderson, Viruses in bats and potential spillover to animals and humans. *Curr. Opin. Virol.* **34**, 79-89 (2019).

53. E. C. Teeling *et al.*, Bat biology, genomes, and the Bat1K project: to generate chromosome-level genomes for all living bat species. *Annu. Rev. Anim. Biosci.* **6**, 23-46 (2018).

54. T. J. O'Shea *et al.*, Bat flight and zoonotic viruses. *Emerging Infect. Dis.* **20**, 741-745 (2014).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

55. J. Kacprzyk *et al.*, A potent anti-inflammatory response in bat macrophages may be linked to extended longevity and viral tolerance. *Acta Chiropt.* **19**, 219-228 (2017).

56. M. Celestino *et al.*, Feline tetherin is characterized by a short N-terminal region and is counteracted by the feline immunodeficiency virus envelope glycoprotein. *J. Virol.* **86**, 6688-6700 (2012).

57. H. L. Schubert *et al.*, Structural and functional studies on the extracellular domain of BST2/tetherin in reduced and oxidized conformations. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 17951-17956 (2010).

58. A. J. Andrew, E. Miyagi, S. Kao, K. Strebel, The formation of cysteine-linked dimers of BST-2/tetherin is important for inhibition of HIV-1 virus release but not for sensitivity to Vpu. *Retrovirology* **6**, 80 (2009).

59. M. W. McNatt *et al.*, Species-specific activity of HIV-1 Vpu and positive selection of tetherin transmembrane domain variants. *PLoS Pathog* **5**, e1000300 (2009).

60. S. Venkatesh, P. D. Bieniasz, Mechanism of HIV-1 virion entrapment by tetherin. *PLoS Pathog.* **9**, e1003483 (2013).

61. T. Sakuma, A. Sakurai, J. Yasuda, Dimerization of tetherin is not essential for its antiviral activity against Lassa and Marburg viruses. *PLoS One* **4**, e6934 (2009).

62. J. A. Hayward *et al.*, Identification of diverse full-length endogenous betaretroviruses in megabats and microbats. *Retrovirology* **10**, 35 (2013).

63. E. C. Skirmuntt, M. Escalera-Zamudio, E. C. Teeling, A. Smith, A. Katzourakis, The potential role of endogenous viral elements in the evolution of bats as reservoirs for zoonotic viruses. *Annu. Rev. Virol.* **7**, 103-119 (2020).

64. J. Cui *et al.*, Identification of diverse groups of endogenous gammaretroviruses in mega and microbats. *J. Gen. Virol.* **93**, 2037-2045 (2012).

65. B. M. Hause, E. A. Nelson, J. Christopher-Hennings, North American big brown bats (*Eptesicus fuscus*) harbor an exogenous deltaretrovirus. *mSphere* **5** (2020).

66. M. Swiecki *et al.*, Structural and biophysical analysis of BST-2/tetherin ectodomains reveals an evolutionary conserved design to inhibit virus release. *J. Biol. Chem.* **286**, 2987-2997 (2011).

67. Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586-1591 (2007).

68. E. C. Teeling *et al.*, A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* **307**, 580-584 (2005).

69. K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725-2729 (2013).

70. A. Krogh, B. Larsson, G. Von Heijne, E. L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567-580 (2001).

71. A. Pierleoni, P. Martelli, R. Casadio, PredGPI: a GPI-anchor predictor. *BMC Bioinformatics* **9**, 1 (2008).

72. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nuc Acid Res* **32**, 1792-1797 (2004).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

73. C. Cowled *et al.*, Molecular characterisation of Toll-like receptors in the black flying fox *Pteropus alecto*. *Dev. Comp. Immunol.* **35**, 7-18 (2011).

74. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular cloning* (Cold spring harbor laboratory press New York, 1989), vol. 2.

75. T. A. Brown, "Synthesis and processing of RNA" in Genomes. (Oxford: Wiley-Liss, 2002).

76. W. J. Kent, BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656-664 (2002).

77. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. *Methods* **25**, 402-408 (2001).

78. A. T. Irving *et al.*, Optimizing dissection, sample collection and cell isolation protocols for frugivorous bats. *Methods Ecol. Evol.* **11**, 150-158 (2020).

79. S. Andrews (2010) FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

80. N. Jouvenet *et al.*, Broad-spectrum inhibition of retroviral and filoviral particle release by tetherin. *J. Virol.* **83**, 1837-1844 (2009).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figures and Legends

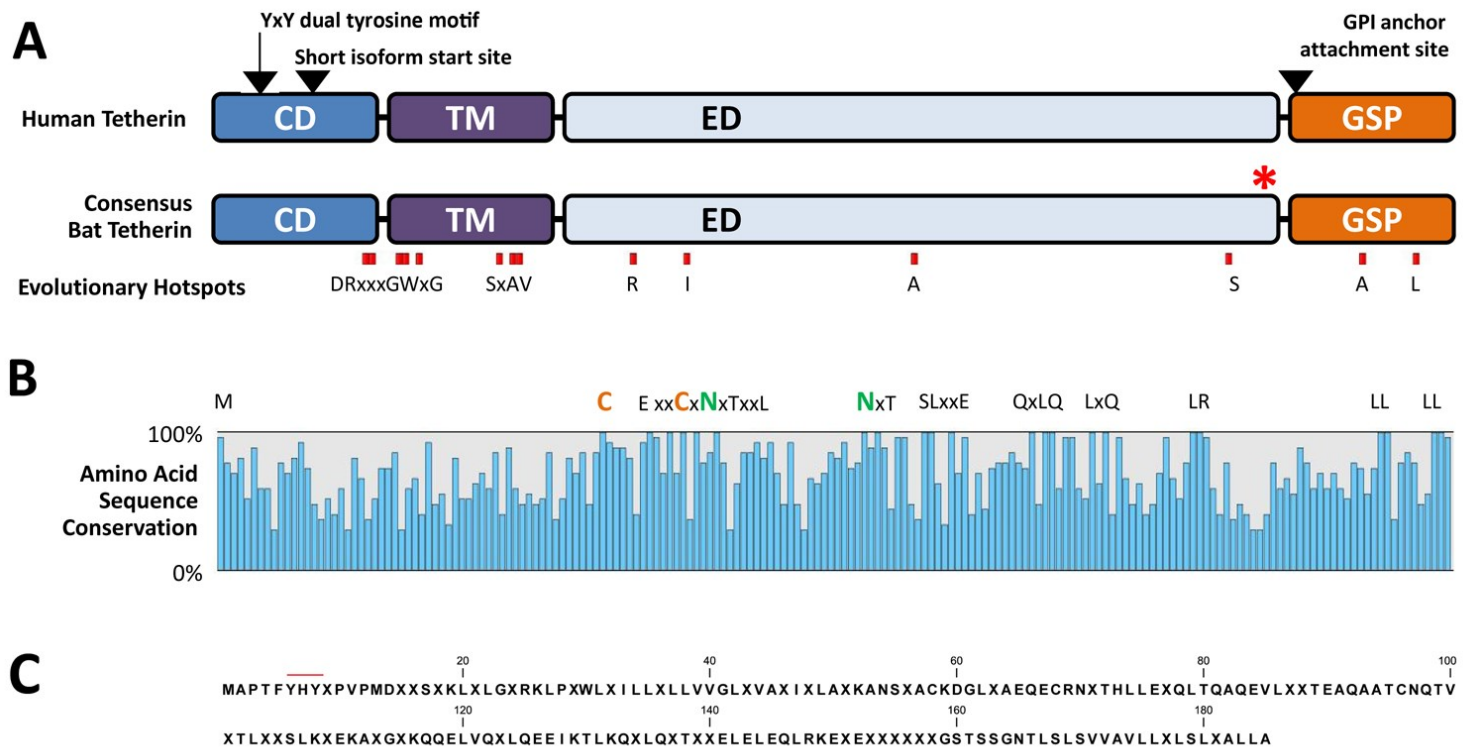

Figure 1. Bat Tetherin structure and sequence diversity. A. The consensus bat tetherin amino acid sequence structures and motifs generated through the multiple sequence alignment of tetherin from 27 bat species (SI Data 1) and compared against human tetherin. Strong positive selection of bat tetherin is revealed by the ratio of non-synonymous to synonymous mutations (Table 2), and is represented as evolutionary hotspots. The YxY dual tyrosine motif, the alternative start site for the short isoform of tetherin, and the GPI anchor attachment site positions are indicated by arrows. The red asterisk (*) indicates the location of the inserted molecular tags in the *P. alecto* and *M. macropus* tetherin expression constructs. CD, cytoplasmic domain; TM, transmembrane domain; ED, extracellular domain; GPI, glycophosphatidylinositol; GSP, GPI signal peptide. **B.** Significant sequence diversity exists among bat tetherin proteins, indicated by the percentage of amino acid sequence conservation at each site of the consensus bat tetherin. Amino acids conserved in all 27 sequences are represented by their letters. Amino acids represented by X are variable and are included to indicate the sequence distance between closely positioned conserved residues. **C.** The consensus bat tetherin sequence. Amino acid residues represented in > 50% of bat tetherin sequences are indicated with their letter. Positions at which no amino acid residue is represented in > 50% of bat tetherin sequences are indicated with 'X'. The red line indicates the position of the dual tyrosine motif.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

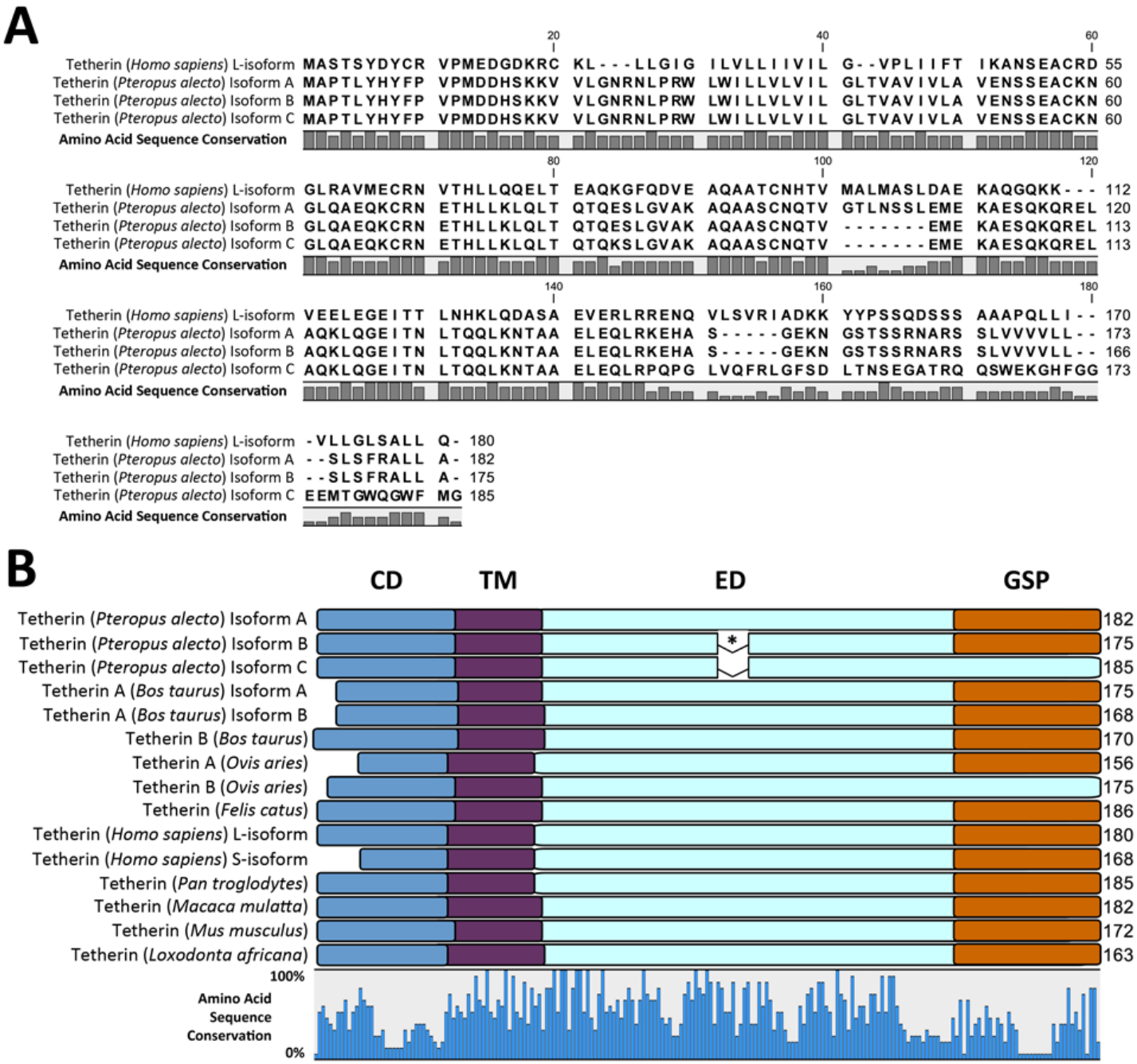

Figure 2. Multiple sequence alignment (MSA) and protein domains of mammalian tetherin amino acid sequences. A. MSA reveals amino acid residues differing between human and *P. alecto* sequences which are indicated by the amino acid sequence conservation bar graph below the sequences. **B.** The protein domains of mammalian tetherin proteins are depicted as overlays of an MSA. Sequence conservation is indicated by the bar graph below the alignment. Protein domains are colour coded: CD (blue), cytoplasmic domain; TM (purple), transmembrane domain; ED (light blue), extracellular domain; GSP (orange), glycophosphatidylinositol signal peptide; * a 7 AA exclusion in *P. alecto* tetherin isoforms B and C relative to isoform A is indicated.

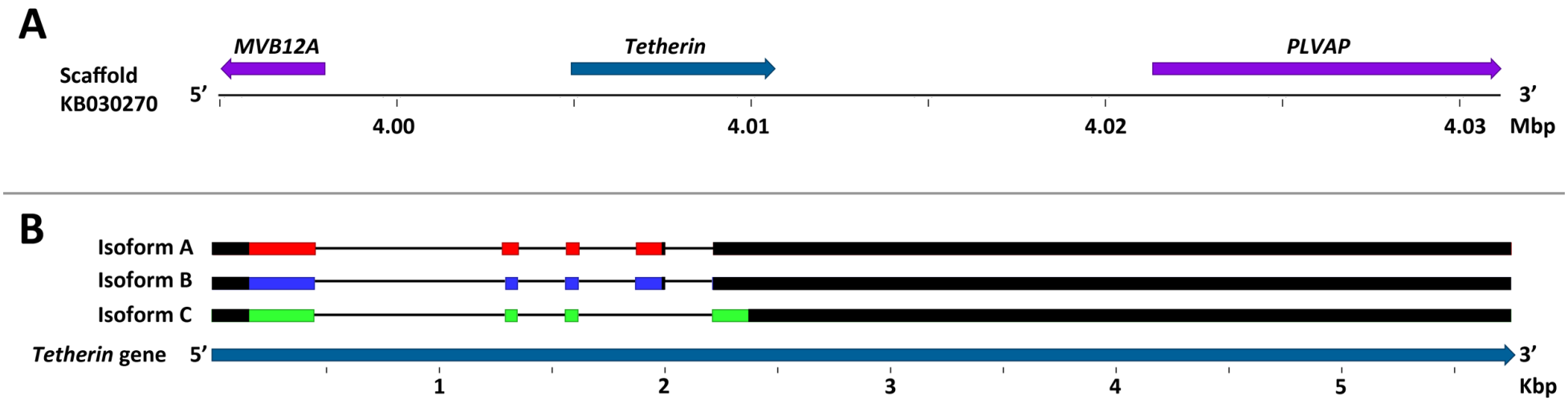

Figure 3. Mapping the tetherin gene to the *P. alecto* genome. A. The tetherin gene location and neighbouring genes, coloured purple, are depicted. Arrows indicate the orientation of each gene's protein coding domain sequence. Sequence numbers and gene orientations are relative to the beginning of the gene scaffold in the 5' to 3' direction. **B.** The tetherin mRNA exons are depicted and colour coded: Isoform A, red; Isoform B, blue; Isoform C, green. Exons are mapped to scale against the tetherin gene. Sequence numbers are relative to the beginning of the gene. The 5' and 3' untranslated regions of the mRNA are coloured black.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

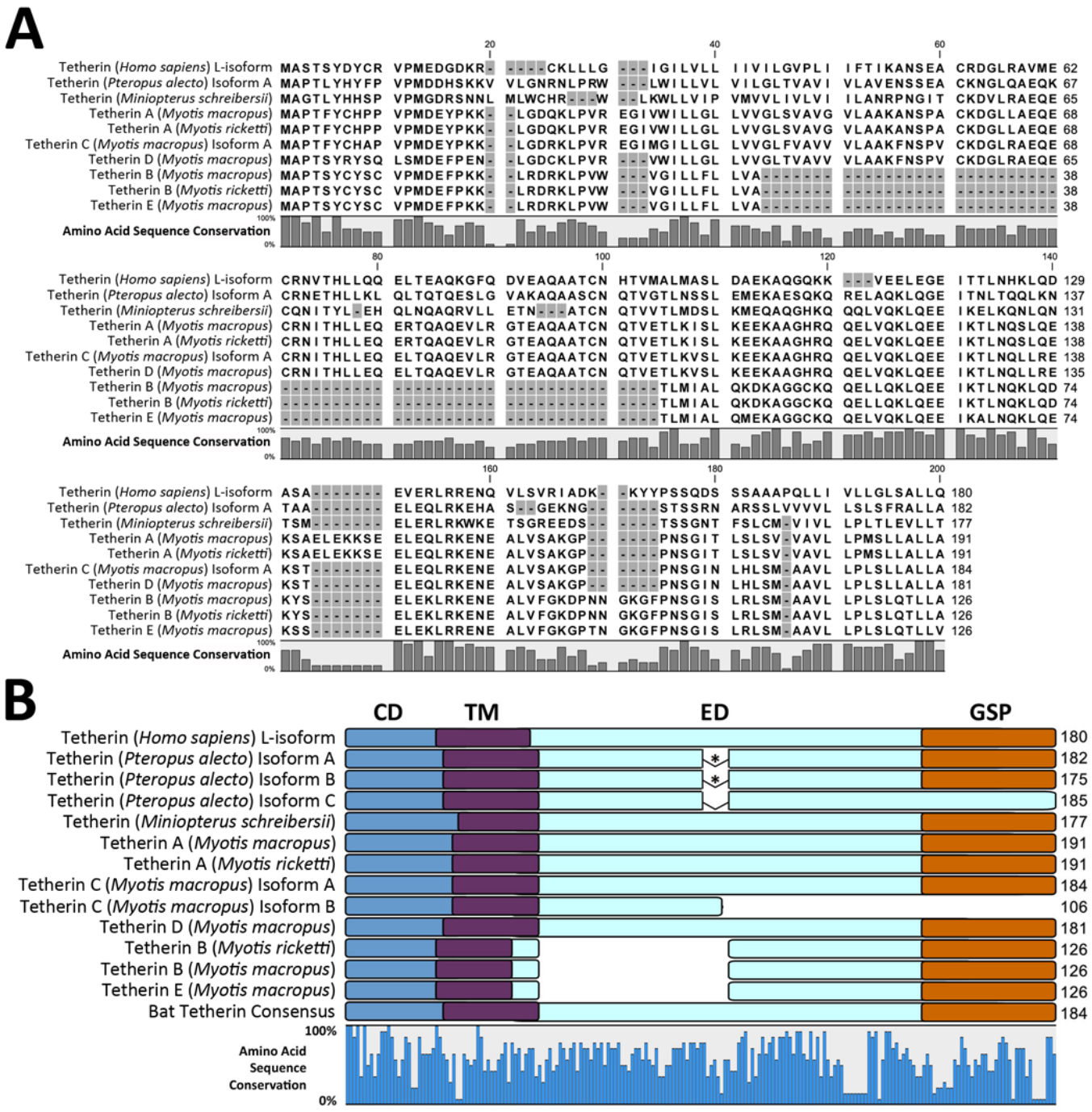

Figure 4. Multiple sequence alignment (MSA) and protein domains of fruit bat and vesper bat tetherin amino acid sequences. A. The percentage of sequence conservation is indicated by a bar plot below the MSA. Gap positions are shaded in grey. **B.** The protein domains of bat tetherin proteins are depicted as overlays of an MSA. Amino acid sequence conservation is indicated by the bar graph below the alignment. Protein domains are indicated and colour coded as follows: CD (blue), cytoplasmic domain; TM (purple), transmembrane domain; ED, (light blue) extracellular domain; GSP (orange), glycophosphatidylinositol signal peptide. * The 7 AA exclusion in *P. alecto* tetherin isoform B and C relative to isoform A is indicated.

A

		1	2	3	4	5	6	7
Tetherin (*Homo sapiens*) L-isoform	1		37.50	34.59	43.55	43.55	43.24	43.17
Tetherin (*Pteropus alecto*) Isoform A	2	115		43.48	53.23	53.23	53.51	52.46
Tetherin (*Miniopterus schreibersii*)	3	121	104		49.73	49.73	49.46	49.18
Tetherin A (*Myotis macropus*)	4	105	87	93		100.00	89.13	84.24
Tetherin A (*Myotis ricketti*)	5	105	87	93	0		89.13	84.24
Tetherin C (*Myotis macropus*) Isoform A	6	105	86	93	20	20		91.26
Tetherin D (*Myotis macropus*)	7	104	87	93	29	29	16	

B

		1	2	3	4	5	6
Tetherin (*Homo sapiens*) L-isoform	1		32.03	31.01	37.50	37.50	39.06
Tetherin (*Pteropus alecto*) Isoform A	2	87		42.40	44.62	44.62	44.62
Tetherin (*Miniopterus schreibersii*)	3	89	72		40.77	40.77	42.31
Tetherin B (*Myotis macropus*)	4	80	72	77		100.00	92.06
Tetherin B (*Myotis ricketti*)	5	80	72	77	0		92.06
Tetherin E (*Myotis macropus*)	6	78	72	75	10	10	

Figure 5. Pairwise comparisons between human and bat tetherin variants. **A.** *M. macropus* tetherin A, C, and D. **B.** *M. macropus* tetherin B and E analysed separately to A, C and D to account for the inclusion of 60 AA in *M. macropus* tetherin A, C, and D that is otherwise removed following gap treatment of the alignment of *M. macropus* B and E. Protein sequences of human and bat tetherins were aligned and gaps were removed. The horizontal axis represents percentage identity between pairs of amino acid sequences, coloured with a white to red gradient indicating increasing similarity as red intensity increases. The vertical axis represents the number of amino acid residue identity differences between sequence pairs, coloured with a white to blue gradient indicating increasing difference as blue intensity increases.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

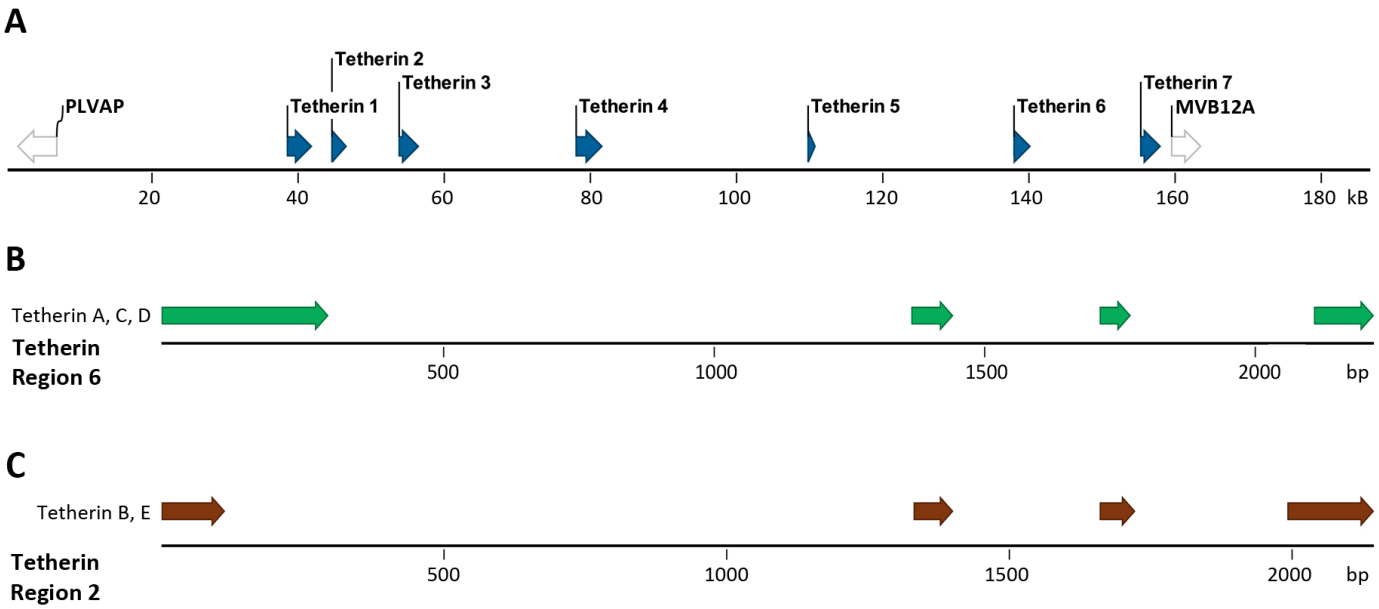

Figure 6. Mapping the tetherin genes of *M. macropus* to the *M. lucifugus* genome. Tetherin sequences derived from *M. macropus* cDNA were mapped against the publicly available genome of *M. lucifugus* (Myoluc2.0 genome assembly). **A.** The tetherin gene locus of *M. lucifugus* is contained within scaffold GL430608 between the PLVAP and MVB12A genes. **B.** *M. macropus* tetherin A, C, and D mapped against *M. lucifugus* tetherin region 6. **C.** *M. macropus* tetherin B and E mapped against *M. lucifugus* tetherin region 2.

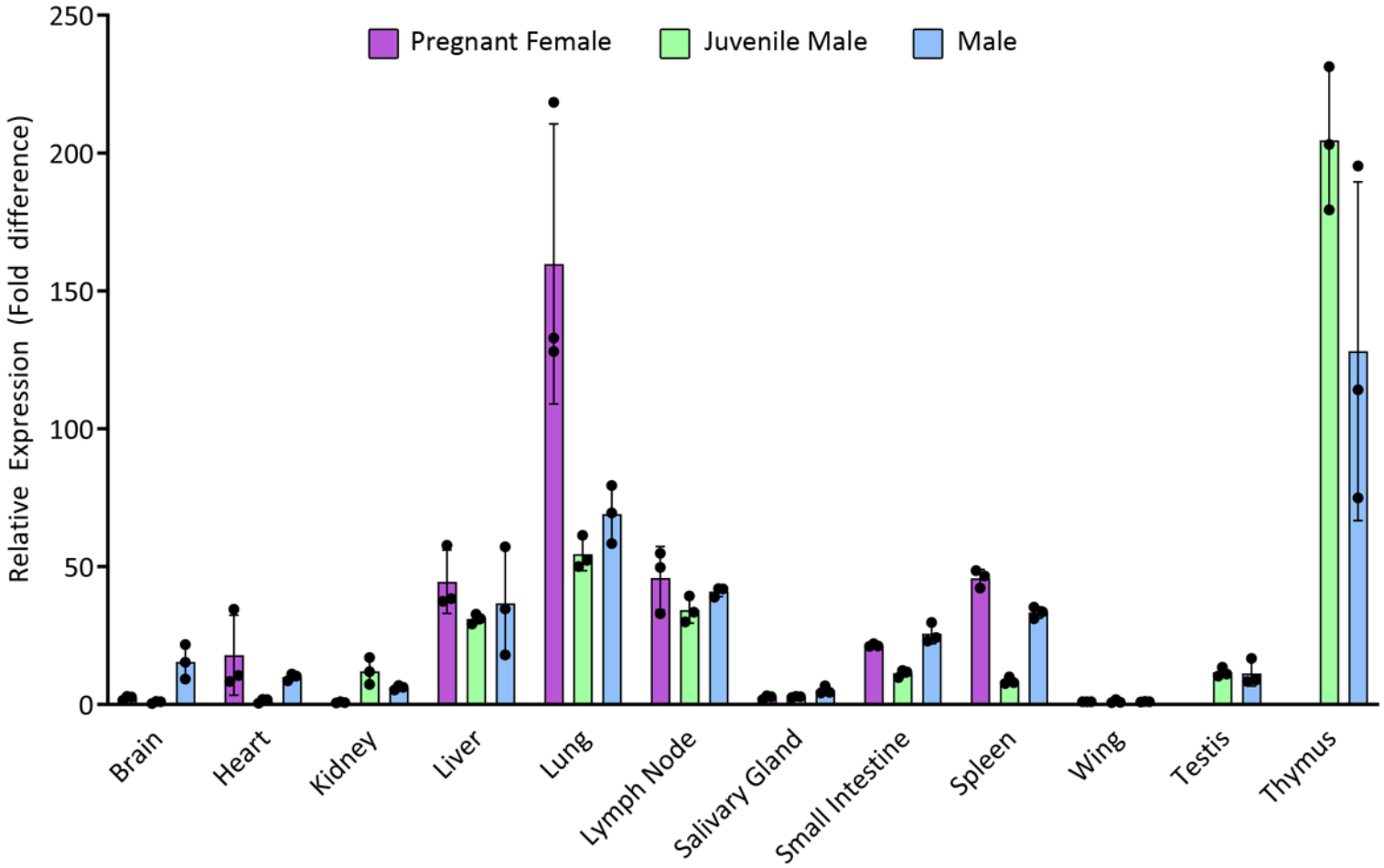

Figure 7. Relative tetherin expression across tissues from *P. alecto*. Expression of tetherin was assessed by qRT-PCR analysis for three bats, and calculated using the Livak/ΔΔCt method. In all cases, cycle threshold (Ct) values for tetherin were normalised against expression of the 18S rRNA gene. For all individuals, relative tetherin expression in each tissue is calibrated against expression in wing tissue. Error bars represent the standard error (N = 3).

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

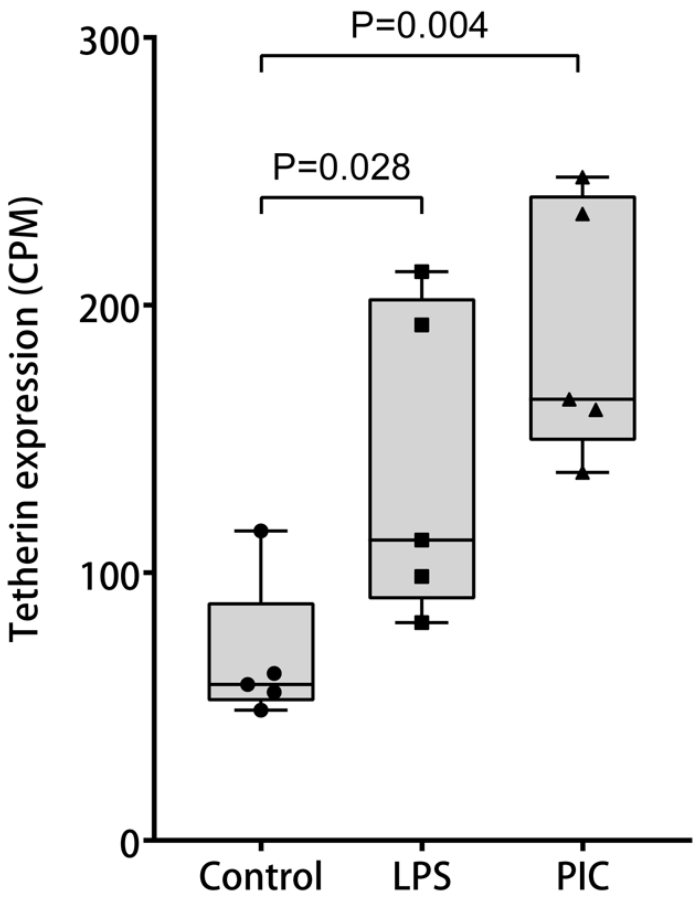

Figure 8. Stimulated expression of tetherin from *P. alecto* spleen tissue. Spleen tissue from 15 bats that were unstimulated (control), or stimulated with lipopolysaccharide (LPS) or polyinosinic:polycytidylic acid (PIC) (N = 5 for each treatment). Tetherin expression levels were from RNA-Seq data mapped against the *P. alecto* genomic scaffold, KB030270, which contains the tetherin gene. Expression levels are denoted as normalised counts-per-million (CPM) reads. P-values were determined using the one-tailed Mann-Whitney test.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

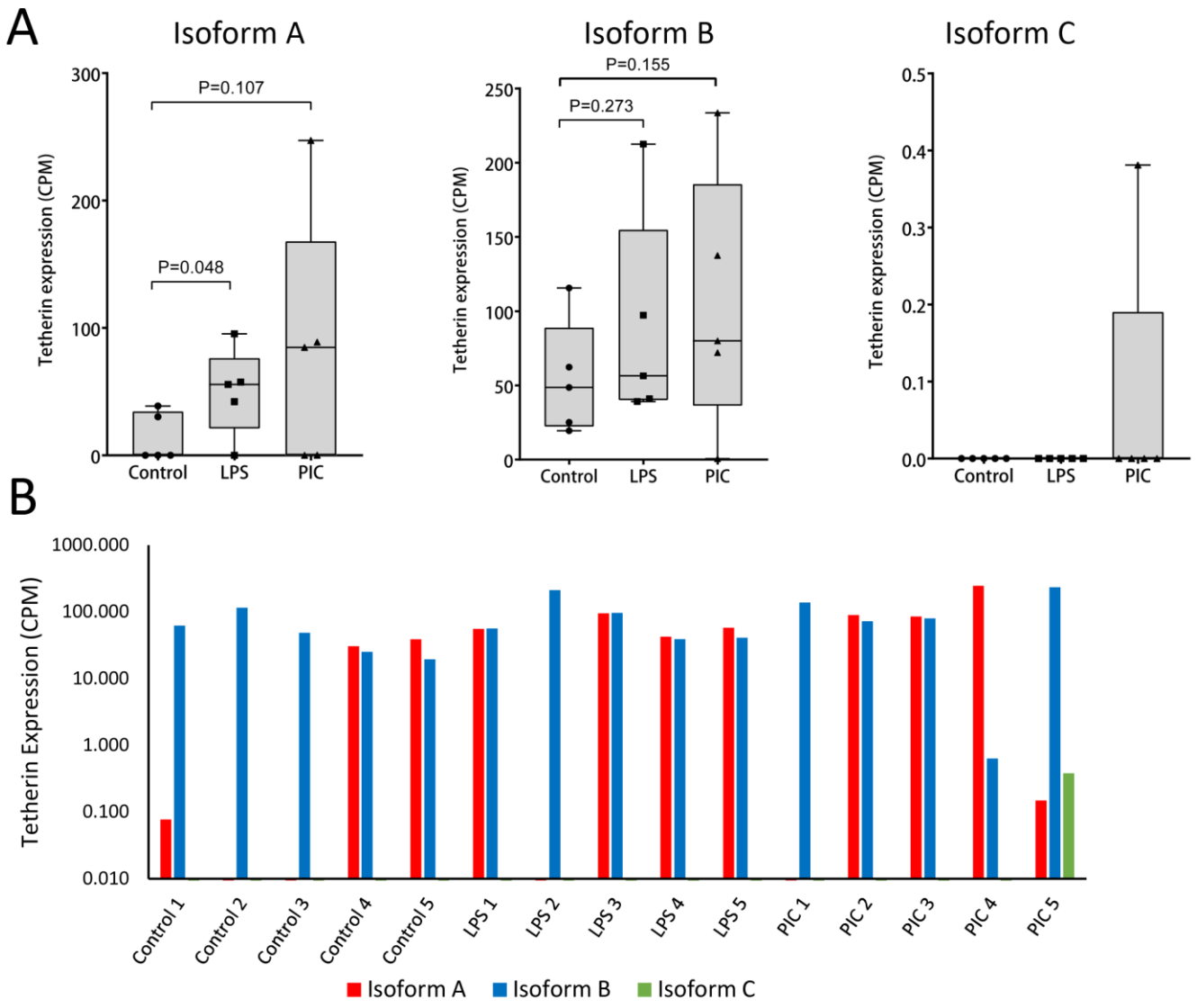

Figure 9. Expression of alternative isoforms of tetherin in stimulated and unstimulated *P. alecto* spleen tissue. Spleen tissue from 15 individuals were unstimulated, or stimulated with lipopolysaccharide (LPS) or polyinosinic:polycytidylic acid (PIC) (N = 5 for each treatment). Tetherin expression was determined from RNA-Seq sequence reads mapped against the *P. alecto* genomic scaffold, KB030270, which harbours the tetherin gene. Tetherin expression levels are represented as normalised counts-per-million (CPM) reads. **A.** Tetherin expression by isoform. P-values were determined using the one-tailed Mann-Whitney test. **B.** Expression of tetherin isoform in each individual sample.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

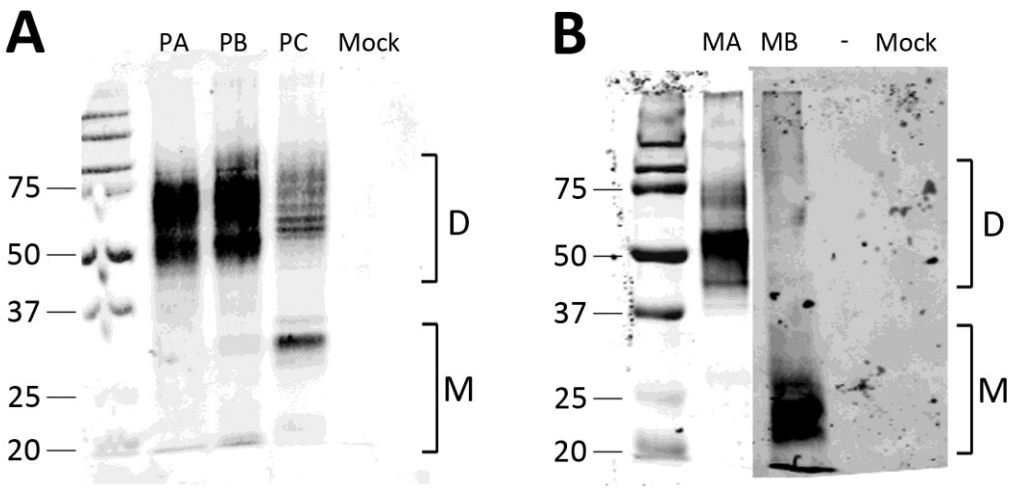

Figure 10. SDS-PAGE and Western blot analysis of the expression of modified bat tetherin constructs. A. Haemagglutinin (HA) tagged *P. alecto* tetherin isoforms A, B, and C (PA, PB, and PC, respectively). **B.** HA tagged *M. macropus* tetherin A and B (MA and MB, respectively). The left side of the membrane is displayed at low exposure while the right side is displayed at high exposure to enable clear images of the bands in each lane. Tetherin was expressed in mammalian HEK293T cells and cell lysates were collected 24 h after transfection of expression plasmids. Tetherin is present in dimeric (D) and monomeric (M) forms. Protein size scale is expressed in kDa.

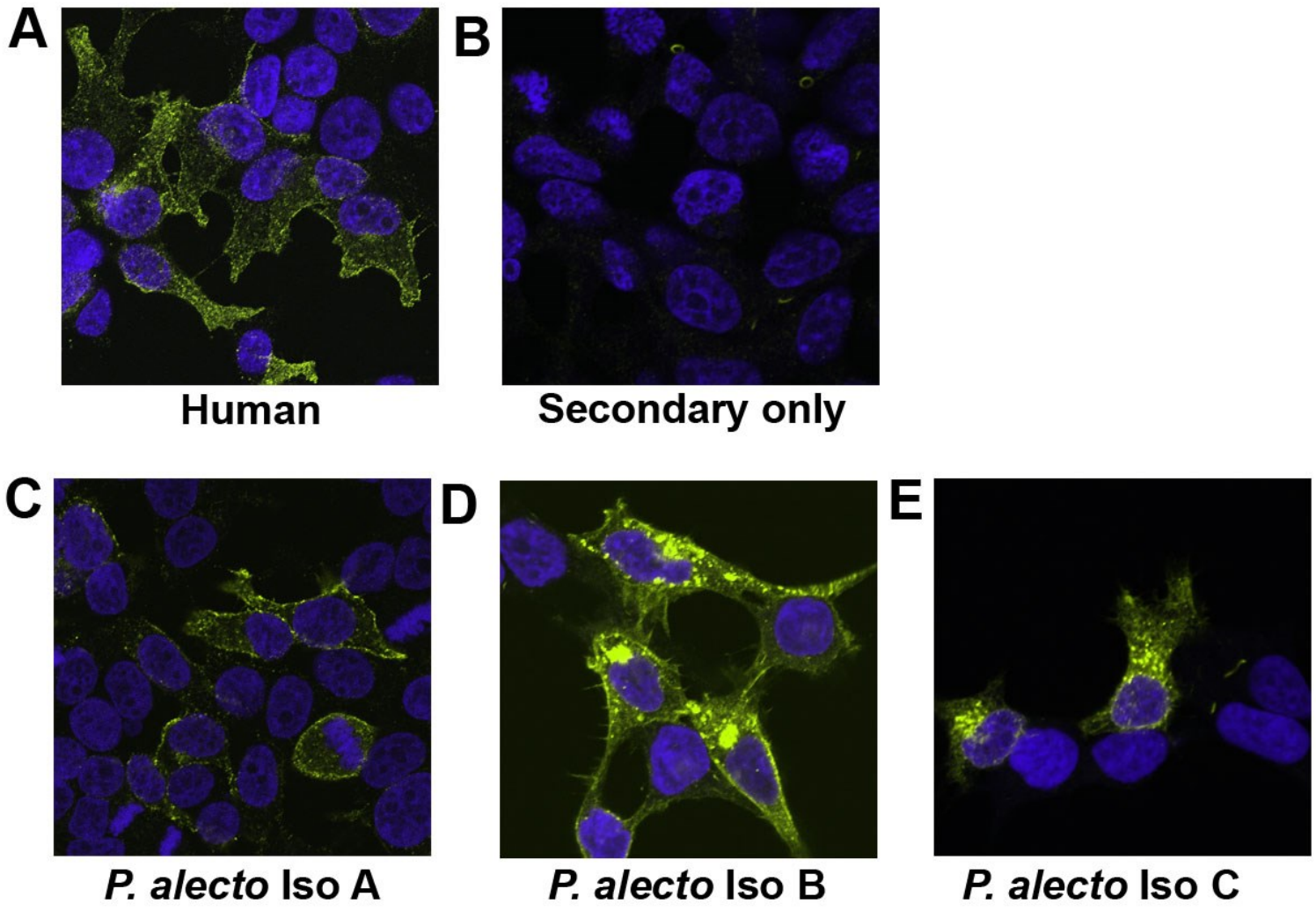

Figure 11. Expression of HA-tagged tetherin proteins. Human and *Pteropus alecto* tetherin isoforms A, B, and C in HEK293T cells. Tetherin localisation in cells was detected using anti-HA-tag rabbit monoclonal Ig and anti-rabbit AlexaFlour 488 secondary Ig. Nuclei are blue stained with Hoescht. Fixed and permeabilised cells were imaged on Nikon AR1 confocal microscope. **A.** Human l-tetherin, **B.** Negative control; HEK293T cells treated without the inclusion of the primary antibody, **C.** *P. alecto* tetherin isoform A, **D.** *P. alecto* tetherin isoform B, **E.** *P. alecto* tetherin isoform C. HA, haemagglutinin.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

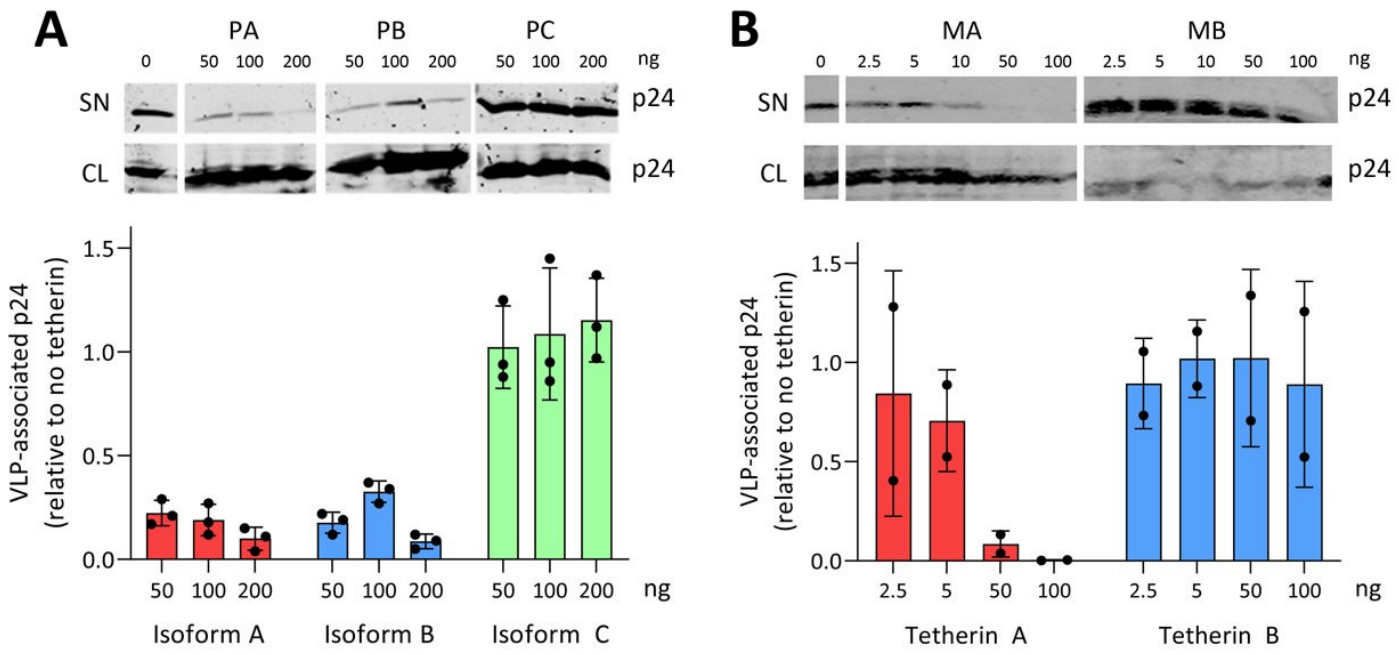

Figure 12. Bat tetherin restricts the release of HIV-1 virus-like particles (VLPs). A. *P. alecto* tetherin isoforms A (PA), B (PB), and C (PC), cotransfected with a HIVΔVpu construct. **B.** *M. macropus* tetherin A (MA) and tetherin B (MB) cotransfected with a HIVΔVpu construct. Mammalian HEK293T cells were cotransfected with 200 ng of the HIVΔVpu plasmid expression construct encoding the HIV Gag-Pol polyprotein, which generates HIV-1 VLPs that do not include the tetherin antagonist Vpu, and 0 – 200 ng of the tetherin plasmid expression vector. VLPs were harvested at 48 h and concentrated by ultracentrifugation using a sucrose cushion. VLP and cell lysates were subjected to SDS-PAGE and Western blot analysis. HIV-1 VLPs were detected with a mouse anti-p24 primary antibody and goat anti-mouse Alexa-Fluor 680 fluorophore-conjugated fluorescent secondary antibody. Representative Western blots are shown. The extent of VLP restriction was quantitated by densitometric analysis of Western blots comparing the relative ratios of VLPs present in the viral lysates and cell lysates from N=3 (*P. alecto*) or N=2 (*M. macropus*) independent assays. Error bars represent the standard deviation. HIV-1, human immunodeficiency virus type 1; SN, cell culture supernatant; CL, cell culture lysate.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

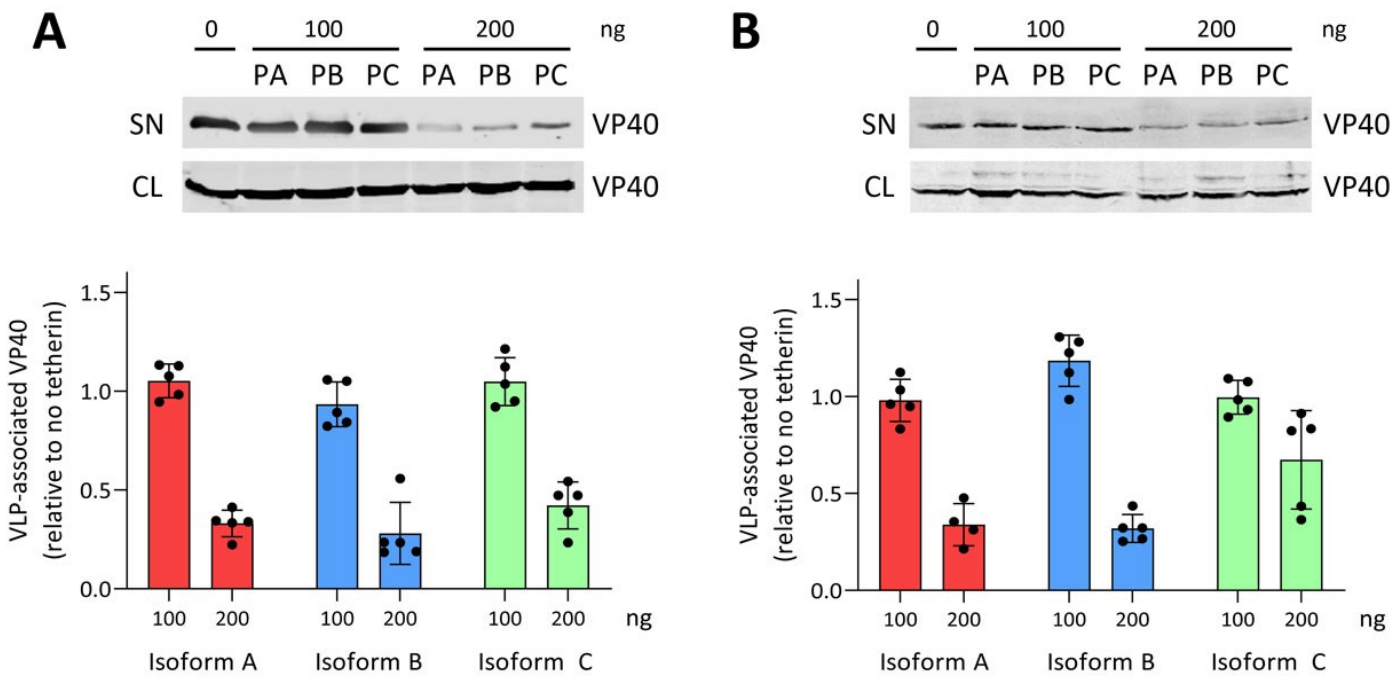

Figure 13. *Pteropus alecto* tetherin restricts the release of Ebola and Marburg virus-like particles (VLPs). *P. alecto* tetherin isoforms A (PA), B (PB), and C (PC), cotransfected with an **A.** Ebola virus construct or **B.** Marburg virus construct. Mammalian HEK293T cells were cotransfected with 200 ng of Ebola or Marburg virus plasmid expression construct encoding a VP40-eGFP protein which generates VLPs, and 0 – 200 ng of the tetherin plasmid expression vector. VLPs were harvested at 48 h and concentrated by ultracentrifugation through a sucrose gradient. VLP and cell lysates were subjected to SDS-PAGE and Western blot analysis. VLPs were detected with a mouse anti-GFP 4B10 primary antibody and a goat anti-mouse Alexa-Fluor 680 fluorophore-conjugated fluorescent secondary antibody. Representative Western blots are shown. The extent of VLP restriction was quantitated by densitometric analysis of Western blots comparing the relative ratios of VLPs present in the viral lysates and cell lysates from N = 5 independent assays. Error bars represent the standard deviation. The non-parametric Wilcoxon Rank Sum test was performed to calculate the statistical significance of the restriction of Ebola and Marburg VLPs in the 200 ng tetherin isoform C treatment groups; p = 0.008 for both. SN, cell culture supernatant; CL, cell culture lysate.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1. Sequence accessions with BioProject and cDNA sources for the prediction and/or confirmation of tetherin homologues

Sequence	Species	Common name	GenBank Accession	UniProt Identifier	NCBI Project Accession	cDNA source
Bats						
Tetherin	*Artibeus jamaicensis*	Jamaican fruit-eating bat	MT274379	-	PRJNA61227	-
Tetherin	*Carollia brevicauda*	Silky short-tailed bat	MT274380	-	PRJNA139591	-
Tetherin	*Carollia perspicillata*	Seba's short-tailed bat	MT274381	-	PRJNA291081	-
Tetherin	*Cynopterus sphinx*	Indian short-nosed fruit bat	MT274382	-	PRJNA222415	-
Tetherin	*Desmodus rotundus*	Common vampire bat	MT274383	-	PRJNA178123	-
Tetherin	*Eidolon helvum*	Straw-coloured fruit bat	MT274384	-	PRJNA209406	-
Tetherin	*Eonycteris spelaea*	Cave nectar bat	MT274385	-	PRJNA255191	-
Tetherin	*Hipposideros armiger*	Great roundleaf bat	MT274386	-	PRJNA260965	-
Tetherin	*Macrotus californicus*	California leaf-nosed bat	MT274387	-	PRJNA226078	-
Tetherin	*Miniopterus natalensis*	Natal long-fingered bat	MT274388	-	PRJNA270639	-
Tetherin	*Miniopterus schreibersii*	Common bent-wing bat	MT274389	-	PRJNA218524	
Tetherin	*Miniopterus schreibersii*	Common bent-wing bat	MT274405	-	-	Kidney cell line
Tetherin	*Murina leucogaster*	Hilgendorf's tube-nosed bat	MT274390	-	PRJNA182766	-
Tetherin	*Myotis brandtii*	Brandt's bat	MT274391	-	PRJNA218631	-
Tetherin	*Myotis davidii*	David's Myotis	MT274392	-	PRJNA172130	-
Tetherin	*Myotis laniger*	Chinese Water Myotis	MT274393	-	PRJNA255191	-
Tetherin	*Myotis lucifugus*	Little Brown Bat	MT274394	-	PRJNA246229	-
Tetherin A	*Myotis macropus*	Large-footed myotis	MT274406	-	-	Kidney cell line
Tetherin B	*Myotis macropus*	Large-footed myotis	MT274407	-	-	Kidney cell line
Tetherin C Isoform A	*Myotis macropus*	Large-footed myotis	MT274408	-	-	Kidney cell line
Tetherin C Isoform B	*Myotis macropus*	Large-footed myotis	MT274409	-	-	Kidney cell line
Tetherin D	*Myotis macropus*	Large-footed myotis	MT274410	-	-	Kidney cell line
Tetherin E	*Myotis macropus*	Large-footed myotis	MT274411	-	-	Kidney cell line
Tetherin	*Myotis myotis*	Greater mouse-eared bat	MT274395	-	PRJNA267654	-
Tetherin A	*Myotis ricketti*	Rickett's big-footed bat	MT274396	-	PRJNA198831	-
Tetherin A	*Myotis ricketti*	Rickett's big-footed bat	MT274412	-	-	Spleen tissue
Tetherin B	*Myotis ricketti*	Rickett's big-footed bat	MT274413	-	-	Spleen tissue
Tetherin Isoform A	*Pteropus alecto*	Black flying fox	MT274397	-	PRJNA73831	-
Tetherin Isoform A	*Pteropus alecto*	Black flying fox	MT274414	-	-	Spleen tissue
Tetherin Isoform B	*Pteropus alecto*	Black flying fox	MT274415	-	-	Spleen tissue
Tetherin Isoform C	*Pteropus alecto*	Black flying fox	MT274416	-	-	Spleen tissue
Tetherin	*Pteropus vampyrus*	Large flying fox	MT274398	-	PRJNA20325	-
Tetherin	*Rhinolophus ferrumequinum*	Greater horseshoe bat	MT274399	-	PRJNA231230	-
Tetherin	*Rhinolophus macrotis*	Big-eared horseshoe bat	MT274400	-	PRJNA261657	-
Tetherin	*Rousettus aegyptiacus*	Egyptian fruit bat	MT274401	-	PRJNA300284	-
Tetherin	*Tadarida brasiliensis*	Mexican free-tailed bat	MT274402	-	PRJNA184055	-
Tetherin	*Taphozous melanopogon*	Black-bearded tomb bat	MT274403	-	PRJNA255191	-
Tetherin	*Uroderma bilobatum*	Tent-making bat	MT274404	-	PRJNA268573	-
Other mammals						
Tetherin A Isoform A	*Bos taurus*	Cow	-	J7M5G6	-	-
Tetherin A Isoform B	*Bos taurus*	Cow	-	J7MAQ7	-	-
Tetherin B	*Bos taurus*	Cow	-	J7M2B2	-	-
Tetherin	*Felis catus*	Cat	-	F8R0X8	-	-
L-tetherin	*Homo sapiens*	Human	-	Q10589-1	-	-
S-tetherin	*Homo sapiens*	Human	-	Q10589-2	-	-
Tetherin	*Loxodonta africana*	Elephant	-	G3UKK9	-	-
Tetherin	*Macaca mulatta*	Rhesus Macaque	-	C4P4A2	-	-
Tetherin	*Mus musculus*	Mouse	-	Q8R2Q8	-	-
Tetherin A	*Ovis aries*	Sheep	-	D5JZS7	-	-
Tetherin B	*Ovis aries*	Sheep	-	D5JZS8	-	-
Tetherin	*Pan troglodytes*	Chimpanzee	-	D7RVC2	-	-

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Evolutionary hotspots in bat tetherin

Tetherin amino acid[A]	dN/dS value
D24	2.668
R25	2.626
G29	2.668
W30	2.660
G32	2.680
S43	2.680
A45	2.672
V46	2.636
R63	2.679
I71	2.671
A105	2.641
S151	2.664
A171	2.678
L179	2.651

[A] The consensus amino acid identity is indicated at the position number relative to the *Cynopterus sphinx* tetherin protein sequence.

bioRxiv preprint doi: https://doi.org/10.1101/2020.04.08.031203; this version posted November 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 3. cDNA amplification and vector sequencing primers for fruit bat and vesper bat tetherins.

A

		cDNA Amplification Primers	Primer sequence (5' > 3')	Direction	Target region	Successful combinations	PCR T_A°C
Pteropus alecto	**Tetherin**	PaF1_JH_119	TCACTGCAAGGGGTTCTCTC	Forward	5' UTR	119/121-2	60
		PaF2_JH_120	GGAAACTTCACTGCAAGGGG	Forward	5' UTR	120/121-2	60
		PaR1_JH_121	CTTCTCCCAGCTTTGTTGCC	Reverse	3' UTR		
		PaR2_JH_122	CTCCTCTCCCCCAAAATGTC	Reverse	3' UTR		
Miniopterus schreibersii	**Tetherin**	MiniF1_JH_466	CCCACAAACTCCCTACACCC	Forward	5' UTR	466/468	60
		MiniF2_JH_467	ATGCTAATGAAGGGGCGGGG	Forward	5' UTR	466/469	62
		MiniR1_JH_468	CTGTCTGTCTTCCTGGGAC	Reverse	3' UTR	467/469	62
		MiniR2_JH_469	GGACAGGTCAGGGAAACCAA	Reverse	3' UTR		
Myotis macropus & Myotis ricketti	**Tetherin A, C, D**	MyoF1_JH_479	TCCACTGCATCCCTCTG	Forward	5' UTR	479/481-4	60
		MyoF2_JH_480	ATGGCACCCACTTTTTAC	Forward	5' UTR	480/481-4	60
		MyoR1_JH_481	TCAGCCAGGTTAGAATGTG	Reverse	3' UTR		
		MyoR2_JH_482	TCCTTGGGCAAACAGCTCTC	Reverse	3' UTR		
		MyoR3_JH_483	CAGGAAACTCTCAGAAAAG	Reverse	3' UTR		
		MyoR4_JH_484	CATCTTTCCAAGACCACA	Reverse	3' UTR		
	Tetherin	MyoF3_JH_474	GCTCCTGTGCATCCCTCTGG	Forward	5' UTR	474/478	60
		MyoR5_JH_478	CCTGGTTAGAATGTGCTTT	Reverse	3' UTR		

B

Sequencing Primers	Primer sequence (5' > 3')	Direction	Target region
M13F	GTAAAACGACGGCCAG	Forward	pCR-Blunt-II-TOPO vector
M13R	CAGGAAACAGCTATGAC	Reverse	pCR-Blunt-II-TOPO vector
T7F	TAATACGACTCACTATAGGG	Forward	pcDNA3.1 vector
BGHR	TAGAAGGCACAGTCGAGG	Reverse	pcDNA3.1 vector

C

	qPCR Primers	Primer sequence (5' > 3')	Direction
Tetherin	Teth_F	TGACTGTGGCCGTGATCGT	Forward
	Teth_R	CCATTTTTGCAGGCCTCACT	Reverse
18S rRNA	18S_rRNA_F	CGGCTACCACATCCAAGGAA	Forward
	18S_rRNA_R	GCTGGAATTACCGCGGCT	Reverse
	qPCR Probes	**Probe sequence (5' > 3')**	
Tetherin	Teth_Probe	*FAM* _TCGCCGTCGAGAACA_*MGB*	
18S rRNA	18S_Probe	*VIC* _TGCTGGCACCAGACTTGCCCTC_*TAMRA*	

A Primers used to amplify tetherin homologues from cDNA generated from *Myotis macropus* , *M. ricketti* , and *Miniopterus schreibersii* . Combinations of forward and reverse primers identified as capable of amplifying tetherin are indicated alongside the highest annealing temperature at which amplification was successful. **B** The sequencing primers used to confirm the sequence of tetherin clones and modified tetherin constructs inserted into plasmid vectors are listed. **C** The qPCR primers and probes used to analyse tetherin expression in bat tissues are listed. Probe 5' reporters and 3' quenchers are indicated with italics. UTR, untranslated region; PCR, polymerase chain reaction; TA°C, primer annealing temperature.

Table 4. Primers used for the generation of haemagglutinin-tagged *Pteropus alecto* and *Myotis macropus* tetherin expression constructs through a 2-Step PCR process.

Pteropus alecto	Construct name	2 Step PCR usage	Primer name	Primer sequence (5' > 3')	Direction	Encoded RE site	Encoded Epitope
Tetherin Isomer A	**pD-PaTAH1**	Step 1, 5' half	TethXF_JH_353	CTCTCTCGAGAGCTTCTTCTCCTGACTCC	Forward	*Xho* I	-
			TethAB-HA1_JH_355	CGTATGGGTACCCGCTGGCATGCTCTTTCCTTAGCTG	Reverse	-	HA
		Step 1, 3' half	TethAB-HA2_JH_356	CAGCGGGTACCCATACGATGTTCCAGATTACGCTGGCAGCTCTGGCGAGAAAAATGG	Forward	-	HA
			TethXR_JH_354	TTTTTCTAGAATGTTTCTCCACCCCTAAGGC	Reverse	*Xba* I	-
		Step 2	TethXF_JH_353	CTCTCTCGAGAGCTTCTTCTCCTGACTCC	Forward	*Xho* I	-
			TethXR_JH_354	TTTTTCTAGAATGTTTCTCCACCCCTAAGGC	Reverse	*Xba* I	-
Tetherin Isomer B	**pD-PaTBH1**	Step 1, 5' half	TethXF_JH_353	CTCTCTCGAGAGCTTCTTCTCCTGACTCC	Forward	*Xho* I	-
			TethAB-HA1_JH_355	CGTATGGGTACCCGCTGGCATGCTCTTTCCTTAGCTG	Reverse	-	HA
		Step 1, 3' half	TethAB-HA2_JH_356	CAGCGGGTACCCATACGATGTTCCAGATTACGCTGGCAGCTCTGGCGAGAAAAATGG	Forward	-	HA
			TethXR_JH_354	TTTTTCTAGAATGTTTCTCCACCCCTAAGGC	Reverse	*Xba* I	-
		Step 2	TethXF_JH_353	CTCTCTCGAGAGCTTCTTCTCCTGACTCC	Forward	*Xho* I	-
			TethXR_JH_354	TTTTTCTAGAATGTTTCTCCACCCCTAAGGC	Reverse	*Xba* I	-
Tetherin Isomer C	**pD-PaTCH1**	Step 1, 5' half	TethXF_JH_353	CTCTCTCGAGAGCTTCTTCTCCTGACTCC	Forward	*Xho* I	-
			TethC-HA1_JH_359	TATGGGTACCCGCTCCTTAGCTGTTCCGGCTCCG	Reverse	-	HA
		Step 1, 3' half	TethC-HA2_JH_360	AGCGGGTACCCATACGATGTTCCAGATTACGCTGGCAGCCCACAACCTGGACTGGTCC	Forward	-	HA
			TethXR_JH_354	TTTTTCTAGAATGTTTCTCCACCCCTAAGGC	Reverse	*Xba* I	-
		Step 2	TethXF_JH_353	CTCTCTCGAGAGCTTCTTCTCCTGACTCC	Forward	*Xho* I	-
			TethXR_JH_354	TTTTTCTAGAATGTTTCTCCACCCCTAAGGC	Reverse	*Xba* I	-

Myotis macropus	Construct name	2 Step PCR usage	Primer name	Primer sequence (5' > 3')	Direction	Encoded RE site	Encoded Epitope
Tetherin A	**pD-MmTAH1**	Step 1, 5' half	MmacXF1_JH_510	CTGCAGAATTCGCCCTTATG	Forward	*Eco* RI	-
			MmacAH1_JH_512	GGGACGTCGTATGGGTATGGGTATGGCCGGCCGACCAAGGCCTCATTCTC	Reverse	-	HA
		Step 1, 3' half	MmacAH2_JH_513	TACCCATACGACGTCCCAGACTACGCTGCTAGCTCTGCCAAGGGTCCCC	Forward	-	HA
			MmacXR1_JH_511	CAGTGTGCTGGAATTCGCC	Reverse	*Eco* RI	-
		Step 2	MmacXF1_JH_510	CTGCAGAATTCGCCCTTATG	Forward	*Eco* RI	-
			MmacXR1_JH_511	CAGTGTGCTGGAATTCGCC	Reverse	*Eco* RI	-
Tetherin B	**pD-MmTBH1**	Step 1, 5' half	MmacXF2_JH_516	CTGCAGAATTCGCCCTTTCC	Forward	*Eco* RI	-
			MmacBH1_JH_518	CTGGGACGTCGTATGGGTATGGGTATGGCCGGCCGTTGGGGTCCTTGCCGAACAC	Reverse	-	HA
		Step 1, 3' half	MmacBH2_JH_519	CCCATACGACGTCCCAGACTACGCTGCTAGCAATGGCAAGGGCTTCCCTAAC	Forward	-	HA
			MmacXR2_JH_517	CAGTGTGCTGGAATTCGCC	Reverse	*Eco* RI	-
		Step 2	MmacXF2_JH_516	CTGCAGAATTCGCCCTTTCC	Forward	*Eco* RI	-
			MmacXR2_JH_517	CAGTGTGCTGGAATTCGCC	Reverse	*Eco* RI	-

HA, haemaggultinin; PCR, polymerase chain reaction