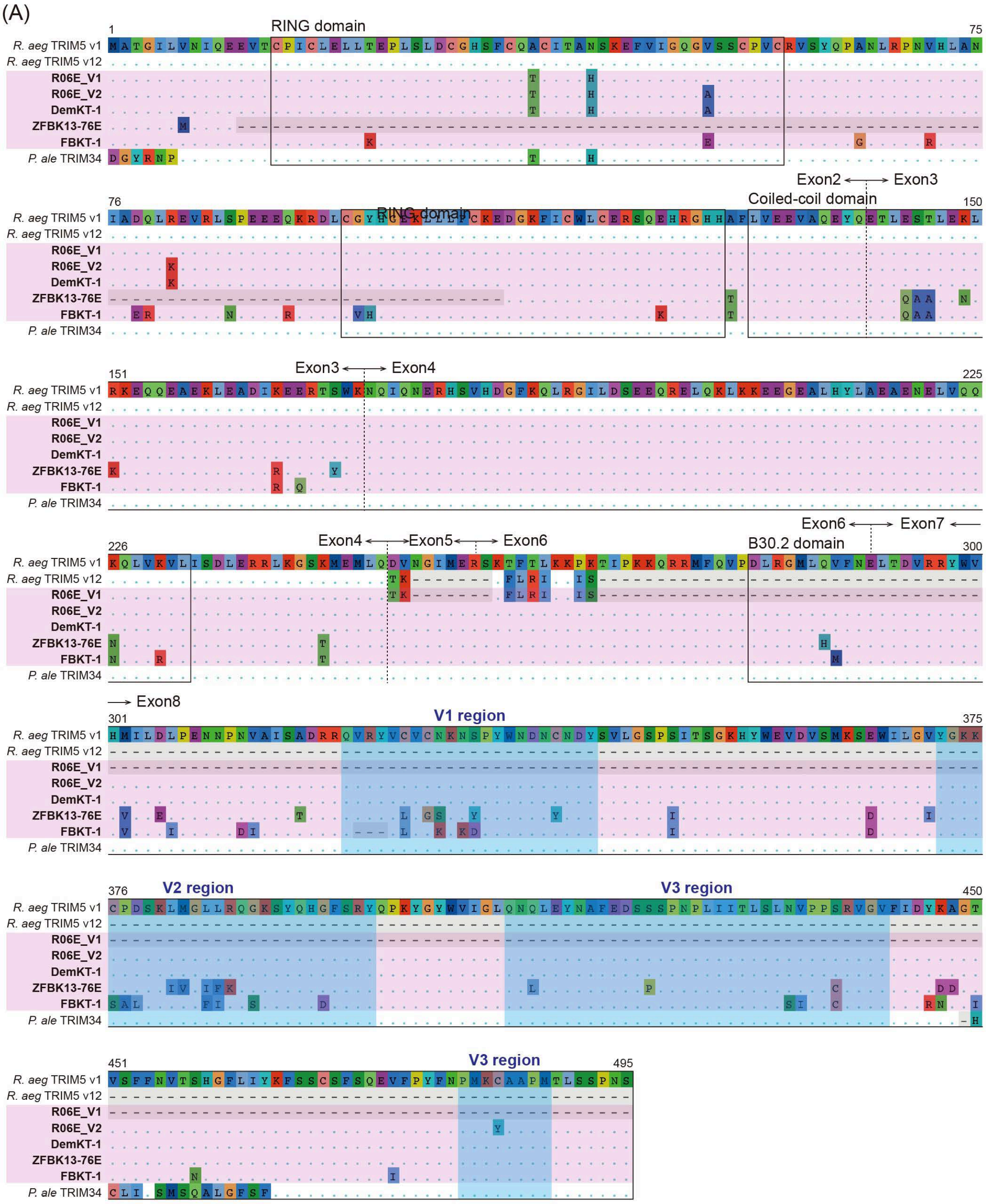
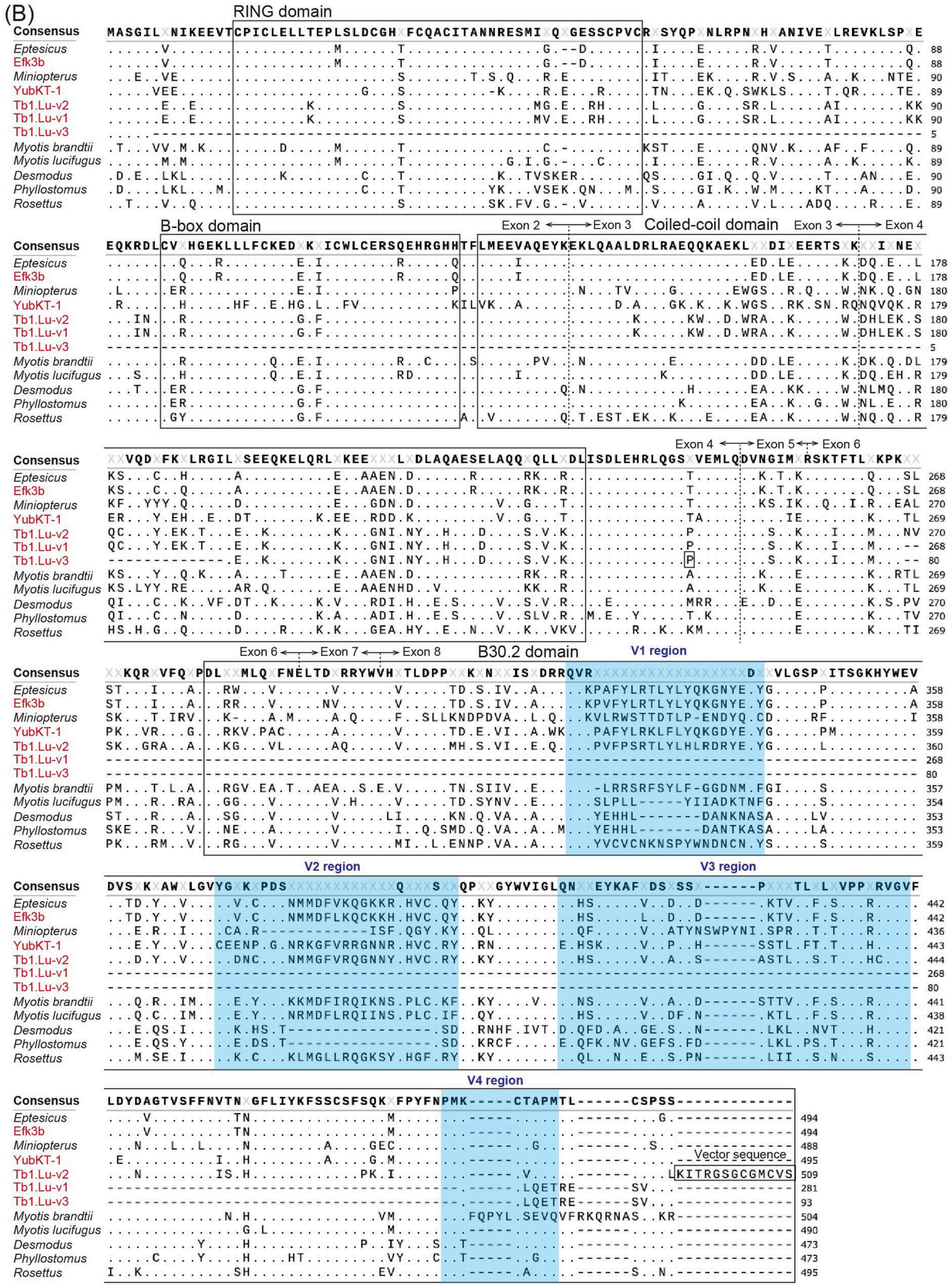
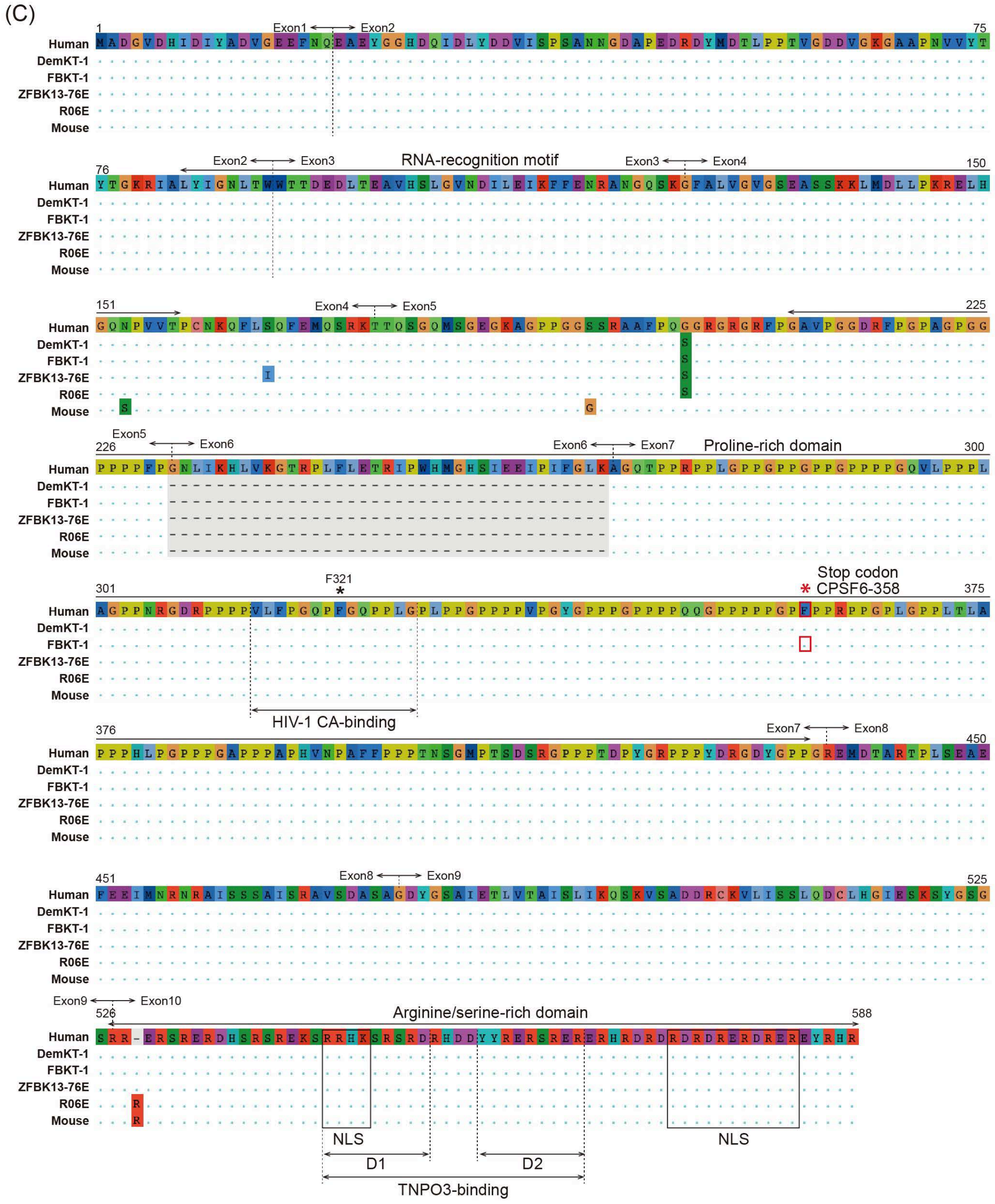
Supplemental information captions

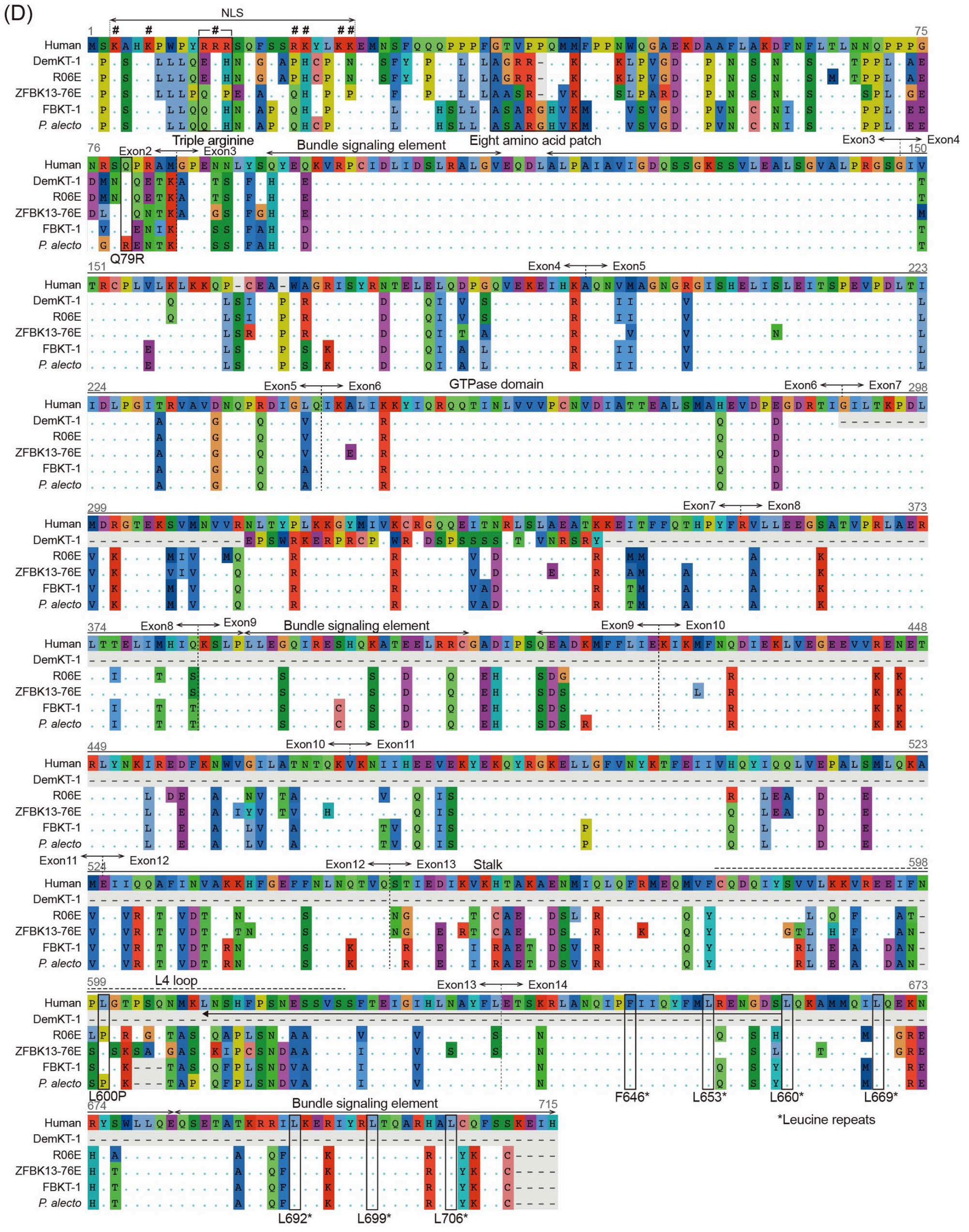



(B)
RING domain
B-box domain
Coiled-coil domain
B30.2 domain
Exon 2
Exon 3
Exon 4
Exon 5
Exon 6
Exon 7
Exon 8
V1 region
V2 region
V3 region
V4 region
Vector sequence
Consensus
Eptesicus
Efk3b
Miniopterus
YubKT-1
Tb1.Lu-v2
Tb1.Lu-v1
Tb1.Lu-v3
Myotis brandtii
Myotis lucifugus
Desmodus
Phyllostomus
Rosettus

3


(C)
Human
DemKT-1
FBKT-1
ZFBK13-76E
R06E
Mouse
1
75
Exon1
Exon2
76
150
Exon2
Exon3
RNA-recognition motif
Exon3
Exon4
151
225
Exon4
Exon5
226
300
Exon5
Exon6
Exon6
Exon7
Proline-rich domain
301
375
F321
Stop codon
CPSF6-358
HIV-1 CA-binding
376
450
Exon7
Exon8
451
525
Exon8
Exon9
526
588
Exon9
Exon10
Arginine/serine-rich domain
NLS
NLS
D1
D2
TNPO3-binding

4


(D)
NLS
Human
DemKT-1
R06E
ZFBK13-76E
FBKT-1
P. alecto
Triple arginine
Bundle signaling element
Eight amino acid patch
Exon2
Exon3
Exon4
Exon5
Exon6
Exon7
Exon8
Exon9
Exon10
Exon11
Exon12
Exon13
Exon14
Q79R
GTPase domain
Stalk
L4 loop
L600P
F646*
L653*
L660*
L669*
L692*
L699*
L706*
*Leucine repeats

(E)

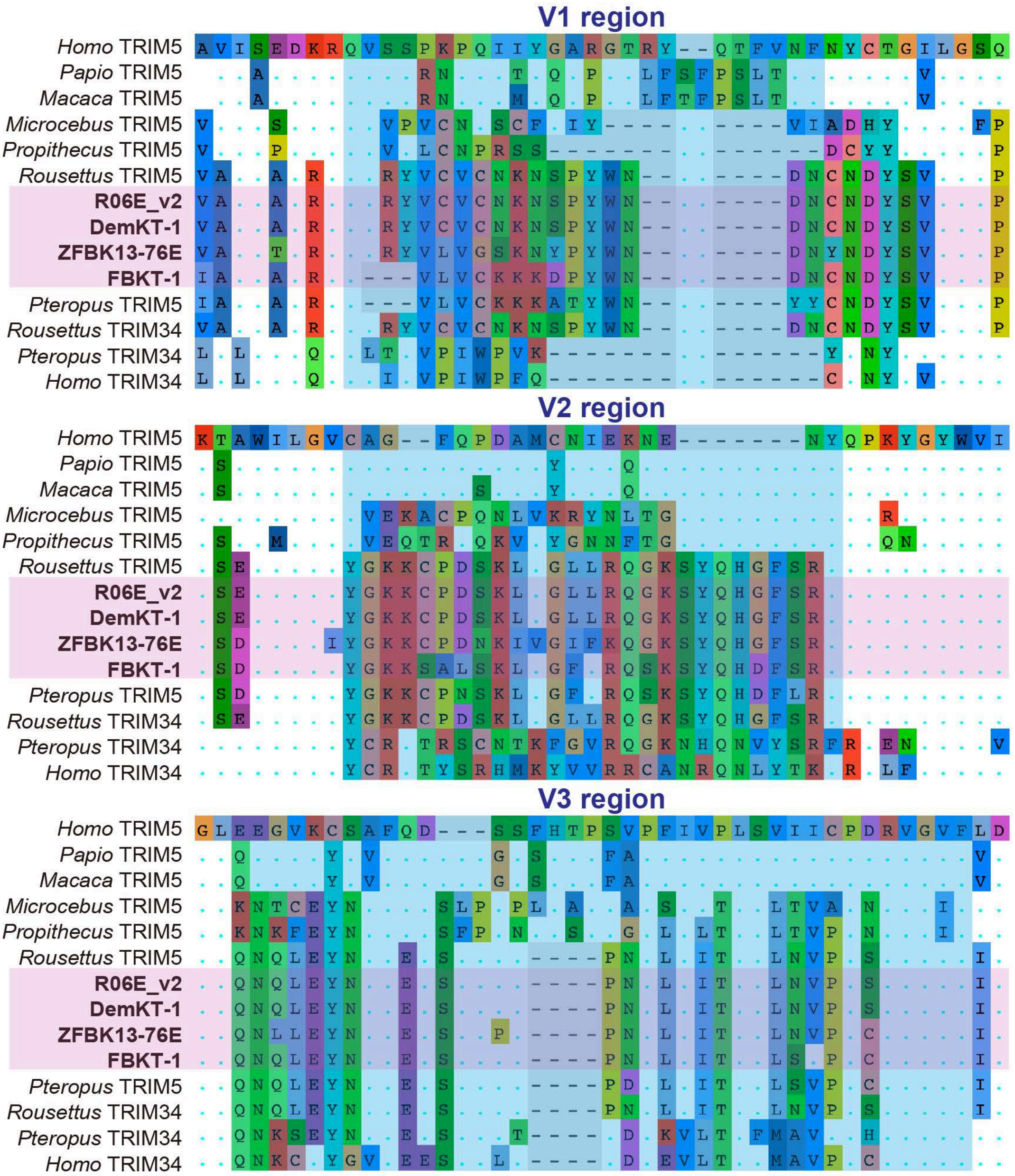


Fig. S1 Amino acid alignments of pteropid megabat TRIM5α, CPSF6, and Mx2/MxB, related to Figs. 1, 5, 7, and 8.

(A) Amino acid alignments of pteropid megabat TRIM5α, related to Fig. 5A. Amino acid sequences of newly identified megabat cDNAs were predicted computationally based on their nucleotide sequences and aligned with their homologs of other pteropid megabats collected from the NCBI GenBank database. Each dot represents an amino acid identical to the top sequence, and each hyphen exhibits an amino acid deletion. The position numbering is based on the positions of the top sequence. Exon boundaries predicted from the NCBI database are shown with dotted lines. Newly identified sequences are highlighted with a magenta background. The tripartite motif and the B30.2 domain are highlighted in boxes, and the variable regions in the B30.2 domain are highlighted with a cyan background. Note that the new TRIM5 cDNAs from four megabat cell lines contain the three characteristic domains, RING, B-box2, and Coiled-coil domains, confirming that they are members of the TRIM family. The TRIM5α cDNA cloned from ZFBK13-76E is an exception, lacking the entire RING domain and the amino-terminal half of the B-box2 domain, whose structure is similar to the predicted *TRIM34*-like gene of *P. vampyrus* (LOC102888261, Accession # XM_015592331, Reference ASM32557v1). From R06E cells, we successfully cloned two variants, R06E_v1 and R06E_v2. The R06E_v1 isoform was closely related to the predicted transcript variant 12 of *Rousettus TRIM5*-like gene (Accession # XM_016150076). While the nucleotide length of the R06E_v1 was longer than that of the R06E_v2 (1,578 bp vs. 1,488 bp), the insertion of a premature stop codon makes the translation product of the R06E_v1 variant lose the carboxy-terminal half of the protein including the entire B30.2 domain. (B) Amino acid alignments of microbat TRIM5α, related to Figs. 5A and D. Names of the newly identified genes are highlighted in red. Tb1.Lu-v1, -v2, and -v3 indicates Tb1.Lu TRIM5α isoforms 1, isoform 2, and isoform 3, respectively. Functional domains and the variable regions in the B30.2 domain are highlighted with black boxes and a cyan background, respectively. The consensus with a threshold of >50% homology is shown at the top of the alignment; otherwise, shown as "X." Predicted exon boundaries are shown based on the boundaries indicated for *Rousettus aegyptiacus* TRIM5α (Fig. S1). At the carboxy-terminal end, Tb1.Lu-v2 has an additional 13 amino acid extension derived from the delivery MLV vector, likely a result of a PCR or cloning error. This insertion of the vector sequence was most likely caused by a single cytosine nucleotide deletion in the third codon of the second-to-last serine (AG*<u>C</u>*). This deletion would result in the frameshift of the last serine, changing the stop codon to a leucine with the addition of the vector-derived sequence. (C) Amino acid alignments of pteropid megabat CPSF6, related to Figs. 5B and 7. Functional domains and predicted nuclear localization signals (NLS) are as per the publication by Lee et al. (1). A black box with an asterisk indicates the stop codon experimentally introduced in the Proline-rich domain for truncation. Note that the new megabat CPSF6 include three characteristic domains; the RNA-recognition motif, the Proline-rich (P-rich) domain, and the Arginine/serine-rich (RS) domain with two NLS, confirming that they are members of the CPSF6. The four megabat CPSF6 contain 37 amino acids-deletions in the Proline-rich domain, similar to the human CPSF6 isoform 1 (2). Therefore, the megabat CPSF6 showed high similarities to both human and mouse CPSF6 at an amino acid sequence level with a few lineage-specific variabilities. (D) Amino acid alignments of pteropid megabat Mx2/MxB, related to Figs. 5C and 8. A black box with an asterisk indicates a position of an amino acid substitution (Q79R and L600P) assessed in this study (Fig. 8). The arginine/lysin residues in the amino-terminal NLS, phenylalanine/leucine residues in the carboxy-terminal leucine repeats, and the triple-arginine motif/the 8-amino acid segment in the amino-terminal region are highlighted with #, asterisks with position numbers, and black boxes, respectively. Note that the newly identified pteropid bat Mx2 cDNAs contain three Bundle signaling elements (BSEs), the GTPase domain, and the Stalk domain, suggesting that they are close relatives to antiviral Mx proteins. The DemKT-1 Mx2 contains a premature stop codon in the GTPase domain. (E) Amino acid alignment of three variable regions in the B30.2 domain of primate and pteropid megabat TRIM5α and TRIM34 collected from the NCBI GenBank database.

Table S1. GenBank accession numbers of the samples utilized for constructing phylogenetic trees, calculating *dN/dS* ratios, and sequence analysis of megabat CPSF6 isoforms, related to Figs. 1A, 5, 7E, and 8D.

Mitochondrial cytochrome b (Fig 1A)	
Species	Accession number
Pteropus vampyrus	FJ561401
Pteropus alecto	KJ532406
Rousettus aegyptiacus	EU624124
Eonycteris spelaea	AB062476
Hipposideros armiger	EU434946
Cloeotis percivali	FJ457616
Rhinolophus sinicus	EU434941
Rhinolophus ferrumequinum	EU436673
Megaderma lyra	DQ888678
Rhinopoma microphyllum	KF874547
Nycteris leporinus	AF330802
Emballonura alecto	HQ693711
Phyllostomus hastatus	FJ155479
Desmodus rotundus	FJ155477
Mormoops megalophylla	AF330808
Pteronotus parnellii	AY604456

Noctilio albiventris	AF330806
Furipterus horrens	AY621004
Thyroptera tricolor	AY621005
Mystacina tuberculata	AY960981
Myotis (Cistugo) lesueuri	AY485687
Myotis lucifugus	AF376854
Myotis brandtii	AM261886
Myotis davidii	AB106591
Eptesicus fuscus	AF376835
Tadarida brasiliensis	JF489129
Natalus major	AY621021
Newly cloned bat restriction factor homologs (Fig 5)	
Sample name	Accession number
DemKT-1 TRIM5alpha	OP272517
EfK3B TRIM5alpha	OP272518
FBKT-1 TRIM5alpha	OP272519
R06E TRIM5alpha-v1	OP272520
R06E TRIM5alpha-v2	OP272521
Tb1Lu TRIM5alpha-v1	OP272522
Tb1Lu TRIM5alpha-v2	OP272523

Tb1Lu TRIM5alpha-v3	OP272524
YubKT-1 TRIM5alpha	OP272525
ZFBK13-76E TRIM5alpha	OP272526
DemKT-1 CPSF6	OP272527
FBKT-1 CPSF6	OP272528
R06E CPSF6	OP272529
ZFBK13-76E CPSF6	OP272530
DemKT-1 Mx2/MxB	OP272531
FBKT-1 Mx2/MxB	OP272532
R06E Mx2/MxB	OP272533
ZFBK13-76E Mx2/MxB	OP272534
Mammalian TRIM, Mx2 and CPSF6 (Fig 5)	
Sample name	Accession
Rousettus aegyptiacus TRIM5	XM_016150065
Homo TRIM5	NM_033034
Macaca TRIM5	NM_001032910
Bos TRIM5	DQ381151
Sus TRIM5	AK235927
Papio TRIM5	NM_001112632
Nomascus TRIM5	NM_001280113

Microcebus TRIM5	HQ413345
Ovis TRIM5	NM_001318072
Capra TRIM5	XM_005689800
Eptesicus fuscus TRIM5	XM_028137626
Callithrix TRIM5	XM_009007572
Saimiri TRIM5	XM_010334225
Rhinopithecus TRIM5	XM_010364548
Propithecus TRIM5	XM_012654429
Homo TRIM34	NM_021616
Rousettus aegyptiacus TRIM34	XM_016126824
Phyllostomus discolor TRIM34	XM_028515501
Myotis lucifugus TRIM34	XM_023758764
Eptesicus fuscus TRIM34	XM_008160439
Pteropus alecto TRIM34	XM_015592357
Pteropus vampyrus TRIM34	XM_011383134
Miniopterus natalensis TRIM34	XM_016215034
Homo TRIM7	NM_203293
Bos TRIM7	NM_001205446
Papio TRIM7	XM_003900656
Macaca TRIM7	XM_005558834
Homo TRIM26	NM_003449

Pan TRIM26	XM_003829748
Macaca TRIM26	NM_001114967
Bos TRIM26	XM_002697339
Homo TRIM39	BT007370
Gorilla TRIM39	XM_004043563
Homo TRIM47	NM_033452
Papio TRIM47	XM_003913451
Bos TRIM47	NM_001193197
Homo TRIM65	NM_173547
Macaca TRIM65	XM_028836809
Bos TRIM65	XM_027518688
Homo TRIM21	NM_003141
Macaca TRIM21	NM_001258045
Bos TRIM21	NM_182655
Homo TRIM22	NM_006074
Macaca TRIM22	EU124697
Canis TRIM22	XM_542402
Homo TRIM6	NM_001003818
Sus TRIM6	XM_021062360
Canis TRIM6	NM_001205190
Myotis davidii Mx2*	XM_006754321.2

Eptesicus fuscus Mx2*	XM_028151882.1
Miniopterus natalensis Mx2*	XM_016221319.1
Phyllostomus discolor Mx2*	XM_036017448.1
Desmodus rotundus Mx2*	XM_024559822.1
Rhinolophus ferrumequinum Mx2*	XM_033089263.1
Pteropus alecto Mx2*	XM_006916730.3
Bos taurus Mx2	XM_015473641.2
Ovis aries Mx2	XM_012175764.3
Macaca mulatta Mx2	NM_001079696.1
Cercocebus atys Mx2	XM_012037318.1
Papio anubis Mx2	XM_003895463.5
Homo sapiens Mx2	NM_002463.2
Pan troglodytes Mx2	XM_016938556.2
Gorilla gorilla Mx2	XM_019017898.2
Pongo abelii Mx2	XM_002830701.2
Saimiri boliviensis Mx2	XM_010338253.1
Callithrix jacchus Mx2	XM_035283249.1
Aotus nancymaae Mx2	XM_021669412.1
Miniopterus natalensis CPSF6	XM_016199385.1
Eptesicus fuscus CPSF6	XM_028156487.1
Rhinolophus ferrumequinum CPSF6	XM_033116537.1

Molossus molossus CPSF6	XM_036250921.1
Hipposideros armiger CPSF6	XM_019640530.1
Desmodus rotundus CPSF6	XM_024576197.1
Myotis myotis CPSF6	XM_036298067.1
Phyllostomus discolor CPSF6	XM_028531694.2
Rousettus aegyptiacus CPSF6	XM_016133567.2
Pteropus vampyrus CPSF6	XM_011374011.2
Homo sapiens CPSF6	NM_001300947.2
Mus musculus CPSF6	NM_001310609.1
Oryctolagus cuniculus CPSF6	XM_008256767.2
Felis catus CPSF6	XM_023257220.1
Sus scrofa CPSF6	XM_021090900.1
Ovis aries CPSF6	XM_027967466.1
Capra hircus CPSF6	XM_018047962.1
Bos taurus CPSF6	XM_005206464.4
Pan troglodytes CPSF6	XM_009425807.3
Gorilla gorilla CPSF6	XM_019038993.1
Pongo abelii CPSF6	XM_024256338.1
Papio Anubis CPSF6	XM_009181190.3
Macaca mulatta CPSF6	XM_015152234.2
Callithrix jacchus CPSF6	XM_009004183.2

Saimiri boliviensis CPSF6	XM_010338496.1
Aotus nancymaae CPSF6	XM_021665939.1
*: These samples were included in the calculation of dN/dS ratio for Fig. 8 together with FBKT-1 and R06E Mx2 sequences.	

SRA accession numbers (Fig. 7E)
SRA accession number
SRR2913598
SRR2914059
SRR2914282
SRR2914295
SRR2914359
SRR2914360
SRR2914366
SRR2914368
SRR2914369
SRR2914370
SRR2914371
SRR2914372
SRR7609226
SRR7548028

SRR7548029
SRR7548030
SRR7548031
SRR7548032
SRR7548033
SRR7548034
SRR7548035
SRR7548036
SRR7548037
SRR7548038
SRR7548039
SRR7548040
SRR7548041
SRR7548042
SRR7548043
SRR7548044
SRR7548045
SRR7548046
SRR7548047
SRR7548048
SRR7548049

SRR7548050
SRR7548051
SRR7548052
SRR7548053
SRR7548054
SRR7548055
SRR7548056
SRR7548057
SRR7548058
SRR7548059
SRR7548060
SRR7548061
SRR7548062
SRR7548063
SRR6453215
SRR6453209
SRR6453210
SRR6453214
SRR6453216
SRR6453213
SRR6453212

SRR6453208
SRR6453211
SRR7735101
SRR7735102

Supplementary Information References

1. Lee KE, Ambrose Z, Martin TD, Oztop I, Mulky A, Julias JG, et al. Flexible use of nuclear import pathways by HIV-1. Cell Host Microbe. 2010;7(3):221–233.

2. Engelman AN. HIV capsid and integration targeting. Viruses [Internet]. 2021 Jan 18 [cited 2021 May 31];13(1):125. Available from: https://www.mdpi.com/1999-4915/13/1/125