

# Supplementary Materials for

## A Reptilian Endogenous Foamy Virus Sheds Light on the Early Evolution of Retroviruses

Xiaoman Wei, Yicong Chen, Guangqian Duan, Edward C. Holmes and Jie Cui

### The file includes:

- **Fig. S1.** Detailed descriptions of consensus ERV-Spuma-Spu genome.
- **Fig. S2.** Conserved domain alignment of the consensus ERV-Spuma-Spu proteins and foamy virus proteins in CDD (Conserved Domain Database).
- **Fig. S3.** Conserved domain alignment of the original ERVs-Spuma-Spu proteins and foamy virus proteins in CDD (Conserved Domain Database).
- **Fig. S4.** Phylogenetic tree of foamy viruses and foamy-like viruses.
- **Table S1.** The information of 28 reptile genomes used for data mining.
- **Table S2.** The information of 130 bird genomes used for data mining.
- **Table S3.** The information of representative retroviruses used for phylogenetic analysis.
- **Table S4.** The important ERVs-Spuma-Spu sequences used for constructing consensus genome.
- **Table S5.** Dating of ERVs-Spuma-Spu insertion based on LTR-LTR divergence.
- **Table S6.** The matching contigs identified in *Sphenodon punctatus* genome.
- **References.**

## ERV-Spuma-Spu consensus genome (ERV-Spuma.0-Spu)

| 5'LTR start

1 ACTCAGGCAGGGCAGGAAAAGGCCTTGTGTTGTTGGAGTTTACTCATTATTAATCCTATCTTATACGATAGGAACCTGCTTAGTATGTAT

91 CAACAGACGTCTTACTATGTATTTTGAGAAAAAACCTCCCTGTCAGCAGTCGTGACCTGACGTCATTTTTAGCTAACCTCCAAGCATCT

181 TTCATTCTATAAGAAGAGAGAGGAAAGATGGTTTTTTCAGAAAAGCTTCCTTTCCCTGAAGCTTTTCTCATCTTCCATCTGGCAGAATCGGG

271 TAAAGGACTGGGTAATATGTTATATCCAATATGATGCACTGAGTCAACCATAGCAATAATACTTAGAGCATATGTTATATCTTATAGATG

361 TTAAGATATTTGAATTTGTAACAGAGCTGTTCTTTGAACAAGACATGTCTTTCTAAGTCATGTACTTTTAACTGGCACTGTTTCTTTTGT

451 ACTATGTCTTTATTTGAACTTAACTATTGCTTCATATTTTTATAATATTCTTAAAGTCTCATAGCCTCATTATATAAATGAATTGGTCT

541 TTTCTATAATTTGCTGAGCTGGCTGGCTGATTAGCAAAGTGAAACTCGAATTCACATATTGGAAAAGAGCGCCCTATATCCAGACCAG

5'LTR end |

PBS

631 ATCATGTCAGGCCTGGACCACTAAGCGCAAATTCCTTTTGGAGAACTGAGGCACAACGTGACAATTGGCACCCATTTCGTGGGGCTCGAG

721 ACATAAGAATATATAATTAAGTACACTGGGGTGTCCACCAGACCGTAACTACTTTTGTCTCTTGATATATTCAAGTTGTAAAGCTTA

| Gag start

M A I Q L N Q V P Y A L W V A N L Q N

811 ACTAAAAGAAAAGGATTTTGTGCATTTTATATCATGGCTATTCAATTAACCAAGTGCCTTATGCACTATGGGTTGCTAATCTGCAAAAT

V T V R D G D D Y C L E I R N G E W G I G H R F L I V S F E

901 GTAACAGTCAGAGATGGAGATGATTATTGTTTAGAAAATAAGAAATGGAGAATGGGGAATTGGACACAGATTCTAATTGTTTCTTTTGAA

A N D A G V I V S M R L R D V V F N P L T M P V L P I N R Q

991 GCAAATGATGCAGGTGTAATAGTTTCTATGAGACTAAGAGATGTAGTTTTTAATCCATTGACAATGCCTGTGTTACCTATCAATAGGCAA

D L N L M A G I V V E I P K N I R R H G P I S I S N A D Y I

1081 GATTTGAACCTAATGGCAGGGATAGTAGTTGAAATTCCEAAAAATATAAGGAGGCATGGGCCTATTTCTATAAGTAATGCTGATTATATT

S G R Y S S Y H R G I A W L Q C S P V G T G V H S I K K R I

1171 TCTGGAAGATACTCCAGCTACCACCGAGGAATTGCGTGGCTTCAATGCAGCCCAGTGGGAACAGGAGTACACTCTATTAAGAAGAGAATA

F K S E E F A K S N V G R Q P I V P G V H P A E I S V L N M

1261 TTTAAATCTGAGGAATTTGCAAAAAGTAATGTTGGAAGACAACCTATTGTGCCAGGAGTACACCCAGCTGAAATTTCTGTCTTAAATATG

A N I R A V T G P T P K D F K E I P M W F E S H L S A L E A

1351 GCAAATATAAGGGCTGTAACCTGGGCCACTCCAAAGGATTTTAAAGAAATTCCTATGTGGTTTGAATCCATTTATCTGCTTTAGAAGCT

V T S T A S P L Q K M R L C N S L V P A S A S L I E Q E C N

1441 GTCACTTCCACAGCATACCATTGCAGAAGATGAGGTTATGCAATTCCTCTAGTCCCAGCATCTGCTTCTTTGATAGAGCAAGAATGTAAT

N W E S V L A N L Y V K T H G Q V G I A D L N E I L R K I T  
1531 AATTGGGAAAGTGTATTGGCTAATTTATATGTA AAAA ACTCATGGCCAAGTGGGCATAGCTGACTTAAATGAAATCTAAGAAAAGATAACT

Q E Q G I V R A Y G V G M K F L S N H D L I W G I L K A L C  
1621 CAAGAGCAAGGCATTGTAAGGGCCTATGGGGTTGGAATGAAATTTCTCTCTAATCATGATTTGATATGGGGAATCCTTAAAGCCCTCTGT

K G D V L K A A I Q S K L D L L I T E Q E K I R S F P K I V  
1711 AAAGGTGATGTATTGAAAGCTGCCATTCAATCTAAATTAGATTTGCTGATTACTGAGCAGGAAAAATAAGGTCCTTCCCAAAAATTGTC

Q D I Y K T L G R D Y L G N N P N R K N L E Q E G S K S K K  
1801 CAGGATATTTATAAGACTCTGGGAGAGATTATTTGGGAAATAATCCAAACAGAAAGAATCTTGAACAGGAAGGTTCTAAATCTAAGAAA

S P S N I N S R K K I L P Q N S Q P N Y R G K G G K I Q G S  
1891 TCTCCCTCAAATATTAATTTCTAGAAAAAAATTCTTCCCAGAATTCTCAGCCAAATTATAGAGGCAAAGGAGGGAAGATCCAAGGCTCA

N R K Q F Q A Q K D Q D S E N P E G A A T Y D L R K G S Q F  
1981 AATAGAAAGCAGTTCCAAGCTCAGAAAGACCAGGATTCTGAAAAATCCAGAAGGTGCAGCTACATATGACTTAAAGAAAAGGATCACAGTTT

**Gag end |**

P H R D F D K S K Q D F N S R G G F R G N K \*  
2071 CCCCATAGAGACTTTGATAAATCAAAGCAAGATTTTAATTCAGAGGTGGATTTCAGAGGAAATAAATGAGGTTTAATAAAAAGAGGAAAT

**| Pol start**

**M A E N R Q**

2161 TATCAGACCTCTAAAGACTATGATCAACTGACATCCTCACAAAAAGCTTCAGAAACTCCCTCTGTGACACAGGATGGCAGAGAACAGACA

A Q V S L E F P M I S V E I A D K A F T A D R H R G K S I L  
2251 AGCTCAGGTCAGCCTTGAATTTCCAATGATCTCTGTAGAAATAGCAGATAAGGCTTTCACAGCTGATAGACACAGGGGCAATCAATCTT

Y S I K M F A T C I Y N S K V N V Y V E S Y F L E E K V V P  
2341 GTATTCAATCAAATGTTTGCCACATGTATATACAACTCAAAGGTCAATGTATACGTTGAGTCTTATTTTTAGAAAGAGAAAAGTAGTACC

**Aspartic protease**

L V K L P V Q I A G Q C L E I E F A I C D F T K H D V V I A  
2431 TCTTGTA AAAATTACCTGTTCAGATAGCAGGACAATGTTTGGA AATTGAATTTGCAATATGTGATTTTACAAAACATGATGTTGTCATAGC

H E R V K D L F P I G S I N I I G T Q D N R G E Q I K E Q I  
2521 TCATGAAAGGGTTAAGGATTTGTTTCCAATAGGTTCAATAAACATAATTGGA ACTCAAGATAATAGGGGTGAGCAAATCAAAGAGCAAAT

A S A D C A K N E K T K L R D I L Y S L K P Y F Q Q F D N Q  
2611 TGCATCAGCTGACTGTGCAAAAAATGAAAAACAAAATTGAGAGATATTTTATACTCTCTGAAACCATATTTTCAGCAATTTGATAATCA

**Q K Q Y P I N K A A I**

I G H R K I K P H D L S V K T Q P K P Q K Q Y P I N K A A I  
2701 AATTGGCCACAGGAAAATAAGCCCTCATGATTTATCTGTCAAAACACAGCCCAAACCGCAAAAAGCAATATCCTATTAACAAAAGCTGCAAT

**N D I Q K V I N D L I A Q G A L I R Q Y S S M N T P V Y P V**

2791 TAATGATATTCAAAAAGTAATTAATGATTTAATTGCACAAGGAGCATTGATTAGACAATATAGCTCCATGAATACTCCAGTATATCCAGT

P K P N G K W H M V L D Y R A L N R V S P S F N V Q N L H V

2881 GCCTAAACCTAATGGGAAATGGCATATGGTATTAGACTACCGCGCCCTAAACAGGGTTTCACCTCTTTTAATGTGCAAAACCTACACGT

W H A R K F G K A Q I Q N Y T G F I Q W I L G S S Y R E Q D

2971 CTGGCATGCTAGGAAATTTGGAAAGGCACAAATACAAAACACTACTGGATTTATCCAATGGATTTTGGGCTCATCCTATAGAGAACAAGA

**Reverse transcriptase domain**

R Y L T A F S W Q G T Q Y C W T R L P Q G Y L N S P A L F S

3061 CAGATATTTGACTGCGTTTAGTTGGCAGGGAACACAGTACTGTTGGACACGCCTCCACAAGGCTATTTAAATAGTCCAGCGTTGTTTTTC

A D V I Q L L K N I P G V H S Y M D D I Y F T N E D L D Q H

3151 AGCTGACGTAATTCAGCTTTTGGAAAAATATTCCAGGTGTACATTCATATATGGATGACATCTACTTTACAAATGAAGATTTAGATCAACA

L A T L K Q I V T V L G E A G Y I I N L K K S Q I C R S K V

3241 TTTAGCTACATTAAGCAAATTTGCACTGTTTTGGGAGAAGCAGGATATATTATCAATCTGAAAAAATCACAGATTTGTAGAAGCAAAGT

K F L G F L L T D S G R G L T Q E F K E K L L T L Q P P K T

3331 AAAGTTCCTTGGTTTCTTATTAACAGACTCTGGCAGAGGATTAACCTCAGGAATTCAAAGAAAAACTGTTGACATTGCAACCTCCTAAAAAC

L K E L Q S V L G F L N V A R I L V P D Y A Q R T K P L Y N

3421 TTTAAAAGAATTGCAGTCAGTTTTAGGTTTTTTAAATGTTGCAAGAATATTGGTTCCTGACTATGCTCAGAGAACCAAACCCCTTTATAA

L I P L A S K G N F W T L E A Q Q T L D N L I T L I N Q A V

3511 TCTAATTCATTAGCAAGCAAAGGAATTTTGGACCTTGAAGCTCAGCAAACATTGGATAATTTGATAACCTTAATTAACCAAGCTGT

E L N T R N S T V S L E I L V G A T Q K G G F A S Y F N Y G

3601 TGAATTAACACAAGAAACAGCACAGTTTCCCTAGAAATTCTAGTTGGAGCCACACAAAAAGGAGGATTTGCTTCATATTTCAATTATGG

E S K P L Q Y I S Y V F S N A E Q K F L P I E R I L C M C N

3691 AGAATCTAACCTCTGCAGTACATATCATATGTCTTCTCAAATGCAGAACAAAAGTTTTTGCCTATTGAAAGAATCTTATGTATGTGTAA

L A I L K G K D L A Q G Q K M I V K T P I A S L R Q V K K G

3781 TCTTGCAATTTTAAAAGGTAAAGATCTTGACAGGGCCAAAAGATGATAGTAAAAACTCCTATTGCCTCCTTAAGGCAGGTTAAAAAAGG

S I P N A K A L H S R W V Q W M S H F E N P Q I E F E Y V E

3871 ATCAATACCCAATGCAAAAGCTCTTCATAGCAGATGGGTACAATGGATGTCACATTTTGAAAATCCACAAATAGAATTTGAATATGTGGA

P S N D L E N L P A F T I E P V S S K N H T P D K P L H E Y

3961 ACCAAGCAATGATTTAGAAAATTTGCCAGCTTTTACTATTGAACCTGTATCTTCTAAAAATCATACCCAGATAAACCTCCTCATGAATA

Q K V I Y T D G S A M S C K Q G K H M W K A G C A V V I G T

4051 TCAAAAAGTTATCTATACTGATGGTTCAGCCATGTCATGTAAGCAAGGCAACACATGTGGAAAGCAGGTTGTGCTGTTGTAATTTGGAC

F N D K G E Y H M S D S I Q M P L G N N T A Q Y A E L M A V

4141 ATTTAATGATAAAGGGGAATATCATATGTCTGACAGTATTCAAATGCCTTTAGGAAACAATACTGCGCAATACGCTGAGTTGATGGCAGT

**RNaseH domain**

H K A I E I S P P D E T V L I C T D S F Y I A R G I N E E L

4231 ACATAAAGCAATAGAGATCTCTCCTCTGATGAAACTGTTCTCATTGTACTGATTCCTTCTATATAGCTAGAGGAATAAATGAAGAATT

P I W R S N G F L D N K R K P L K H A H R W Q K L A T L L D

4321 GCCAATCTGGCGGTCTAATGGTTTTCTAGATAATAAAAGAAAACCTTTGAAACATGCCCATAGATGGCAAAAATTAGCAACACTTTTGA

D K P L I T V M H V P G H S K Y G S H V N G N T L V D L L A

4411 TGACAAGCCATTGATTACTGTAATGCACGTGCCAGGCCATTCCAAATATGGATCTCATGTTAACGGGAACACCCTTGTAGACTTACTGGC

K E A M K A S S V C V L T R S Q V K K C L D G E L T Q C I S

4501 AAAGGAAGCAATGAAAGCTTCTCAGTCTGTACTTACCCGGTCACAAGTTAAAAAATGCTTGGATGGGGAATTGACTCAATGCATCAG

P D S I N P K G Y P S A Y D Y A L K D G K C V V T F T N G E

4591 CCCAGATTCCATTAACCCCTAAGGTTACCCCTTCTGCCTATGATTATGCCCTAAAAGATGGGAAGTGTGTTGTAACATTACTAATGGGGA

K R V I P P V D T R P N L I Q E A H N S L G H V H Q G V N A

4681 GAAACGTGTAATACCCCTGTTGACACAAGACCTAACTTAATACAAGAGGCTCACAAATAGTCTTGGGCATGTTACCAAGGTGTTAATGC

T V E S L Q R S Y W W P G L R K Q V Q R H L A Q C E P C L R

4771 CACAGTTGAATCTTTACAGCGTTCTTACTGGTGGCCTGGACTGCGCAACAAGTCCAACGGCATTGGCCAGTGTGAACCTTGCTAAG

T N P G P V T R P P Y L K N P K P L S P F D K V Y M D Y I G

4861 AACAAAACCTGGTCCAGTTACCAGACCCTTACTTAAAAATCCTAAGCCTCTGTCTCCTTTTGATAAGGTGTATATGGATTATATTGG

P L P P S H G H N H C L V I V D A C T G F V W I Y P T R D Q

4951 TCCTTTGCCACCTTCCATGGGCATAATCACTGTTTGGTTATTGTTGATGCTTGTACTGGTTTTGTATGGATTACCCACACGAGATCA

**Integrase core domain**

S A S T T V R T L T S F I S L G L P C I L H S D K G G A F T

5041 ATCTGCCTCTACCACTGTAGAACTCTCACCTCCTCATTTCGCTTGGGCTTCCATGTATCCTACACTCTGACAAGGGAGGTGCCTTCAC

S H Q M Q S F A K S F G I V L E Y S T P Y H P Q S A G V V E

5131 CTCTACCAAATGCAAAGTTTTGCAAAGAGTTTTGGGATAGTGTGGAATATAGCACACCTTATCACCCCAAAGTGCTGGGTTGTGGA

R K N G E I K R A L M K L L V G R S R Q W Y S L L P L V Q L

5221 AAGGAAAAATGGAGAGATAAACGAGCTTTAATGAAGCTATTGGTGGGAGGTCCCGCAGTGGTATTCTTTGCTTCTCTGGTACAGCT

G L N N L P R S D C H L T P Y K L L F A K D M T T P L E Q L

5311 TGGACTTAATAACCTTCTAGGAGTGACTGTCACCTAACCTTATAAATTGCTATTTGCAAAGGACATGACCACTCCTTTAGAACAACT

A L S S P I S Q Q E Q L A L I D E L R E E K N P V P I L K V

5401 GGCTTTATCTTCTCTATCTCTCAACAGGAGCAGTTGGCACTGATAGATGAATTGAGAGAAGAAAAAACCTGTACCTATTCTCAAAGT

L S P R A V E I Q T S P G N S K I V S I D N L K R T P I H Y  
5491 TTTGAGTCCCCGAGCAGTAGAAATTCAGACATCACCAGGTAATTCGAAAATTGTGTCTATTGACAACCTAAAGCGCACTCCCATCCACTA  
Pol end |  
G Q G N D G L L S N S T S T P T S P T N S T \*  
| Env start  
M V R V M M D Y C Q I P L Q P Q P L P P T A P D Y Y N F Y Y  
5581 TGGTCAGGTAATGATGGATTATTGTCAAATTCACCTTCAACCCCAACCTCTCCACCAACAGCACCTGACTATTATAACTTTTACTATT  
Y F F S P S F N M G S Y C F I T R Y S N S L E P L R G T Q W  
5671 ATTTTTTCAGTCTTCGTTAACATGGGCTCATACTGTTTCATTACCCGCTACTCCAATTCATTGGAACCTCTCCGAGGCACACAATGGT  
Y V V Y R R M H N A R P K R A L H I E V V P V L V E T A G I  
5761 ATGTTGTGTATCGAAGGATGCATAATGCCAGACCTAAACGTCTTTCATATAGAGGTTGTTCCAGTACTCGTGAGACTGCAGGTATTC  
P F G I I H N P F P K P I V S Q R S E L L V P F T L N I D T  
5851 CATTGGGATAATTCATAATCCCTTTCCTAAGCCTATTGTTTCTCAACGAAGTGAACCTGTTGGTTCCTTTTACTTTAAACATAGACACAA  
R A L A Y C S G L F S K D A N T H L A K T I E E D L Q D L D  
5941 GAGCACTTGCTTATTGTTCTGGCTGTTTTCAAAAGATGCTAACACACATCTTGCAAAGACTATAGAAGAAGACTTGCAAAGATTGGACA  
S R N A H F L V P G T D P W H Q T S Y A D K M C F A S Y G H  
6031 GTAGAAATGCACATTTTCTGTTCCAGGTACAGATCCTTGGCATCAAACAAGCTATGCTGATAAAATGTGCTTTGCTTCTATGGACATT  
C Y F V S Y G K P R K W P R P H V Y A D H C D R P Q F W T D  
6121 GTTATTTTGTTCCTATGGGAAACCGAGGAAATGGCCAAGACCTCATGTGTATGCAGATCACTGTGACAGACCTCAATTTTGGACTGACA  
I K T A T Q G L P Q W Y L A I D D F S D H L R Y A K Q Q R S  
6211 TTA AAACTGCAACACAAGGCCTGCCTCAGTGGTACCTAGCTATTGATGATTTTTCTGATCACTTGCCTATGCTAAGCAGCAACGTTCTG  
G G E D R E Y R V P G G Q L P Y T G A I F C T S F L Y N T S  
6301 GAGGTGAAGACCGTGAATACAGAGTTCAGGGGGCAATTGCCCTATACTGGAGCAATTTTTTGCACATCTTTTTTATATAATACCTCTT  
W W D E S N L S V D G S L E L K S I L T S C L A N S T T G K  
6391 GGTGGGATGAATCAAATTTATCTGTGGATGGCTCTTTAGAATTAAGTCCATTTTGACTTCTGTCTTGCAAATTTCTACTACTGGGAAAC  
L K P K C L A S Q W H D N G A N E M F I G V T G T S F C D I  
6481 TCAAACCTAAGTGTAGCCTCTCAATGGCATGATAATGGTGCTAATGAGATGTTTATTGGTGTTACTGGGACTTCATTTTGTGATATCC  
P R Y P I F L N R S E S I V S C K S T F V N P R Q Q P L E C  
6571 CTAGGTATCCTATTTTTCTCAATAGATCTGAGAGTATTGTATCTTGTA AAAAGTACTTTTTGTTAATCTCGACAACAGCCTCTTGAATGTG  
G N N K T L A A K G L H S W N C G P C S V N I T A N L M G N  
6661 GCAATAATAAAACACTGCTGCTAAGGGTTTACATTCATGGAATTGTGGGCCATGTAGTGTTAATATTACTGCAAATCTCATGGGTAATT

Y T A K E R A S L G N K R W F N L I Q G P L F V N A T P F F  
6751 ATACAGCCAAAGAGAGGGCGAGTTTGGGAAATAAGCGTTGGTTAACTTGATACAAGGCCATTATTTGTCAACGCTACTCCATTTTTTTG

A D N Y A I Y S L Y Q K C K T L S E K Y S L F S V L Q A L E  
6841 CTGATAATTATGCCATTTATTCGCTGTATCAGAAATGTA AACCTTATCTGAAAAATATTCTTTATTTTCAGTGCTTCAAGCATTAGAAG

E F I M V P Q E N E D Y P C T H S C I N A S L L Q M N P K R  
6931 AATTTATCATGGTACCTCAGGAAAATGAAGATTACCCTTGACACATTCTTGTATTAATGCTTCTTTATTACAGATGAACCCCAAAAGAG

A I W G T N K T L N D I H I L A T P D A T S E S F P S S T Y  
7021 CAATATGGGGCACAAATAAACTCTTAATGATATTCATATTTTAGCAACACCTGATGCTACTTCAGAGAGTTTCCCTTCTTCCACTTATA

K S R K I L Q V E S L N N A Q I F R K T N F L L A K S M E K  
7111 AATCTAGGAAGATTTTACAAGTAGAATCTTTGAATAATGCACAAATATTTAGAAAACTAATTTCTTGCTCGCAAAATCTATGGAAAAGA

I S R L Q D A N N I N L R N G V Y L V K D A L T Q V A L I V  
7201 TTTCTAGATTGCAGGATGCCAATAATATAAAATTTAAGAAATGGAGTTTATTTGGTCAAGGATGCTTAACTCAAGTGGCCTTGATAGTCA

K H D L A V L S D E L I M E I I V T Q L Q K I I F S L S N G  
7291 AACATGATTTGGCAGTTTATCTGATGAATTGATTATGGAAATTATTGTCACCTCAGTTGCAAAAAATTATTTTTTCTTTAAGCAATGGGC

H V P W T I G Y F Q I L R D H S D L H G E Y P P V L T V T Q  
7381 ACGTTCCTGGACTATTGGTTATTTTCAGATATTACGAGATCACAGTGACCTTCAGGTGAATACCCACCTGTGCTCACTGTGACACAGC

H V L R M K L L P D S V L T N P L N M M V Y S Y Y L M E Y S  
7471 ACGTGCTTCGAATGAACTATTACCAGATTCAGTCTTACTAACCCCTAAACATGATGGTATATTCCTACTACTTGATGGAATATTCAT

F F Y S F E N F Y Y G L T N W N I L N M G F L V A T G T N I  
7561 TTTTTTATTCATTTGAAAATTTCTACTATGGTTAACTAATTTGGAATATTTTGAATATGGGTTTTTTAGTAGCTACTGGAACAAATATTG

A H A H M K L H M N M S P Q M F L E T Q H T Y I L T N V K I  
7651 CTCATGCTCATATGAAATTACATATGAATATGTCTCCACAGATGTTTTTGGAAACACAACATACTTACATACTGACCAATGTCAAGATTT

L D I C F A N K L H I S P C G L V P P S T T S C P I R I Q S  
7741 TGGATATTTGTTTTGCAAACAAATTACACATTTCCCTTGTTGGACTGGTACCACCTAGTACCCTTCTTGTCCCTATTAGGATTCAATCAA

S N I S F V F I D S L T N G S Y L I L A G K S E C N I P A L  
7831 GTAATATTTCAATTTGTTTTATTGATTCCCTTACAAATGGATCATATTTAATCCTTGCTGGAAAATCTGAATGCAATATTCCAGCTTTGC

Q P S I V T V N T T I T C Y G R Q L F P P P N L G S M H S Q  
7921 AACCATCTATAGTAACAGTGAATACCACAATCACTTGTTATGGCAGACAACCTTTCCACCTCCTAATTTGGGATCAATGCATTCTCAGG

V T F F V P H F S L Q F P L L T G I I A K L Q R S E I T L F

8011 TTACATTTTTGTACCTCACTTTTCTTTGCAATTTCCACTATTAACCTGGCATCATAGCTAAATTGCAAAGGTCAGAAATTACTCTTTTTA

N T H D A T E D I L Q E V K Q L L Q R I D I H E G D F P L W

8101 ATACACATGATGCAACTGAAGATATCTTGCAGGAAGTAAAGCAACTACTTCAAAGGATTGATATACATGAAGGAGATTTTCTCTGTGGC

L N R L A S A V S A A W P S L A N M A N S I A H A D T S I G

8191 TCAATAGATTGGCTTCAGCTGTTTCTGCAGCCTGGCCTTCTTTAGCTAACATGGCTAATCCATTGCTCACGCTGATACTTCTATAGGTA

T S I L G T G L Q I L T Y L K P I F I A I V L I V L L I I V

8281 CTTCAATTTTAGGAACGGGTTTACAAATACTTACCTACCTCAAACCAATCTTTATTGCTATAGTGTTGATTGTACTATTAATCATCGTTA

I K I F R F F S G L H L K K I S A P N E D W M S L P S P L V

8371 TTAAGATTTTtaggttcttctctgactgcactgaaaaaaatATCAGCCCCAACGAAGATTGGATGTCTCTCCCTCACCACCTGTGA

**Env end |**

K K V P V L C T F L M S S \*

8461 AGAAGTACCAGTGCCTTGTACCTTCTGATGTCTTCTGAAGACATGACTGATTGGACATTTGTTGAATTTCTTTCTGTATTGTATA

**| ORF1 start**

M Y P V Y R H I C L T I C A S I Y

8551 CTACATTTGAAACACATTCTCGCCAAAGAGAGGGTGTATGTATCCTGTATATAGGCATATATGTTTGACCATCTGTGCTTCAATATA

W K S A S L Y D G F L K E L S Y S D L S T E L L T R Y W K I

8641 TTGAAATCAGCCTCTCTGTATGATGGCTTCTTAAAAGAGCTGTCTACAGTGACCTTTCTACTGAACTCTTGACCAGATACTGGAAAAT

R L Y T G I D N V I K S V S Y E I N S F C R L V S R Y M T L

8731 CAGACTGTATACCGAATAGACAATGTAATCAAGTCTGTCTCTTACGAAATTAACAGTTTCTGTGACTAGTCTCAAGATACATGACTCT

R Q G R K R P C L F G V Y S L L I L S Y T I G T C L V C I N

8821 CAGGCAGGCAGGAAAAGGCCTTGTATTGAGGTTTACTCATTGTTAATCCTATCTTATACGATAGGAACCTGTTTAGTATGTATCAA

**3'LTR start**

**|ORF1 end |**

R L L T M Y F E K S P S L S A V V T \*

8911 CAGGCTTCTACTATGTATTTTGAGAAAAGCCCTCCCTATCAGCAGTTGTGACCTGACGTCATTTTTAGCTGACCTCCAAGCATCTTTC

9001 ATTCTATAAGAAGAGAGAGGAAAGATGGATTTTTTCAGAAAGCTTCCCTTCTGAAGCTTTTCTCATCTTCTATCTGGCAGAATCGGGTAA

9091 AGGACTGGGTAATATGTTATATCCAATATGATGCACTGAGTCAACCATAGCAATAATACTTAGAGCATATGTTATATCTTATAGATGTTA

9181 AGATATTTGAATTTGTAACAGCTGTTCTTTGAACAAGACATGTCTTCTAAGTCATGTACTTTTAACTGGCACTGTTCTTTTGTACTAT

9271 GCCTTTATTTGAACTTAACTATTGCTTCATATTTTTATAATATTCTTAAAGTCTCATAGCCTCATTATTTAAATGAATTGGTCTTTTCT

9361 ATATTTTGCTGAGCTGGCTGGCTGATTAGCAAAGTGAACCTCGAATTCACATATTGGAAAAGAGCGCCCTATATCCAGACCAGATCAT

**3'LTR end |**

9451 GTCAGGCCTGGACCACTAAACGCAAATTCCTTTTGGGAAACTGAGGCACAACGCGACAGGCC



**Fig. S1. Detailed descriptions of consensus ERV-Spuma-Spu genome.** Protein Open reading frame (ORF) of gag, pol, env and ORF1 locations were determined by sequence similarity to representative foamy viruses using BLASTP and based on the distribution of stop and start codons, determined by ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>). The conserved domains were identified by searching the conserved domain database ([www.ncbi.nlm.nih.gov/Structure/cdd](http://www.ncbi.nlm.nih.gov/Structure/cdd)) and highlighted in darker colors. The ERV-Spuma-Spu PBS is nearly identical to the PBS of human foamy virus (HFV) (ERV-Spuma-Spu is 5'-TGGC**A**CCCA**T**TCGTGGGG-3' and HFV is 5'-TGGCGCCCAACGTGGGG-3'). The inferred pre-substitution nucleotides are indicated in red.

(a)

**Spuma virus Gag domain (pfam03276)**

**E-value:9.16e-24**

Consensus\_seq 435 PIVPG<sup>vh</sup>PAEISVLNMANIRAVTGPTPKDFKEIPMWFESHLSALEAVTSTASPLQKMRLCNSLVPASA--SLIEQECNNW 512  
Cdd:pfam03276 291 PVPPP--PPVGAVIPIQHIRSVTGEPPRNPREIPIWLG RNAP AIDGVFPTTPDLRCRIINALLGGNLgISLTPGDCITW 368

513 ESVLANLYVKTHGQVGIADLNEILRKITQEQQIVRAYGVGMKFL-SNHDLIWGI LKALCKGDVLKAAIQSKLDDLITEQE 591  
369 DSAVATLFI RTYGGYPLHQLGNVLKGIADQEGVATAYTLGMMLSGQNYQLVSGIIRGYLPGQAVVTAMQRLDQEIDDQT 448

592 KIRSFPKIVQDIYKTLGRDYLGNPNRKNLEQEGSKSKKSPSNINSRKKILPQNSQ<sup>pn</sup>yRGGGKIQGSNRKQFQAQKDQ 671  
449 RAETF IQHLNAVYEILGLNARGQSIRASVTPQPRPSRGRGQSAPESQGPVNSG---RGRQCPAPGQNDRGSNIQNG 525

672 DSENPEGA<sup>at</sup>YDLRKGSGFPHR 693  
526 QENSSQGG--YNLRSRTYQPQR 545

(b)

**Spuma aspartic protease (A9) domain (pfam03539)**

**E-value:1.02e-08**

Consensus\_seq 864 EQIKEQIASADCAKNEKTKLRDILYSLKPYFQQFDNQIGHRKIKPHDLSVKT-QPKPQKQVPINKAA 929  
Cdd:pfam03539 95 EQQETLLQQSALSKEGKELLKFLKYDALWQHVENQVGHRRIKPHKIATGTIKPRPQKQYHINPKA 161

(c)

**Foamy virus envelope protein domain (pfam03408)**

**E-value:1.96e-77**

Consensus\_seq 1922 VYRRMHNARPKRALHIE-VVPVLVET--AGIPFGIHNPFKPIVSQRSELLVPFTLNIDTRALAYCSGLfSKDANTHLA 1998  
Cdd:pfam03408 111 VYQPLQTRRIARSLRMQhPVPKYIEV<sup>nm</sup>TSIPQGVVYEPHPEPIV<sup>VT</sup>TERVLGLSQVLMINSENIANNANL-TQEVKKLLA 189

1999 KTI EEDLQDLDSRNAHFLVPGTDPWHQTSYADKMCFASYGHCYFVSYGKPRKWP<sup>RP</sup>HVYADHCDRPFQWTDIKTATQGLP 2078  
190 EVVNEEMQSLSDVMIDFEIPLGDPRDQEYIHRKCYQEF<sup>AH</sup>CYLVKYKTPKSW<sup>TE</sup>GLIADQCPLPGYHAGLSYKQSIW 269

2079 QWYLAID----DFSDHLRYAkQQRSGgedrEYRVPGG--QLPYTGAI<sup>F</sup>CTSFLYn<sup>TS</sup>WWD<sup>esn</sup>IsVDGSLEL-KSILTS 2150  
270 DYYIKVEi<sup>trpa</sup>NWSSQAVYG-QARLG----SFYVPGKi<sup>r</sup>QNNYSHVLF<sup>C</sup>SDQLY-SKWYN----IENSIEQnEKFLLN 338

2151 CLANSTTGK--LKP<sup>K</sup>CLASQWHDNGANEMF IGVTGTSFCDIPRYPIFLNRSESIVS----CKSTFVNPRQQPLEC---- 2219  
339 KLDNLTGGSiLKKRALPKEWSSQGNALFKEINVL<sup>D</sup>CVCKPELVILLNTSYYSFSIwegdCNFTKNMISQLVPECegfy 418

2220 GNNKTLAakg<sup>L</sup>HSWNCGPC-SVNITANLMGNYTAKERA-----SLGNKRW<sup>F</sup>NLIQ-----GPLFVNATPFFADNYAI 2285  
419 NNSKWMH---MHPYACRFWR-SKNEKEETKCRPGEKEKCl<sup>yy</sup>pyqdSLESTYDFGFLA<sup>yq</sup>knfpAPICIEQQEIRDKDYEV 495

2286 YSLYQKCKTLEKYSLSV<sup>L</sup>QALEEFImvpqenedypc<sup>TH</sup>SeiNASLLQMNPKRAIWGTnktLNDIHILATPDATSESEFP 2365  
496 YSLYQECKLASKVHGIDTVL<sup>F</sup>SLKNFL-----NHT--GRPVNEMP<sup>N</sup>ARAFVGL---VDPKFPSPNVN<sup>T</sup>REHYT 559

2366 SSTYKSRKi<sup>lq</sup>vESLNAQIFRKTN<sup>F</sup>LAKSMEKISRLQDANNINLRNGVYLKVDAL<sup>T</sup>QVALIVK<sup>H</sup>DLAVLSDELIMEII 2445

560 SCNNRKR---STDNNYAKLKSMGYALTGAVQTLSDINDENLQGGIYLLRDHVITLMEATLHDISVMEGMFAVQHL 635

2446 VTQLQKIIFSLSNHVPWTI---GYFQILRDHSDIhGEYPPVLTVTQH-VLRMKLLPDSVLTNPLNMMVYsYYLMEYSFF 2521

636 HTHLNHLKTMLEERRIDWTYmSSAWLQQQLKSD--DEMKVIKRIAKSI VYVVKQTYNSPTATAWEIGLY-YELTIPKHV 712

2522 YsfenfyygLTNWNILNMGFLVATGTNIAHAMKLMNMSPQMLETQHTYiLTNVKILDICFANKLHI-SPCGI VPPST 2600

713 Y-----LNNWVNVNIGHLVQSAGQLHVTIAHPYEI INKECTETKYLH-LKDCRRQDYVICDVVEIvQPCG-NSTDT 782

2601 TSCPRIQSSNISFVIDSLTNGSYLILAGKSECNIPALQPSIVTVNTTITCYGRQIFPPPNLGSMHSQVTFVPHFSLQ 2680

783 SDCPVWAEAVKEPFVQVNPVKNGSYLVLASSTDCQIPPYVPSIVTVNETTSCYGLN-FKKPLVAEERLGFEPRLPNLQLR 861

2681 FPLLTGIIAKLQRSEITLFNTHDATEDILQEVKQLLQRIDIHEGDFPLWLNRLASAVSAAWPSLANMANSIAHADTSIGT 2760

862 LPHLVGIIAKIKGLKIEVTSSGESIKDIERAKAELLRLDIHEGDTPAWIQQLAAATKDVWPAASALQGI GNFLSAAH 941

2761 SILGTGLQILT YLKP 2775

942 GIFGTAFSLG YLKP 956

**Figure S2. Conserved domain alignment of the consensus ERV-Spuma-Spu proteins and foamy virus proteins in CDD (Conserved Domain Database).** (a) The alignment of consensus ERV-Spuma-Spu Gag proteins and Spuma virus Gag domain (pfam03276); (b) The alignment of consensus ERV-Spuma-Spu pro proteins and Spuma aspartic protease (A9) domain (pfam03539); (c) The alignment of consensus ERV-Spuma-Spu envelope proteins and foamy virus envelope protein domain (pfam03408). Numbers refer to the position in the consensus ERV-Spuma-Spu each protein or conserved domain. Identical amino acid residues are highlighted in red, and black and blue indicate gaps or different amino acid residues, respectively. The E-value was generated by Conserved Domain search.

(a)

**Spuma virus Gag domain (pfam03276)**

**E-value:2.16e-12**

QEPC01003194.1\_seq 488 RLLNSLVPASA--SLMELECNNLENVLANLYVKTHGXVGIADLNXLIRKIQEQGIVRAYVGMKFL-SNHDLIWGILKT 564  
Cdd:pfam03276 346 RIINALLGGNLgLSLTPGDCITWDSAVATLFIIRTYGQYPLHQLGNVLKGIADQEGVATAYTLGMMLSGQNYQLVSGIIRG 425

565 LCKGDALKAAIQSKLDLLITEQEQRISFPKIVQDIYKILGRDYLGDNPNRKNLEQEGsksksnSPSNTNFRKKILPQNSQ- 643  
426 YLPQAVVTAMQQRLDQEIDDQTRAETFQHLNAVVEILGLNARGQSIRASVTPQP-----RPSRGRGRGQSAPEPSQg 499

644 PNHRGKGGKI----QDSNRKQFQAQKDKDSenpeSAATYDLRN\*SQFPHR 689  
500 PVNSGRGRQCpapgQNDRGSNIQNGQENS----SQGGYNLRSRTYQQR 545

(b)

**Spuma aspartic protease (A9) domain (pfam03539)**

**E-value:8.72e-08**

QEPC01003194.1\_seq 859 EQIKEQIASADCAKNEKIKLRDILYTLKPYFQQFDNQIGHRKIKPHDLSVKT-QPKPQKQYPIKAA 924  
Cdd:pfam03539 95 EQQETLLQQSALSKEGKELKKLFLKYDALWQHWENQVGHRRIKPHKIATGTIKPRPQKQYHINPKA 161

(c)

**Foamy virus envelope protein domain (pfam03408)**

**E-value:7.69e-19**

QEPC01003194.1\_seq 1926 SVLRLTWAHTVSLPATPIHWNLSEahngTVVYRRMHNAPKRALHIE-VIPVLVET--AGMPFGIHNPFKPIILnqvnc 2002  
Cdd:pfam03408 85 TISRQWNRDIQVLGPVIDWNVTD----RAVYQPLQTRRIARSLRMqhPVPKYIEVnmTSlPQGVYVEPHPEPIV----- 155

2003 wclll\*t\*TSER\*piVLACSQKKL-----TNILQRL\*KMTCEIWTVEMH-----IFLFQVQI---LGIKQAMLKIC 2065  
156 -----VTER---VLGLSQVLMinseniannANLTQEVKLLAEVVNEEMQslsdvMIDFEIPLgdpRDQEQYIHRKC 224

2066 aLLLIHCYFVSDGKPRKWPRPHVYADHCDRPQFWTDIKTATQGLPQWYLAID----DFSDHLRYAkQQRSGgedrEYR 2140  
225 -YQEFACHYLVKYKTPKSWPTEGLIADQCPLPGYHAGLSYKPKQSIWDYIYKVEitrapaNWSSQAVYG-QARLG----SFY 298

2141 VPGG--QLPYTGAICTSFYnTSWWDesnlsVDGSLEL-KSVLTSCLNSTTGK--LKPCLSSQWHDNGANEMFVGV 2215  
299 VPKGirQNNYSHVLFCSQDLY-SKWYN-----IENSIEQnEKFLNKLNLTTGSsILKKRALPKWSSQGNALFKEIN 372

2216 GTSFCDIPRYPIFLNR-----SESIVSCKSTFVNHQ--QQPLECGTNKTLTAKgIHSWNCGPCSV 2273  
373 VLDVCSKPELVILLNTsyysfslwegdcnftknmiSQLVPECEGFYNSKwmhMHPYACRFWRSKNEK--EETKCRPGEK 450

2274 NITTNLMGNYTAKERVSLGNKRWFNLIQSPLFVNATPFFADNYAIYSLYQKCKTLESEKHSLSFVLTLE\*FI 2345  
451 EKCLYYPYQDSLESTYDFGLAYQKNFPAPICIEQQEIRDKDYEVYSLYQECKLASKVHGIDTVLFSKLNFL 522

**E-value:3.41e-19**

QEPC01003194.1\_seq 2683 EHAILQHTARALIFNMKLLPDSvlt nplnmwyiPT-TXWNIHFFIHLKISNMVYLTNWNILNMGFLVATGTNIAHAHMKL 2761  
Cdd:pfam03408 671 EMKVIKRIAKSLVYYVKQTYNS-----PTaTAWELGLYYELTIPKHVYLNWVNIHGLVQSAGQLTHVTIAH 739

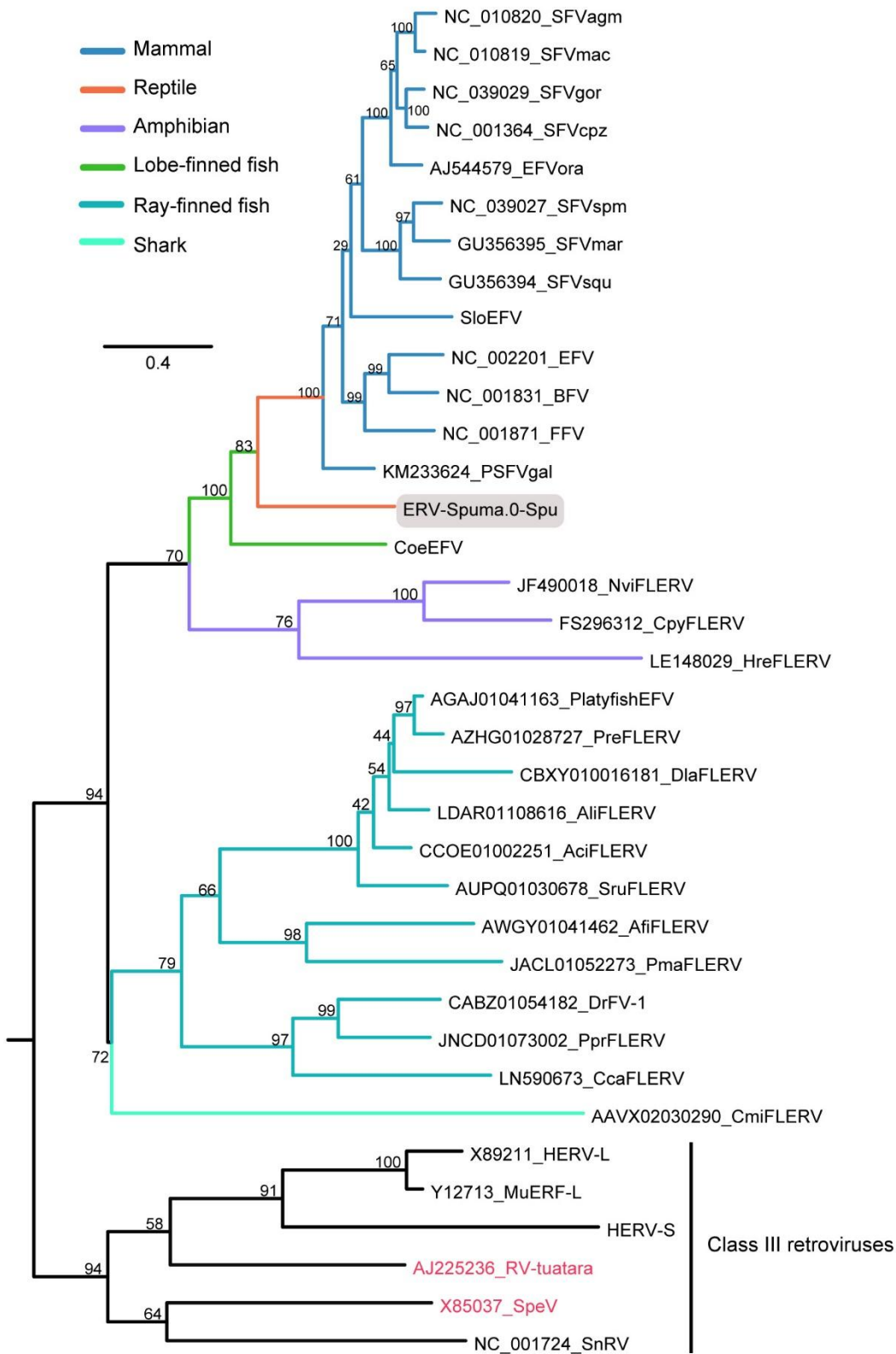
2762 PYEWSTDVfGNITTYLHTDECQDFGYLFCKQI-----THFPLWVgtt\*yyflsy\*DSIK\*YFICFY\*fPYKwi 2829  
740 PYEILINKEC-TE TKYLHLKDCRRQDYVICDVVeivqpcgnstdtSDCPVWA-----EAVKEPFVQVN--PLK-- 803

2830 ifNSCWTI\*m\*ysgfanIYHNSDYHSHLLWQTAFSTS\*fEINAFSGYIFRNSLF-----LTGIIA 2889  
804 --NGSYLV-----LASSTDCQIPPYVPSIVTVN--ETTSCYGLNFKKPLVaeerlgfeprlpnlqlrlphLVGIIA 870

2890 KLQRSEITLFINHDATEDTLQEVKQLLQRIDIEHGD FPLWFNRLATAISAAWPSLASMHSIAHAATSLGNSILGTGLQI 2969  
871 KIKGLKIEVTSSGESIKDQIERAKAELLRLDIEHGDTPAWIQQLAATKDVWPAASALQGI GNFLSGAAHGIFGTAFSL 950

2970 ITYLPK 2975  
951 LGYLPK 956

**Figure S3. Conserved domain alignment of the original ERVs-Spuma-Spu proteins and foamy virus proteins in CDD (Conserved Domain Database).** (a) The alignment of original ERV-Spuma-Spu Gag proteins and Spuma virus Gag domain (pfam03276); (b) The alignment of original ERV-Spuma-Spu pro proteins and Spuma aspartic protease (A9) domain (pfam03539); (c) The two parts of alignment of original ERV-Spuma-Spu envelope proteins and foamy virus envelope protein domain (pfam03408). Numbers refer to the position in the original ERV-Spuma-Spu each protein or conserved domain. Identical amino acid residues are highlighted in red, and black and blue indicate gaps or different amino acid residues, respectively. The E-value was generated by Conserved Domain search.



**Fig S4. Phylogenetic tree of foamy viruses and foamy-like viruses.** ML tree of foamy viruses and foamy-like virus including consensus sequences of ERV-Spuma-Spu (ERV-Spuma.0-Spu), inferred

using amino acid sequences of the Pol gene. Preceding viral names are the contig accession numbers containing viral sequences. Class III retroviruses were used to root the viral tree. The consensus ERV-Spuma-Spu was labelled using a grey-shaded box. Previously reported sequences found in tuatara were labelled in red. The scale bar indicates the number of amino acid changes per site.

**Table S1. The information of 28 reptiles used for data mining**

| No. | Species name                        | Common name                       | Accession no.   |
|-----|-------------------------------------|-----------------------------------|-----------------|
| 1   | <i>Alligator mississippiensis</i>   | American alligator                | GCA_000281125.4 |
| 2   | <i>Alligator sinensis</i>           | Chinese alligator                 | GCA_000455745.1 |
| 3   | <i>Gavialis gangeticus</i>          | Gharial                           | GCA_001723915.1 |
| 4   | <i>Gekko japonicus</i>              | Schlegel's Japanese gecko         | GCA_001447785.1 |
| 5   | <i>Lacerta bilineata</i>            | Western green lizard              | GCA_900245895.1 |
| 6   | <i>Lacerta viridis</i>              | European green lizard             | GCA_900245905.1 |
| 7   | <i>Paroedura picta</i>              | Ocelot gecko                      | GCA_003118565.1 |
| 8   | <i>Sphenodon punctatus</i>          | Tuatara                           | GCA_003113815.1 |
| 9   | <i>Anolis carolinensis</i>          | Green anole                       | GCA_000090745.2 |
| 10  | <i>Apalone spinifera</i>            | Spiny softshell turtle            | GCA_000385615.1 |
| 11  | <i>Chelonia mydas</i>               | Green sea turtle                  | GCA_000344595.1 |
| 12  | <i>Chrysemys picta</i>              | Painted turtle                    | GCA_000241765.2 |
| 13  | <i>Crocodylus porosus</i>           | Saltwater crocodile               | GCA_001723895.1 |
| 14  | <i>Crotalus horridus</i>            | Timber rattlesnake                | GCA_001625485.1 |
| 15  | <i>Crotalus pyrrhus</i>             | Southwestern speckled rattlesnake | GCA_000737285.1 |
| 16  | <i>Crotalus viridis</i>             | Prairie rattlesnake               | GCA_003400415.1 |
| 17  | <i>Gopherus agassizii</i>           | Desert tortoise                   | GCA_002896415.1 |
| 18  | <i>Malaclemys terrapin</i>          | Diamondback terrapin              | GCA_001728815.2 |
| 19  | <i>Ophiophagus hannah</i>           | King cobra                        | GCA_000516915.1 |
| 20  | <i>Pantherophis guttatus</i>        | Corn snake                        | GCA_001185365.1 |
| 21  | <i>Pelodiscus sinensis</i>          | Chinese softshell turtle          | GCA_000230535.1 |
| 22  | <i>Pogona vitticeps</i>             | Central bearded dragon            | GCA_900067755.1 |
| 23  | <i>Protobothrops flavoviridis</i>   | Okinawa habu                      | GCA_003402635.1 |
| 24  | <i>Protobothrops mucrosquamatus</i> | Brown spotted pit viper           | GCA_001527695.3 |
| 25  | <i>Python bivittatus</i>            | Burmese python                    | GCA_000186305.2 |
| 26  | <i>Terrapene mexicana</i>           | Mexican box turtle                | GCA_002925995.2 |
| 27  | <i>Thamnophis sirtalis</i>          | Common garter snake               | GCA_001077635.2 |
| 28  | <i>Vipera berus</i>                 | Common European adder             | GCA_000800605.1 |

**Table S2. The information of 130 bird genomes used for data mining**

| <b>No.</b> | <b>Species name</b>                    | <b>Common name</b>              | <b>Accession no.</b> |
|------------|--|---------------------------------|----------------------|
| 1          | <i>Acanthisitta chloris</i>            | Rifleman                        | GCA_000695815.1      |
| 2          | <i>Acridotheres javanicus</i>          | Javan myna                      | GCA_002849675.1      |
| 3          | <i>Agapornis roseicollis</i>           | Peach-faced lovebird            | GCA_002631895.1      |
| 4          | <i>Amazona aestiva</i>                 | Blue-fronted amazon             | GCA_001420675.1      |
| 5          | <i>Amazona vittata</i>                 | Puerto Rican parrot             | GCA_000332375.1      |
| 6          | <i>Anas platyrhynchos</i>              | Mallard                         | GCA_000355885.1      |
| 7          | <i>Anas zonorhyncha</i>                | Eastern spot-billed duck        | GCA_002224875.1      |
| 8          | <i>Anser brachyrhynchus</i>            | Pink-footed goose               | GCA_002592135.1      |
| 9          | <i>Anser cygnoides domesticus</i>      | Domestic goose                  | GCA_000971095.1      |
| 10         | <i>Antrostomus carolinensis</i>        | Chuck-will's-widow              | GCA_000700745.1      |
| 11         | <i>Apaloderma vittatum</i>             | Bar-tailed trogon               | GCA_000703405.1      |
| 12         | <i>Aptenodytes forsteri</i>            | Emperor penguin                 | GCA_000699145.1      |
| 13         | <i>Apteryx australis mantelli</i>      | Southern brown kiwi             | GCA_001039765.1      |
| 14         | <i>Apteryx haastii</i>                 | Great spotted kiwi              | GCA_003342985.1      |
| 15         | <i>Apteryx owenii</i>                  | Little spotted kiwi             | GCA_003342965.1      |
| 16         | <i>Apteryx rowi</i>                    | Okarito brown kiwi              | GCA_003343035.1      |
| 17         | <i>Aquila chrysaetos canadensis</i>    | Golden eagle                    | GCA_000766835.1      |
| 18         | <i>Ara macao</i>                       | Scarlet macaw                   | GCA_000400695.1      |
| 19         | <i>Athene cunicularia</i>              | Burrowing owl                   | GCA_003259725.1      |
| 20         | <i>Balearica regulorum gibbericeps</i> | East African grey crowned-crane | GCA_000709895.1      |
| 21         | <i>Bambusicola thoracicus</i>          | Chinese bamboo-partridge        | GCA_002909625.1      |
| 22         | <i>Buceros rhinoceros silvestris</i>   | Rhinoceros hornbill             | GCA_000710305.1      |
| 23         | <i>Calidris pugnax</i>                 | Ruff                            | GCA_001431845.1      |
| 24         | <i>Calidris pygmaea</i>                | Spoon-billed sandpiper          | GCA_003697955.1      |
| 25         | <i>Callipepla squamata</i>             | Scaled quail                    | GCA_002218305.1      |
| 26         | <i>Calypte anna</i>                    | Anna's hummingbird              | GCA_000699085.1      |
| 27         | <i>Cariama cristata</i>                | Red-legged seriema              | GCA_000690535.1      |
| 28         | <i>Casuaris casuaris</i>               | Southern cassowary              | GCA_003342895.1      |
| 29         | <i>Cathartes aura</i>                  | Turkey vulture                  | GCA_000699945.1      |
| 30         | <i>Chaetura pelagica</i>               | Chimney swift                   | GCA_000747805.1      |
| 31         | <i>Charadrius vociferus</i>            | Killdeer                        | GCA_000708025.2      |
| 32         | <i>Chlamydotis macqueenii</i>          | Macqueen's bustard              | GCA_000695195.1      |
| 33         | <i>Chlamydotis undulata undulata</i>   | Houbara bustard                 | GCA_003400225.1      |
| 34         | <i>Chrysolophus pictus</i>             | Golden pheasant                 | GCA_003413605.1      |
| 35         | <i>Cicinnurus regius</i>               | King bird of paradise           | GCA_003713305.1      |
| 36         | <i>Ciconia boyciana</i>                | Oriental stork                  | GCA_002002965.1      |
| 37         | <i>Colinus virginianus</i>             | Northern bobwhite               | GCA_000599465.2      |
| 38         | <i>Colius striatus</i>                 | Speckled mousebird              | GCA_000690715.1      |
| 39         | <i>Columba livia</i>                   | Rock pigeon                     | GCA_000337935.1      |
| 40         | <i>Corvus brachyrhynchus</i>           | American crow                   | GCA_000691975.1      |
| 41         | <i>Corvus cornix cornix</i>            | Hooded crow                     | GCA_000738735.2      |
| 42         | <i>Corvus hawaiiensis</i>              | Hawaiian crow                   | GCA_003402825.1      |



|    |                                   |                                |                 |
|----|-----------------------------------|--------------------------------|-----------------|
| 43 | <i>Coturnix japonica</i>          | Japanese quail                 | GCA_001577835.1 |
| 44 | <i>Crypturellus cinnamomeus</i>   | Thicket tinamou                | GCA_003342915.1 |
| 45 | <i>Cuculus canorus</i>            | Common cuckoo                  | GCA_000709325.1 |
| 46 | <i>Cyanistes caeruleus</i>        | Blue tit                       | GCA_002901205.1 |
| 47 | <i>Diphyllodes magnificus</i>     | Magnificent bird-of-paradise   | GCA_003713285.1 |
| 48 | <i>Dromaius novaehollandiae</i>   | Emu                            | GCA_003342905.1 |
| 49 | <i>Egretta garzetta</i>           | Little egret                   | GCA_000687185.1 |
| 50 | <i>Empidonax traillii</i>         | Willow flycatcher              | GCA_003031625.1 |
| 51 | <i>Eopsaltria australis</i>       | Eastern yellow robin           | GCA_003426825.1 |
| 52 | <i>Erythrura gouldiae</i>         | Gouldian finch                 | GCA_003676055.1 |
| 53 | <i>Eudromia elegans</i>           | Elegant crested-tinamou        | GCA_003342815.1 |
| 54 | <i>Eurypyga helias</i>            | Sunbittern                     | GCA_000690775.1 |
| 55 | <i>Falco cherrug</i>              | Saker falcon                   | GCA_000337975.1 |
| 56 | <i>Falco peregrinus</i>           | Peregrine falcon               | GCA_000337955.1 |
| 57 | <i>Ficedula albicollis</i>        | Collared flycatcher            | GCA_000247815.2 |
| 58 | <i>Fulmarus glacialis</i>         | Northern fulmar                | GCA_000690835.1 |
| 59 | <i>Gallirallus okinawae</i>       | Okinawa rail                   | GCA_002003005.1 |
| 60 | <i>Gallus gallus</i>              | Chicken                        | GCA_000002315.5 |
| 61 | <i>Gavia stellata</i>             | Red-throated loon              | GCA_000690875.1 |
| 62 | <i>Geospiza fortis</i>            | Medium ground-finch            | GCA_000277835.1 |
| 63 | <i>Grus japonensis</i>            | Red-crowned crane              | GCA_002002985.1 |
| 64 | <i>Haliaeetus albicilla</i>       | White-tailed eagle             | GCA_000691405.1 |
| 65 | <i>Haliaeetus leucocephalus</i>   | Bald eagle                     | GCA_000737465.1 |
| 66 | <i>Hemignathus virens</i>         | Hawaii amakihi                 | GCA_003286495.1 |
| 67 | <i>Hirundo rustica rustica</i>    | Barn swallow                   | GCA_003692655.1 |
| 68 | <i>Junco hyemalis</i>             | Dark-eyed junco                | GCA_003829775.1 |
| 69 | <i>Lepidothrix coronata</i>       | Blue-crowned manakin           | GCA_001604755.1 |
| 70 | <i>Leptosomus discolor</i>        | Cuckoo roller                  | GCA_000691785.1 |
| 71 | <i>Limosa lapponica baueri</i>    | Bar-tailed godwit              | GCA_002844005.1 |
| 72 | <i>Lonchura striata domestica</i> | Bengalese finch                | GCA_002197715.1 |
| 73 | <i>Lyrurus tetrix tetrix</i>      | Black grouse                   | GCA_000586395.1 |
| 74 | <i>Manacus vitellinus</i>         | Golden-collared manakin        | GCA_001715985.2 |
| 75 | <i>Meleagris gallopavo</i>        | Turkey                         | GCA_000146605.3 |
| 76 | <i>Melopsittacus undulatus</i>    | Budgerigar                     | GCA_000238935.1 |
| 77 | <i>Merops nubicus</i>             | Carmine bee-eater              | GCA_000691845.1 |
| 78 | <i>Mesitornis unicolor</i>        | Brown roatelo                  | GCA_000695765.1 |
| 79 | <i>Mixornis gularis</i>           | Striped tit-babbler            | GCA_003546035.1 |
| 80 | <i>Nannopterum auritus</i>        | Double-crested cormorant       | GCA_002173455.1 |
| 81 | <i>Nannopterum brasilianus</i>    | Neotropic cormorant            | GCA_002174335.1 |
| 82 | <i>Nannopterum harrisi</i>        | Galapagos flightless cormorant | GCA_002173475.1 |
| 83 | <i>Nestor notabilis</i>           | Kea                            | GCA_000696875.1 |
| 84 | <i>Nipponia nippon</i>            | Crested ibis                   | GCA_000708225.1 |
| 85 | <i>Nothoprocta perdicaria</i>     | Chilean tinamou                | GCA_003342845.1 |
| 86 | <i>Numida meleagris</i>           | Helmeted guineafowl            | GCA_002078875.2 |

|     |  |  |                 |
|-----|--|--|-----------------|
| 87  | <i>Opisthocomus hoazin</i>                 | Hoatzin                                | GCA_000692075.1 |
| 88  | <i>Paradisaea raggiana</i>                 | Raggiana bird of paradise              | GCA_003713265.1 |
| 89  | <i>Paradisaea rubra</i>                    | Red bird of paradise                   | GCA_003713215.1 |
| 90  | <i>Parotia lawesii</i>                     | Lawes's parotia                        | GCA_003713295.1 |
| 91  | <i>Parus major</i>                         | Great tit                              | GCA_001522545.2 |
| 92  | <i>Passer domesticus</i>                   | House sparrow                          | GCA_001700915.1 |
| 93  | <i>Patagioenas fasciata monilis</i>        | Band-tailed pigeon                     | GCA_002029285.1 |
| 94  | <i>Pelecanus crispus</i>                   | Dalmatian pelican                      | GCA_000687375.1 |
| 95  | <i>Phaethon lepturus</i>                   | White-tailed tropicbird                | GCA_000687285.1 |
| 96  | <i>Phalacrocorax carbo</i>                 | Great cormorant                        | GCA_000708925.1 |
| 97  | <i>Phoenicopterus ruber ruber</i>          | American flamingo                      | GCA_000687265.1 |
| 98  | <i>Phylloscopus plumbeitarsus</i>          | Two-barred warbler                     | GCA_001655115.1 |
| 99  | <i>Phylloscopus trochiloides viridanus</i> | Greenish warbler                       | GCA_001655095.1 |
| 100 | <i>Phylloscopus trochilus acredula</i>     | <i>Phylloscopus trochilus acredula</i> | GCA_002305835.1 |
| 101 | <i>Picoides pubescens</i>                  | Downy woodpecker                       | GCA_000699005.1 |
| 102 | <i>Podiceps cristatus</i>                  | Great crested grebe                    | GCA_000699545.1 |
| 103 | <i>Pseudopodoces humilis</i>               | Tibetan ground-tit                     | GCA_000331425.1 |
| 104 | <i>Psittacula krameri</i>                  | Rose-ringed parakeet                   | GCA_002870145.1 |
| 105 | <i>Pterocles gutturalis</i>                | Yellow-throated sandgrouse             | GCA_000699245.1 |
| 106 | <i>Pterocnemia pennata</i>                 | Darwin's rhea                          | GCA_003342835.1 |
| 107 | <i>Pygoscelis adeliae</i>                  | Adelie penguin                         | GCA_000699105.1 |
| 108 | <i>Pygoscelis antarcticus</i>              | Chinstrap penguin                      | GCA_003264595.1 |
| 109 | <i>Pygoscelis papua</i>                    | Gentoo penguin                         | GCA_003264615.1 |
| 110 | <i>Rhea americana</i>                      | Greater rhea                           | GCA_003343005.1 |
| 111 | <i>Saxicola maurus maurus</i>              | Siberian stonechat                     | GCA_900205225.1 |
| 112 | <i>Serinus canaria</i>                     | Common canary                          | GCA_000534875.1 |
| 113 | <i>Setophaga coronata coronata</i>         | Yellow-rumped warbler                  | GCA_001746935.1 |
| 114 | <i>Spheniscus humboldti</i>                | Humboldt's penguin                     | GCA_003264545.1 |
| 115 | <i>Spheniscus magellanicus</i>             | Magellanic penguin                     | GCA_003264715.1 |
| 116 | <i>Spheniscus mendiculus</i>               | Galapagos penguin                      | GCA_003264655.1 |
| 117 | <i>Sporophila hypoxantha</i>               | Tawny-bellied seedeater                | GCA_002167245.1 |
| 118 | <i>Strix occidentalis caurina</i>          | Spotted owl                            | GCA_002372975.1 |
| 119 | <i>Struthio camelus australis</i>          | African ostrich                        | GCA_000698965.1 |
| 120 | <i>Sturnus vulgaris</i>                    | Common starling                        | GCA_001447265.1 |
| 121 | <i>Syrmaticus mikado</i>                   | Mikado pheasant                        | GCA_003435085.1 |
| 122 | <i>Taeniopygia guttata</i>                 | Zebra finch                            | GCA_000151805.2 |
| 123 | <i>Tauraco erythrolophus</i>               | Red-crested turaco                     | GCA_000709365.1 |
| 124 | <i>Tinamus guttatus</i>                    | White-throated tinamou                 | GCA_000705375.2 |
| 125 | <i>Tympanuchus cupido pinnatus</i>         | Greater prairie chicken                | GCA_001870855.1 |
| 126 | <i>Tyto alba</i>                           | Barn owl                               | GCA_000687205.1 |
| 127 | <i>Uria lomvia</i>                         | Thick-billed guillemot                 | GCA_002289315.1 |
| 128 | <i>Urile pelagicus</i>                     | Pelagic cormorant                      | GCA_002173435.1 |
| 129 | <i>Zonotrichia albicollis</i>              | White-throated sparrow                 | GCA_000385455.1 |
| 130 | <i>Zosterops lateralis melanops</i>        | Silver-eye                             | GCA_001281735.1 |

---

**Table S3. The information of representative retroviruses used for phylogenetic analysis.**

| No. | Virus name                                 | Genus             | Abbreviation | Natural host             | Accession no.<br>/Reference |
|-----|--|-------------------|--------------|--------------------------|-----------------------------|
| 1   | Avian leukemia virus                       | Alpharetrovirus   | ALV          | Chicken                  | NC_015116                   |
| 2   | Lymphoproliferative disease virus          | Alpharetrovirus   | LDV          | Turkey                   | U09568                      |
| 3   | Mouse mammary tumor virus                  | Betaretrovirus    | MMTV         | Mouse                    | NC_001503                   |
| 4   | Mason-Pfizer monkey virus                  | Betaretrovirus    | MPMV         | Primate                  | NC_001550                   |
| 5   | Simian retrovirus 1                        | Betaretrovirus    | SRV1         | Primate                  | M11841                      |
| 6   | Bovine leukemia virus                      | Deltaretrovirus   | BLV          | Cattle                   | NC_001414                   |
| 7   | Human T-lymphotropic virus 1               | Deltaretrovirus   | HTLV1        | Human                    | NC_001436                   |
| 8   | Simian T-lymphotropic virus 2              | Deltaretrovirus   | STLV2        | Non-human primate        | NC_001815                   |
| 9   | Walleye dermal sarcoma virus               | Epsilonretrovirus | WDSV         | Fish                     | NC_001867                   |
| 10  | Walleye epidermal hyperplasia virus type 1 | Epsilonretrovirus | WEHV1        | Fish                     | AF133051                    |
| 11  | Walleye epidermal hyperplasia virus type 2 | Epsilonretrovirus | WEHV2        | Fish                     | AF133052                    |
| 12  | Atlantic salmon swim bladder sarcoma virus | Gamma-epsilon     | SSSV         | Atlantic salmon          | NC_007654                   |
| 13  | Feline leukemia virus                      | Gammaretrovirus   | FeLV         | Cat                      | NC_001940                   |
| 14  | Friend murine leukemia virus               | Gammaretrovirus   | F-MuLV       | Mouse                    | NC_001362                   |
| 15  | Mus dunni endogenous retrovirus            | Gammaretrovirus   | MDEV         | Mouse                    | AF053745                    |
| 16  | Porcine endogenous retrovirus B            | Gammaretrovirus   | PERV-B       | Pig                      | AY099324                    |
| 17  | Rhinolophus ferrumequinum retrovirus       | Gammaretrovirus   | RfRV         | Greater horseshoe bat    | JQ303225                    |
| 18  | Equine infectious anemia virus             | Lentivirus        | EIAV         | Horse                    | NC_001450                   |
| 19  | Feline immunodeficiency virus              | Lentivirus        | FIV          | Cat                      | NC_001482                   |
| 20  | Human immunodeficiency virus 1             | Lentivirus        | HIV1         | Human                    | NC_001802                   |
| 21  | Visna/Maedi virus                          | Lentivirus        | VMV          | Sheep                    | NC_001452                   |
| 22  | Amphilophus citrinellus fomy-like virus    | Spuma-like        | AciFLERV     | Fish (midas cichlid)     | Reference.1                 |
| 23  | Austrofundulus limnaeus fomy-like virus    | Spuma-like        | AliFLERV     | Fish (annual killifish ) | Reference.1                 |
| 24  | Notophthalmus viridescens                  | Spuma-like        | NviFLERV     | Amphibian (eastern newt) | Reference.1                 |
| 25  | Platyfish endogenous retrovirus            | Spuma-like        | PlatyfishEFV | Fish (platyfish)         | Reference.2                 |
| 26  | Danio rerio foamy virus                    | Spuma-like        | DrFV-1       | Zebrafish                | CABZ01054182                |
| 27  | Bovine foamy virus                         | Spumavirus        | BFV          | Cattle                   | NC_001831                   |

|    |  |              |         |                         |             |
|----|--|--------------|---------|-------------------------|-------------|
| 28 | Equine foamy virus                           | Spumavirus   | EFV     | Horse                   | NC_002201   |
| 29 | Feline foamy virus                           | Spumavirus   | FFV     | Cat                     | NC_001871   |
| 30 | Coelacanth endogenous foamy-like virus       | Spumavirus   | CoeEFV  | Fish (coelacanth)       | Reference.3 |
| 31 | Brown greater galago prosimian foamy virus   | Spumavirus   | PSFVgal | Greater galago          | KM233624    |
| 32 | White-tufted-ear marmoset simian foamy virus | Spumavirus   | SFVmar  | Common marmoset         | GU356395    |
| 33 | Squirrel monkey simian foamy virus           | Spumavirus   | SFVsqu  | Squirrel monkey         | GU356394    |
| 34 | Orangutan Simian foamy virus                 | Spumavirus   | SFVora  | Pongo pygmaeus pygmaeus | AJ544579    |
| 35 | Macaque simian foamy virus                   | Spumavirus   | SFVmac  | Macaque                 | NC_010819   |
| 36 | African green monkey simian foamy virus      | Spumavirus   | SFVagm  | African green monkey    | NC_010820   |
| 37 | Western chimpanzee simian foamy virus        | Spumavirus   | SFVcpz  | Western chimpanzee      | NC_001364   |
| 38 | Western lowland gorilla simian foamy virus   | Spumavirus   | SFVgor  | Western lowland gorilla | NC_039029   |
| 39 | Spider monkey simian foamy virus             | Spumavirus   | SFVspm  | Spider monkey           | NC_039027   |
| 40 | Sloth endogenous virus                       | Spumavirus   | SloEFV  | Sloth                   | Reference.4 |
| 41 | Snakehead retrovirus                         | Unclassified | SnRV    | Fish (snakehead fish)   | NC_001724   |

---

**Table S4. The important ERVs-Spuma-Spu sequences used for constructing consensus genome.**

| <b>ERV name</b>  | <b>Contig number</b> | <b>Contig size (bp)</b> | <b>Location (start-end)</b> | <b>Genomic region present</b> |
|------------------|----------------------|-------------------------|-----------------------------|-------------------------------|
| ERV-Spuma.1-Spu  | QEPC01000046.1       | 2,261,982               | 2082874-2085549             | pol                           |
| ERV-Spuma.3-Spu  | QEPC01000841.1       | 6,575,273               | 6473051-6474912             | 5'LTR-gag-pol-env-3'LTR       |
| ERV-Spuma.7-Spu  | QEPC01001126.1       | 3,925,424               | 1140946-1146409             | 5'LTR-gag-pol                 |
| ERV-Spuma.9-Spu  | QEPC01001501.1       | 2,970,112               | 587272-591846               | pol-env                       |
| ERV-Spuma.10-Spu | QEPC01001533.1       | 17,747,792              | 7844865-7846050             | env                           |
| ERV-Spuma.12-Spu | QEPC01001717.1       | 8,015,171               | 3767132-3769195             | env                           |
| ERV-Spuma.18-Spu | QEPC01002186.1       | 2,508,909               | 2432408-2433392             | env                           |
| ERV-Spuma.20-Spu | QEPC01002219.1       | 948,854                 | 864881-870322               | 5'LTR-gag-pol                 |
| ERV-Spuma.23-Spu | QEPC01003194.1       | 8,830,396               | 6038514-6048628             | 5'LTR-gag-pol-env-3'LTR       |
| ERV-Spuma.24-Spu | QEPC01003216.1       | 2,422,517               | 542585-543835               | env                           |
| ERV-Spuma.25-Spu | QEPC01003596.1       | 1,075,095               | 148116-150161               | env                           |
| ERV-Spuma.26-Spu | QEPC01003632.1       | 3,332,090               | 1919183-1920342             | 5'LTR-gag                     |
| ERV-Spuma.28-Spu | QEPC01003743.1       | 7,791,302               | 4870735-4875331             | 5'LTR-gag-pol                 |
| ERV-Spuma.31-Spu | QEPC01004481.1       | 1,713,101               | 1529186-1534239             | pol-env                       |
| ERV-Spuma.35-Spu | QEPC01004609.1       | 304,660                 | 79998-82854                 | gag-pol                       |
| ERV-Spuma.36-Spu | QEPC01004754.1       | 998,333                 | 809887-813532               | 5'LTR-gag-pol                 |
| ERV-Spuma.40-Spu | QEPC01004976.1       | 1,923,601               | 702186-707162               | 5'LTR-gag-pol                 |
| ERV-Spuma.46-Spu | QEPC01006337.1       | 853,782                 | 263997-266031               | 5'LTR-gag                     |
| ERV-Spuma.49-Spu | QEPC01006862.1       | 2,177,237               | 1552398-1555851             | gag-pol                       |
| ERV-Spuma.57-Spu | QEPC01007467.1       | 4,402,635               | 2528830-2534236             | 5'LTR-gag-pol                 |
| ERV-Spuma.59-Spu | QEPC01007499.1       | 9,879,472               | 4175270-4179532             | pol-env                       |
| ERV-Spuma.62-Spu | QEPC01007812.1       | 4,155,962               | 3366177-3370732             | pol-env                       |
| ERV-Spuma.63-Spu | QEPC01007842.1       | 1,455,502               | 628428-630375               | 5'LTR-gag                     |
| ERV-Spuma.65-Spu | QEPC01008061.1       | 4,176,174               | 879218-881207               | 5'LTR-gag                     |
| ERV-Spuma.74-Spu | QEPC01009046.1       | 1,866,647               | 2304064-2305587             | pol                           |
| ERV-Spuma.79-Spu | QEPC01010218.1       | 1,935,863               | 720863-722618               | env                           |
| ERV-Spuma.90-Spu | QEPC01012034.1       | 449,693                 | 259221-263659               | gag-pol                       |

|                   |                |           |                 |               |
|-------------------|----------------|-----------|-----------------|---------------|
| ERV-Spuma.94-Spu  | QEPC01012163.1 | 1,035,380 | 925517-929611   | 5'LTR-gag-pol |
| ERV-Spuma.96-Spu  | QEPC01012627.1 | 76,116    | 40790-45921     | 5'LTR-gag-pol |
| ERV-Spuma.100-Spu | QEPC01013045.1 | 5,599,944 | 3726992-3732001 | 5'LTR-gag-pol |
| ERV-Spuma.108-Spu | QEPC01014179.1 | 180,213   | 153493-156634   | env           |

---

**Table S5. Dating of ERVs-Spuma-Spu insertion based on LTR-LTR divergence.**

| Contig number  | Divergence | Integration time (MYA) |
|----------------|------------|------------------------|
| QEPC01003194.1 | 0.0266     | 17.50                  |
| QEPC01006022.1 | 0.0292     | 19.22                  |
| QEPC01009045.1 | 0.0353     | 23.20                  |
| QEPC01012853.1 | 0.0539     | 35.47                  |
| QEPC01013045.1 | 0.0020     | 1.30                   |

**Table S6. The matching contigs identified in *Sphenodon punctatus* genome.**

| Contig number  | Contig size (bp) | The number of hits* | ERV name   |
|----------------|------------------|---------------------|--|
| QEPC01000046.1 | 2,261,982        | 1                   | ERV-Spuma.1-Spu  |
| QEPC01000755.1 | 2,726,839        | 1                   | ERV-Spuma.2-Spu  |
| QEPC01000841.1 | 6,575,273        | 2                   | ERV-Spuma.3-Spu<br>ERV-Spuma.4-Spu   |
| QEPC01000971.1 | 267,574          | 1                   | ERV-Spuma.5-Spu  |
| QEPC01000978.1 | 4,485,343        | 1                   | ERV-Spuma.6-Spu  |
| QEPC01001126.1 | 3,925,424        | 1                   | ERV-Spuma.7-Spu  |
| QEPC01001380.1 | 1,282,949        | 1                   | ERV-Spuma.8-Spu  |
| QEPC01001501.1 | 2,970,112        | 1                   | ERV-Spuma.9-Spu  |
| QEPC01001533.1 | 17,747,792       | 2                   | ERV-Spuma.10-Spu<br>ERV-Spuma.11-Spu   |
| QEPC01001717.1 | 8,015,171        | 4                   | ERV-Spuma.12-Spu<br>ERV-Spuma.13-Spu<br>ERV-Spuma.14-Spu<br>ERV-Spuma.15-Spu |
| QEPC01002018.1 | 3,274,960        | 1                   | ERV-Spuma.16-Spu   |
| QEPC01002047.1 | 3,931,782        | 1                   | ERV-Spuma.17-Spu   |
| QEPC01002186.1 | 2,508,909        | 1                   | ERV-Spuma.18-Spu   |
| QEPC01002218.1 | 1,036,309        | 1                   | ERV-Spuma.19-Spu   |
| QEPC01002219.1 | 948,854          | 1                   | ERV-Spuma.20-Spu   |
| QEPC01002862.1 | 2,112,741        | 2                   | ERV-Spuma.21-Spu<br>ERV-Spuma.22-Spu   |
| QEPC01003194.1 | 8,830,396        | 1                   | ERV-Spuma.23-Spu   |
| QEPC01003216.1 | 2,422,517        | 1                   | ERV-Spuma.24-Spu   |
| QEPC01003596.1 | 1,075,095        | 1                   | ERV-Spuma.25-Spu   |
| QEPC01003632.1 | 3,332,090        | 1                   | ERV-Spuma.26-Spu   |
| QEPC01003681.1 | 2,869,604        | 1                   | ERV-Spuma.27-Spu   |
| QEPC01003743.1 | 7,791,302        | 2                   | ERV-Spuma.28-Spu<br>ERV-Spuma.29-Spu   |
| QEPC01004016.1 | 1,026,366        | 1                   | ERV-Spuma.30-Spu   |
| QEPC01004481.1 | 1,713,101        | 1                   | ERV-Spuma.31-Spu   |
| QEPC01004587.1 | 2,591,693        | 2                   | ERV-Spuma.32-Spu   |

|                |            |   |                  |
|----------------|------------|---|------------------|
| QEPC01004588.1 | 2,510,880  | 1 | ERV-Spuma.33-Spu |
| QEPC01004609.1 | 304,660    | 1 | ERV-Spuma.34-Spu |
| QEPC01004754.1 | 998,333    | 2 | ERV-Spuma.35-Spu |
|                |            |   | ERV-Spuma.36-Spu |
|                |            |   | ERV-Spuma.37-Spu |
| QEPC01004817.1 | 5,636,056  | 2 | ERV-Spuma.38-Spu |
|                |            |   | ERV-Spuma.39-Spu |
| QEPC01004976.1 | 1,923,601  | 1 | ERV-Spuma.40-Spu |
| QEPC01005318.1 | 1,812,687  | 1 | ERV-Spuma.41-Spu |
| QEPC01005364.1 | 927,510    | 1 | ERV-Spuma.42-Spu |
| QEPC01005564.1 | 1,094,855  | 1 | ERV-Spuma.43-Spu |
| QEPC01005592.1 | 430,143    | 1 | ERV-Spuma.44-Spu |
| QEPC01006022.1 | 1,470,196  | 1 | ERV-Spuma.45-Spu |
| QEPC01006337.1 | 853,782    | 1 | ERV-Spuma.46-Spu |
| QEPC01006337.1 | 853,782    | 1 | ERV-Spuma.47-Spu |
| QEPC01006515.1 | 867,373    | 1 | ERV-Spuma.48-Spu |
| QEPC01006862.1 | 2,177,237  | 2 | ERV-Spuma.49-Spu |
|                |            |   | ERV-Spuma.50-Spu |
| QEPC01007003.1 | 1,535,892  | 1 | ERV-Spuma.51-Spu |
| QEPC01007237.1 | 6,544,613  | 2 | ERV-Spuma.52-Spu |
|                |            |   | ERV-Spuma.53-Spu |
| QEPC01007433.1 | 3,616,258  | 2 | ERV-Spuma.54-Spu |
|                |            |   | ERV-Spuma.55-Spu |
| QEPC01007453.1 | 1,560,635  | 1 | ERV-Spuma.56-Spu |
| QEPC01007467.1 | 4,402,635  | 2 | ERV-Spuma.57-Spu |
|                |            |   | ERV-Spuma.58-Spu |
| QEPC01007499.1 | 9,879,472  | 1 | ERV-Spuma.59-Spu |
| QEPC01007621.1 | 1,817,577  | 1 | ERV-Spuma.60-Spu |
| QEPC01007748.1 | 3,321,910  | 1 | ERV-Spuma.61-Spu |
| QEPC01007812.1 | 4,155,962  | 1 | ERV-Spuma.62-Spu |
| QEPC01007842.1 | 1,455,502  | 1 | ERV-Spuma.63-Spu |
| QEPC01008041.1 | 6,575,273  | 1 | ERV-Spuma.64-Spu |
| QEPC01008061.1 | 4,176,174  | 1 | ERV-Spuma.65-Spu |
| QEPC01008071.1 | 3,181,871  | 1 | ERV-Spuma.66-Spu |
| QEPC01008173.1 | 12,535,514 | 1 | ERV-Spuma.67-Spu |
| QEPC01008309.1 | 13,133,864 | 1 | ERV-Spuma.68-Spu |
| QEPC01008451.1 | 3,157,202  | 2 | ERV-Spuma.69-Spu |
|                |            |   | ERV-Spuma.70-Spu |
| QEPC01008452.1 | 4,018,016  | 1 | ERV-Spuma.71-Spu |
| QEPC01008871.1 | 1,592,225  | 1 | ERV-Spuma.72-Spu |
| QEPC01009045.1 | 1,866,647  | 1 | ERV-Spuma.73-Spu |
| QEPC01009046.1 | 1,866,647  | 1 | ERV-Spuma.74-Spu |
| QEPC01009539.1 | 190,299    | 1 | ERV-Spuma.75-Spu |
| QEPC01009586.1 | 505,356    | 1 | ERV-Spuma.76-Spu |



|                |            |   |                   |
|----------------|------------|---|-------------------|
| QEPC01009783.1 | 1,171,668  | 2 | ERV-Spuma.77-Spu  |
|                |            |   | ERV-Spuma.78-Spu  |
| QEPC01010218.1 | 1,935,863  | 1 | ERV-Spuma.79-Spu  |
| QEPC01010231.1 | 2,180,379  | 2 | ERV-Spuma.80-Spu  |
|                |            |   | ERV-Spuma.81-Spu  |
| QEPC01010332.1 | 3,049,825  | 1 | ERV-Spuma.82-Spu  |
| QEPC01010651.1 | 2,503,641  | 1 | ERV-Spuma.83-Spu  |
| QEPC01010723.1 | 1,079,134  | 1 | ERV-Spuma.84-Spu  |
| QEPC01011263.1 | 1,245,951  | 1 | ERV-Spuma.85-Spu  |
| QEPC01011712.1 | 4,733,220  | 1 | ERV-Spuma.86-Spu  |
| QEPC01011882.1 | 3,029,500  | 2 | ERV-Spuma.87-Spu  |
|                |            |   | ERV-Spuma.88-Spu  |
| QEPC01011901.1 | 29,987,930 | 1 | ERV-Spuma.89-Spu  |
| QEPC01012034.1 | 449,693    | 2 | ERV-Spuma.90-Spu  |
|                |            |   | ERV-Spuma.91-Spu  |
| QEPC01012093.1 | 582,806    | 2 | ERV-Spuma.92-Spu  |
|                |            |   | ERV-Spuma.93-Spu  |
| QEPC01012163.1 | 1,035,380  | 1 | ERV-Spuma.94-Spu  |
| QEPC01012515.1 | 1,668,146  | 1 | ERV-Spuma.95-Spu  |
| QEPC01012627.1 | 76,116     | 1 | ERV-Spuma.96-Spu  |
| QEPC01012853.1 | 722,533    | 1 | ERV-Spuma.97-Spu  |
| QEPC01012962.1 | 1,874,977  | 2 | ERV-Spuma.98-Spu  |
|                |            |   | ERV-Spuma.99-Spu  |
| QEPC01013045.1 | 5,599,944  | 1 | ERV-Spuma.100-Spu |
| QEPC01013253.1 | 1,321,719  | 2 | ERV-Spuma.101-Spu |
|                |            |   | ERV-Spuma.102-Spu |
| QEPC01013568.1 | 4,769,775  | 1 | ERV-Spuma.103-Spu |
| QEPC01013677.1 | 2,261,225  | 1 | ERV-Spuma.104-Spu |
| QEPC01013962.1 | 978,021    | 1 | ERV-Spuma.105-Spu |
| QEPC01013985.1 | 1,532,209  | 1 | ERV-Spuma.106-Spu |
| QEPC01014049.1 | 917,795    | 1 | ERV-Spuma.107-Spu |
| QEPC01014179.1 | 180,213    | 1 | ERV-Spuma.108-Spu |
| QEPC01014350.1 | 2,578,219  | 1 | ERV-Spuma.109-Spu |
| QEPC01014606.1 | 3,568,174  | 2 | ERV-Spuma.110-Spu |
|                |            |   | ERV-Spuma.111-Spu |
| QEPC01014636.1 | 2,612,062  | 1 | ERV-Spuma.112-Spu |
| QEPC01014787.1 | 1,793,453  | 1 | ERV-Spuma.113-Spu |
| QEPC01015038.1 | 2,158,932  | 1 | ERV-Spuma.114-Spu |
| QEPC01015112.1 | 1,559,665  | 1 | ERV-Spuma.115-Spu |
| QEPC01015139.1 | 3,320,937  | 1 | ERV-Spuma.116-Spu |
| QEPC01015182.1 | 2,270,417  | 1 | ERV-Spuma.117-Spu |
| QEPC01015561.1 | 3,676,817  | 1 | ERV-Spuma.118-Spu |

\* The hits within 30kb are regarded as one hit.

## References

1. Aiewsakun P, Katzourakis A. 2017. Marine origin of retroviruses in the early Palaeozoic Era. *Nat Commun* 8:13954.
2. Katzourakis A, Gifford RJ, Tristem M, Gilbert MT, Pybus OG. 2009. Macroevolution of complex retroviruses. *Science* 325:1512.
3. Han GZ, Worobey M. 2012. An endogenous foamy-like viral element in the coelacanth genome. *PLoS Pathog* 8:e1002790.
4. Scharl M, Walter RB, Shen Y, Garcia T, Catchen J, Amores A, Braasch I, Chalopin D, Volff JN, Lesch KP, et al. 2013. The genome of the platyfish, *Xiphophorus maculatus*, provides insights into evolutionary adaptation and several complex traits. *Nat Genet* 45:567-572.