Virology 530 (2019) 27–31

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

Discovery of endogenous retroviruses with mammalian envelopes in avian genomes uncovers long-term bird-mammal interaction

Mingyue Chen, Jie Cui*

CAS Key Laboratory of Special Pathogens and Biosafety, Center for Emerging Infectious Diseases, Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan 430071, China

ARTICLE INFO

Keywords:
Birds
Endogenous retroviruses
Recombination
Cross-order transmission
Envelope protein

ABSTRACT

Endogenous retroviruses (ERVs) arise from the infection and integration of past retroviruses into animal hosts. We performed large-scale genomic mining of 101 avian genomes for discovery of ERVs having none-avian origin and investigated the cross-species transmission events. Phylogenetic analysis of the reverse transcriptase (RT) of polymerase gene (*pol)* and the transmembrane subunit (TM) of the envelope gene (*env*) supported that avian ERVs with a mammalian *env* gene existed in at least 15 avian species and can be divided into two major groups: Group-1 were of recombinant ERVs with an alpha-like *pol* gene and a gamma-like *env* gene, and Group-2 included ERVs with both gamma-like *pol* and *env* genes. Group-1 represented the avian alpharetroviral/mammalian gammaretroviral recombinant while Group-2 documented viral jump from mammals to birds. Molecular dating analysis suggested that Group-1 ERVs had integrated into avian genomes continuously, until recent past. We have expanded the knowledge of ERVs with cross-order transmission.

1. Introduction

Endogenous retroviruses (ERVs) represent remnants of past retroviral infection (Boeke and Stoye, 1997). ERVs are commonly distributed in vertebrate genomes, and thus, can be used to investigate the long-term evolution of retroviruses, including cross-species transmission (CST) events (Zhuo and Feschotte, 2015). Birds have diversified classes of ERVs including alpha-, beta- and gamma-ERVs in their genomes (Cui et al., 2014). There has been limited documentation of the transmission of ERVs between birds and other vertebrates. A well-studied example is the reticuloendotheliosis virus (REV), which originated from mammalian retroviruses and now circulates in birds (Niewiadomska and Gifford, 2013). Another study revealed that a recombinant ERV (named TgERV-F) in the genome of the zebra finch (*Taeniopygia guttata*) had a mammalian gammaretroviral *env* gene and combines features of both avian and mammalian retroviruses (Henzy et al., 2014).

A typical retrovirus genome consists of at least three genes, namely *gag*, *pol*, and *env*. *pol* is the most conserved of the three genes, which encodes the reverse transcriptase (RT) and integrase (IN) proteins. Based on phylogenetic analysis of the RT region, ERVs can be classified into three classes: class I ERVs include gamma- and epsilonretroviruses; class II ERVs include beta-, delta-, and alpharetroviruses and lentiviruses; and class III includes spumaretroviruses (Henzy et al., 2014). *env* is the most divergent of the three retroviral genes and encodes the surface subunit (SU) and the transmembrane subunit (TM), which together form the viral envelope. SU is highly variable and involved in receptor recognition, and TM is highly conserved and directly responsible for cell membrane fusion and virus entry (Benit et al., 2001; Henzy et al., 2014). Based on the sequence of TM region, retroviral envelope glycoproteins can be divided into two groups: the covalent type (typical of alpha-, gamma-, and delta-retroviruses) with a triple-cysteine motif in the TM loop region and the noncovalent type (typical of betaretroviruses and lentiviruses) that lacks the third cysteine (Henzy and Coffin, 2013; Henzy et al., 2014). The viral envelope is thus a key determinant of host range of retroviruses. Therefore, recombinant that acquired a heterologous *env* gene from a distantly related retrovirus could, in theory, facilitate viral entry into a new host (Henzy et al., 2014).

TgERV-F is the only known class II/gammaretroviral recombinant ERV in the avian genome and has only been identified in the genome of the zebra finch (Henzy et al., 2014). In this study, we demonstrated that there were more class II/gammaretroviral recombinants involved with different avian species, and some recombinants infiltrated avian genomes much earlier than TgERV-F. Moreover, we identified a group of ERVs that were gamma-like and may have directly transmitted from mammals to birds.

---

* Corresponding author.
E-mail address: jiecui@wh.iov.cn (J. Cui).

https://doi.org/10.1016/j.virol.2019.02.005
Received 21 November 2018; Received in revised form 7 February 2019; Accepted 7 February 2019

2. Materials and methods

2.1. Initial detection of CST events involving avian ERVs

We first used tBLASTn (Altschul et al., 1990) and performed an in silico genome-scale mining of all avian genomes available in the National Center of Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) whole genome shotgun (WGS) database as of April 2018 (101 avian species, in total). We used SU and TM subunit amino acid sequences (from 71 representative ERVs and infectious retroviruses, including alpha-, beta-, gamma-, delta-, epsilon-, spumaretroviruses, and lentiviruses) as queries. A 65% similarity over 80% of the sequence region and an E-value of 0.001 were used as filter criteria to exclude nonspecific hits. To detect ERVs with none-avian origin, hits to queries of avian RT and TM sequences (accession numbers: NC_015116.1, NC_001407.1, AB522906.1, AJ623291.1) were excluded.

2.2. Identification of putative full-length ERVs

To conduct an in-depth analysis of ERVs identified by genome mining, we focused only on full-length ERVs. First, we extracted the sequences of the hits, along with 30 kb flanking sequences. Then full-length ERVs were identified using both LTR_FINDER (Xu and Wang, 2007) and LTRharvest (Ellinghaus et al., 2008), and only ERVs predicted by both software programs were included in subsequent analysis. The Conserved Domain database (https://www.ncbi.nlm.nih.gov/cdd/) was used to identify ERVs with RT domains. Only ERVs with both RT and TM regions were analyzed for recombination events.

2.3. Phylogenetic analyses

To determine the phylogenetic relationship between the identified ERVs, we performed phylogenetic analyses of their TM and RT regions. Multiple alignments of the nucleotide sequences of the TM and RT regions were conducted using the MAFFT program (Katoh and Standley, 2013) in MEGA 7.0 (Kumar et al., 2016). Phylogenetic analyses were performed in RAxML (Stamatakis, 2014) with the GTRCAT substitution model, incorporating 100 bootstrap replicates to determine the robustness of the result. All alignments were provided upon request.

2.4. Molecular dating of full-length ERVs

The ERV integration time based on LTR divergence was calculated using the equation $T = (D/R)/2$, where T is the invasion time (million years, MY), D is the number of nucleotide differences per site between the two LTRs, and R is the genomic substitution rate (nucleotide substitutions per site per year). For the value of R, we applied previously estimated, neutral substitution rates for the bird genome of $2 \times 10–9$ and $3.9 \times 10–9$ substitutions per site per year (Axelsson et al., 2004; Henzy et al., 2014).

3. Results

3.1. ERVs closely related to Rhinolophus ferrumequinum retrovirus (RfRV) or TgERV-F are distributed among different bird genomes

To detect ERVs with none-avian origin, we performed a large scale of genomic mining of SU and TM subunit sequences within all 101 available avian genomes in GenBank. We excluded hits to queries from avian retroviruses through an in silico similarity search. Then, the most significant hits were obtained from two TM subunit amino acid sequences of two ERVs, namely the RfRV and the TgERV-F. The RfRV is an endogenous gammaretrovirus discovered in bats (Cui et al., 2012), and TgERV-F is an avian recombinant ERV from the zebra finch that acquired a mammalian gammaretroviral *env* sequence (Henzy et al., 2014). Further examination revealed that the hits with TM regions similar to RfRV and TgERV-F were distributed among different bird species, mostly those belonging to the Neognathae superorder such as the turkey (*Meleagris gallopavo*, Galliformes), the zebra finch (*Taeniopygia guttata*, Passeriformes), the carmine bee-eater (*Merops nubicus*, Coraciiformes), the Anna's hummingbird (*Calypte anna*, Trochiliformes), and the peach-faced lovebird (*Agapornis roseicollis*, Psittaciformes). The hits were also found in the brown kiwi (*Apteryx australis*) belonging to the superorder Palaeognathae, which diverged from Neognathae during the early Cretaceous period (110 million years ago) (Yonezawa et al., 2017).

3.2. Phylogenetic analysis of full-length ERVs

To classify avian ERVs included in these hits, we screened full-length ERVs with the RT region of *pol*, and the results generated 91 full-length ERVs (Supplementary Table 1, Supplementary Data 1). Each avian species harbored a relatively small number ($1-9$) of these ERVs, except for the zebra finch, which possessed as many as 55 copies (Supplementary Table 1). This implied that ERVs with the characteristic of a mammalian gammaretroviral *env* had greatly expanded in the zebra finch, generating numerous insertions.

The RT phylogenetic tree suggested that 91 full-length ERVs could be divided into two groups: Group-1 (containing 86 ERVs) that clustered with alpharetroviruses and Group-2 (containing the remaining five ERVs) that clustered with gammaretroviruses (Fig. 1). The RT catalytic site of 84 Group-1 ERVs consisted of the amino acid sequence “YMDD”, which is typical of class II ERVs (Henzy et al., 2014), and the RT catalytic site sequence of the remaining two Group-1 ERVs (accession numbers: LXGM01000049.1 and NDXB01014090.1) was “YMDN”; a similar but mutated version of class II ERVs. Surprisingly, the RT catalytic-site sequence of Group-2 ERVs (accession numbers: LK064666.1, LK071509.1, KN265775.1, KB398124.1, and MVNZ02000048.1) was “YVDD”, which has been shown in the Moloney murine leukemia virus (MuLV) to play a role in the fidelity of DNA synthesis (Kaushik et al., 2000).

Gammaretroviruses are present in vertebrates, mainly in mammals (Hayward et al., 2015). Representative gammaretroviruses in mammals and also birds were used as reference sequences in Fig. 2. The phylogenetic tree of TM region showed that all Groups-1 and -2 ERVs clustered with gammaretroviruses, and RfRV was their closest relative (Fig. 2). Our online BLAST searches also showed that RfRV was the most closely related retrovirus to all Groups-1 and -2 ERVs (59–71% sequence identity). So Groups-1 and -2 ERVs may have mammalian origin. Group-1 (except for the three ERVs, for which the first codon of “CX6CC” motif did not code for cysteine) and -2 ERVs had a covalent envelope structure with a “CX6CC” motif in the TM region. Groups-1 and -2 ERVs also had a conserved immunosuppressive domain (ISD), which is typical of gamma- and alpharetroviral but not betaretroviral, TM proteins.

The TM phylogenetic tree indicated that Groups-1 and -2 ERVs had a heterologous mammalian gammaretroviral *env* sequence (Fig. 2). The similar structure (“CX6CC” motif and ISD) and the close relationship of Groups-1 and -2 ERVs suggested that the *env* gene of the two groups may originate from a common progenitor. For Group-2, the consistency of the RT and TM phylogenetic trees (Figs. 1 and 2) indicated that all five members were gamma-ERVs and may have been acquired from mammals through CST events. But for Group-1, the inconsistency of the RT and TM phylogenetic trees suggested that Group-1 ERVs were the avian alpharetroviral/mammalian gammaretroviral recombinant. The recombinations were also supported using Recombination Detection Program (RDP) (Martin et al., 2015). Group-1 included TgERV-F found in the zebra finch (Henzy et al., 2014), and ERVs in this group were also identified in other bird species belonging to both Palaeognathae (Apterygiformes) and Neognathae (Passeriformes, Psittaciformes, Falconiformes, Apodiformes, and Strigiformes) superorders (Supplementary Table 1). Since ERVs without long terminal repeats (LTRs) were not included in our analysis, the genomic distributions of Groups-1 and -2 in birds are probably underestimated. Table 1

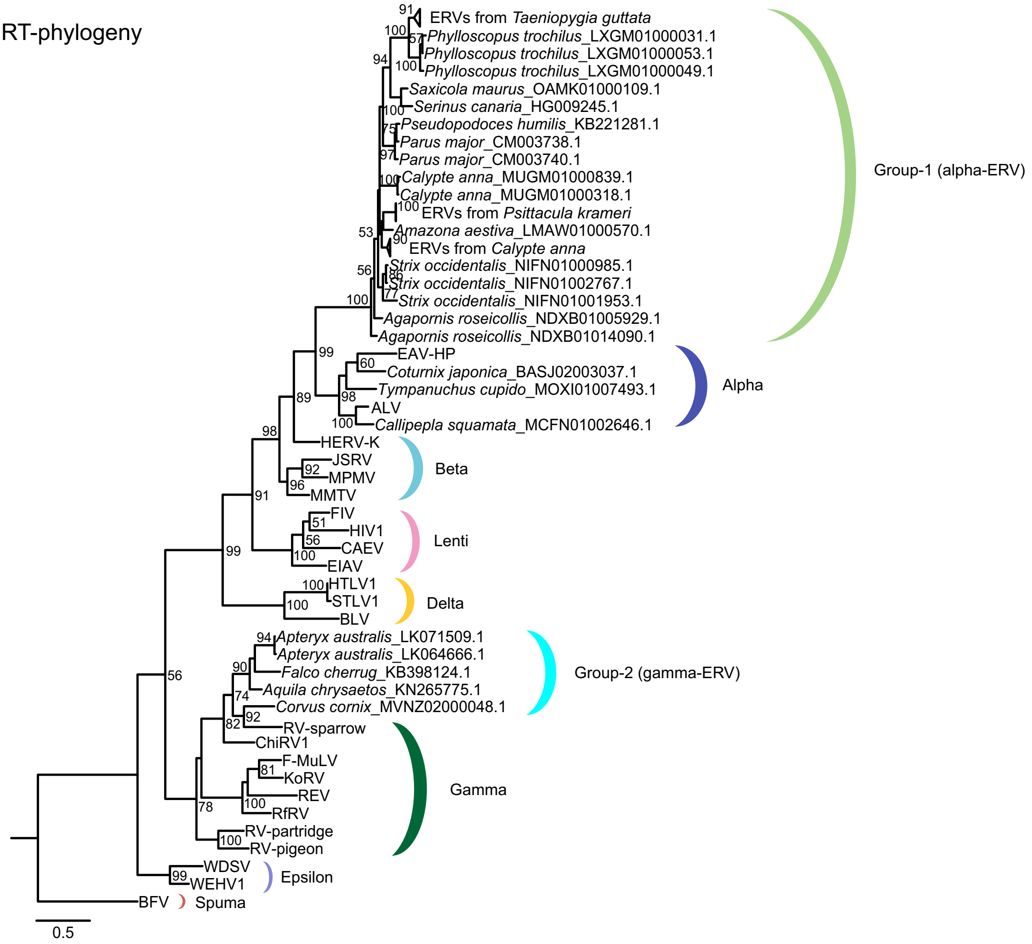

Fig. 1. Phylogenetic tree of RT region constructed using nucleotide sequences of Groups-1 and −2 ERVs identified in this study and other representative retroviruses (alpha-, beta-, gamma-, delta-, epsilon-, spumaretroviruses, and lentiviruses). Bootstrap values < 50% (obtained from 100 replicates) are not shown in the phylogenetic tree. The tree was rooted using Bovine Foamy virus (BFV). Clades with closely similar ERVs from *Taeniopygia guttata*, *Calypte anna*, and *Psittacula krameri* were collapsed in the tree. TgERV-F was included in ERVs from *Taeniopygia guttata*. Alpha, Beta, Gamma, Delta, Epsilon, Spuma, and Lenti represent the representative alpha-, beta-, gamma-, delta-, epsilon-, spumaretroviruses, and lentiviruses, respectively. ALV, avian leukemia virus; EAV-HP, avian endogenous retrovirus EAV-HP; ChiRV1, *Gallus gallus* MLV-related endogenous retrovirus; RV-sparrow, sparrow endogenous retrovirus; RV-partridge, partridge endogenous retrovirus; RV-pigeon, pigeon endogenous retrovirus; HERV-K, human endogenous retrovirus K113; JSRV, jaagsiekte sheep retrovirus; MPMV, Mason-Pfizer monkey virus; MMTV, murine mammary tumor virus; FIV, feline immunodeficiency virus; HIV1, human immunodeficiency virus 1; CAEV, caprine arthritis encephalitis virus; EIAV, equine infectious anemia virus; STLV1, simian T-lymphotropic virus 1; HTLV1, human T-lymphotropic virus 1; BLV, bovine leukemia virus; WEHV1, walleye epidermal hyperplasia virus type 1; WDSV, walleye dermal sarcoma virus; F-MuLV, Friend murine leukemia virus; KoRV, koala retrovirus; RfRV, *Rhinolophus ferrumequinum* retrovirus; REV, reticuloendotheliosis virus. *Coturnix japonica*_BASJ02003037.1, *Tympanuchus cupido_MOXI01007493.1*, and *Callipepla squamata_MCFN01002646.1* are the representative avian alpha-ERVs from *Coturnix japonica*, *Tympanuchus cupido*, and *Callipepla squamata*, respectively.

3.3. Molecular dating analysis

To gain further insight into the evolution of these avian ERVs, we estimated their integration time, which is based on the LTR divergence (Johnson and Coffin, 1999). The oldest ERV in Group-1 (accession number: HG009245.1) dated back to 7.9–15.5 million years ago (MYA; Supplementary Table 1), much earlier than the oldest TgERV-F (2.1–4.2 MYA) (Henzy et al., 2014). And the oldest ERV in Group-2 (accession number: LK064666.1) dated back to 7.1–13.8 MYA. The other four ERVs in Group-2 (accession numbers: LK071509.1, KN265775.1, KB398124.1, and MVNZ02000048.1) infiltrated bird genomes 1.4–2.8, 2.3–4.5, 1.8–3.5, and 1.8–3.5 MYA (Supplementary Table 1). Continuous integration times for Group-1 (e.g. 6.5–12.8, 5.8–11.3, 4.2–8.3, 3.3–6.5, 2.2–4.3, 1.5–3.0, and 0.4–0.8 MYA) are listed in Supplementary Table 1. Among these, 19 ERVs even have identical LTR pairs, indicating recent viral integration. These results implied that Group-1 ERVs have integrated into avian genomes continuously, until recent past.

3.4. Evolutionary history of Groups-1 and −2 ERVs

The similar TM structure and close phylogenetic relationship between the TM regions of Groups-1 and −2, as well as the proximity of the integration times of the oldest ERVs in Groups-1 and −2, suggested that the infection of Group-2 (gamma-ERVs) through cross-order transmission from mammals may have resulted in the acquisition of a mammalian *env* by alpharetroviruses, leading to the appearance of Group-1 (the recombinant alpha-ERVs). We also found that only two ERVs in Group-1 (accession numbers: CM003738.1 and KB221281.1) may be vertically transmitted from a common ancestor, as verified by their flanking sequences (i.e. the host sequences flanked ERVs should be identical for vertical transmission).

4. Discussion

In this study, we identified two groups of avian ERVs (Groups-1 and −2) which have acquired the mammalian gammaretroviral *env* gene. Group-2 ERVs were gamma-like and may have originated from mammalian retroviruses, while Group-1 ERVs may be avian alpharetroviral/mammalian gammaretroviral recombinant. The recombination probably occurred after the infection of birds as alpharetroviruses have been so far described exclusively in birds. Our results showed that RfRV was the most closely related retrovirus to Groups-1 and −2 ERVs. RfRV is an endogenous gammaretrovirus isolated from bats, a species that has been shown to be capable of carrying and transmitting emerging and re-emerging viruses (Cui et al., 2012). Previous research has even implied that mammalian gammaretroviruses may have originated from bats (Cui et al., 2012). Thus, our observation raises the intriguing possibility that the *env* gene of Groups-1–2 may have originated from RfRV.

Until recently, the only well-documented avian retroviruses with a mammalian gammaretroviral *env* were REV and TgERV-F. REV originated as mammalian retroviruses that were accidentally introduced into avian hosts in the late 1930s, and subsequently integrated into the

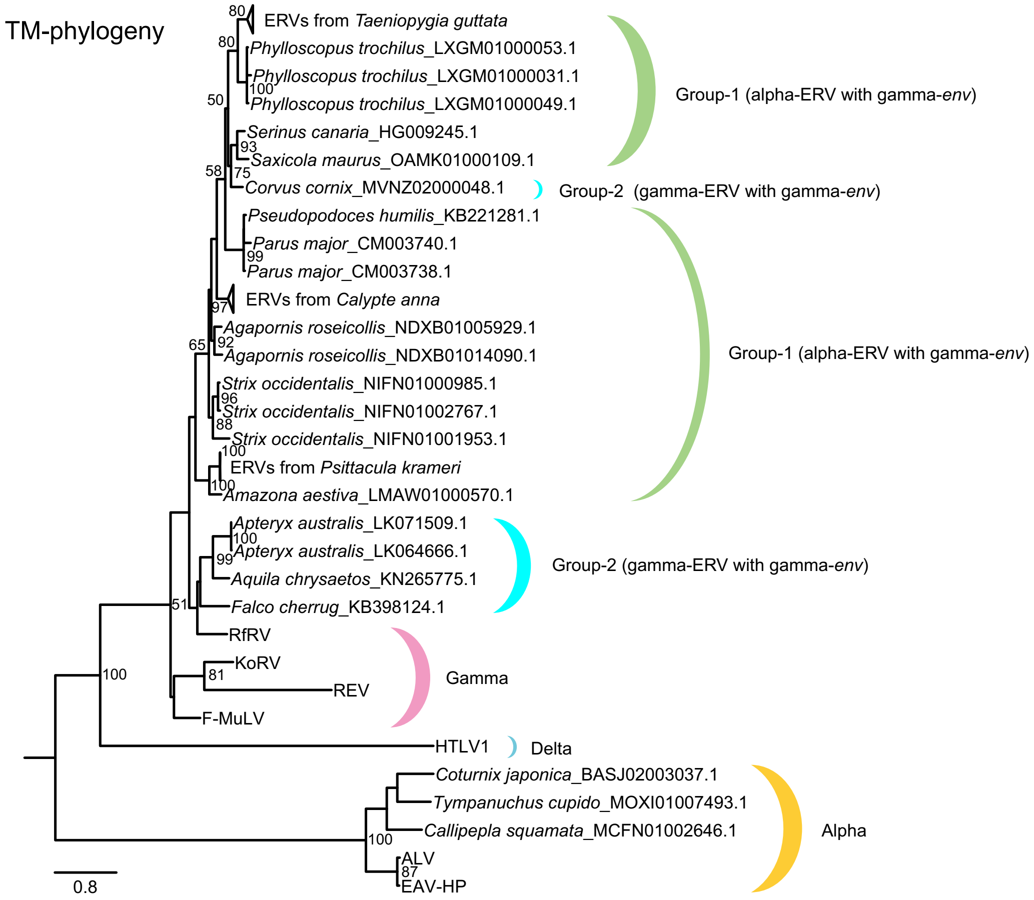

Fig. 2. Phylogenetic tree of TM region constructed using nucleotide sequences of 91 avian ERVs identified in this study and other representative alpha-, gamma-, and deltaretroviruses. Bootstrap values < 50% (obtained from 100 replicates) are not shown in the phylogenetic tree. Clades with closely similar ERVs from *Taeniopygia guttata*, *Calypte anna*, and *Psittacula krameri* were collapsed in the tree. TgERV-F was included in ERVs from *Taeniopygia guttata*. Alpha, Gamma, and Delta represent the representative alpha-, gamma-, and deltaretrovirus, respectively. ALV, avian leukemia virus; EAV-HP, avian endogenous retrovirus EAV-HP; RfRV, *Rhinolophus ferrumequinum* retrovirus; F-MuLV, Friend murine leukemia virus; KoRV, koala retrovirus; REV, reticuloendotheliosis virus. *Callipepla squamata*_MCFN01005272.1 and *Charadrius vociferus*_KL872275.1 are the representative alpha-ERVs from scaled quail (*Callipepla squamata*) and killdeer (*Charadrius vociferus*), respectively.

Table 1
The distribution of avian Groups-1 and −2 ERVs.

Species	Taxonomy	No. of ERVs	Category of ERVs
Taeniopygia guttata	Neognathae, Passeriformes, Passeroidea, Estrildidae, Estrildinae, Taeniopygia	55	Group-1
Phylloscopus trochilus	Neognathae, Passeriformes, Sylvioidea, Phylloscopidae, Phylloscopus	3	Group-1
Parus major	Neognathae, Passeriformes, Paridae, Parus	2	Group-1
Pseudopodoces humilis	Neognathae, Passeriformes, Paridae, Pseudopodoces	1	Group-1
Saxicola maurus	Neognathae, Passeriformes, Muscicapidae, Saxicola	1	Group-1
Serinus canaria	Neognathae, Passeriformes, Passeroidea, Fringillidae, Carduelinae, Serinus	1	Group-1
Psittacula krameri	Neognathae, Psittaciformes, Psittacidae, Psittacula	9	Group-1
Agapornis roseicollis	Neognathae, Psittaciformes, Psittacidae, Agapornis	2	Group-1
Amazona aestiva	Neognathae, Psittaciformes, Psittacidae, Amazona	1	Group-1
Calypte anna	Neognathae, Apodiformes, Trochilidae, Calypte	8	Group-1
Strix occidentalis	Neognathae, Strigiformes, Strigidae, Strix	3	Group-1
Apteryx australis	Palaeognathae, Apterygiformes, Apterygidae, Apteryx	2	Group- 2
Corvus cornix	Neognathae, Passeriformes, Corvoidea, Corvidae, Corvus	1	Group- 2
Aquila chrysaetos	Neognathae, Falconiformes, Accipitridae, Accipitrinae, Aquila	1	Group- 2
Falco cherrug	Neognathae, Falconiformes, Falconidae, Falco	1	Group- 2

fowlpox virus (FWPV) and gallid herpesvirus type 2 (GHV-2) genomes, generating recombinant DNA viruses that now circulate in birds (Niewiadomska and Gifford, 2013). In this study, we described the gamma-like Group-2 ERVs with a mammalian *env*, the oldest of which has integrated into the avian genome 7.1–13.8 MYA, indicating the early viral jump from mammals to birds. In addition, the only known class II/gammaretroviral recombinant in the avian genome is TgERV-F, which has only been identified in the zebra finch (Henzy et al., 2014). Our data showed that more class II/gammaretroviral recombinant ERVs (Group-1) were present in birds, including both Palaeognathae (Apterygiformes) and Neognathae (Passeriformes, Psittaciformes, Falconiformes, Apodiformes, and Strigiformes; Supplementary Table 1). Thus, we demonstrated widespread and long-term bird-mammal interaction.

The number of full-length ERVs in each avian genome was relatively small. But there are much more solitary LTRs. All 91 full-length avian ERVs are distributed in 15 avian genomes. Solitary LTRs of 15 avian genomes are at least five times the number of full-length ERVs except for *Strix occidentalis* (three full-length ERVs and three solitary LTRs). So it is important to note that our Groups-1 and −2 record in avian genomes likely represented only a fraction of the total number of CST events as we only considered full-length ERVs. And some unknown proportion of ERVs may have eventually become extinct, thereby losing their corresponding genomic records. Thus, it is likely that there are more bird-mammal interaction events that took place earlier in evolutionary time. In our study, TgERV-F was included in Group-1. TgERV-F also possesses betaretrovirus-like features of genome organization and demonstrates a unique mix of alpha-, beta-, and gammaretroviral features (Henzy et al., 2014). Thus, we cannot rule out that one or several betaretroviruses were involved in the introduction of the betaretroviral characteristic into Group-1 ERVs. Taken together, our study discovers ERVs with mammalian envelopes in bird genomes and thus uncovers long-term bird-mammal interaction.

Acknowledgment

This research was supported by National Natural Science Foundation of China (31671324) and CAS Pioneer Hundred Talents Program.

Declarations of interest

None.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.virol.2019.02.005.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Axelsson, E., Smith, N.G., Sundstrom, H., Berlin, S., Ellegren, H., 2004. Male-biased mutation rate and divergence in autosomal, z-linked and w-linked introns of chicken and Turkey. Mol. Biol. Evol. 21, 1538–1547.

Benit, L., Dessen, P., Heidmann, T., 2001. Identification, phylogeny, and evolution of retroviral elements based on their envelope genes. J. Virol. 75, 11709–11719.

Boeke, J.D., Stoye, J.P., 1997. In: Coffin, J.M., Hughes, S.H., Varmus, H.E. (Eds.), Retrotransposons, Endogenous Retroviruses, and the Evolution of Retroelements. Cold Spring Harbor (NY), Retroviruses.

Cui, J., Tachedjian, M., Wang, L., Tachedjian, G., Wang, L.F., Zhang, S., 2012. Discovery of retroviral homologs in bats: implications for the origin of mammalian gammaretroviruses. J. Virol. 86, 4288–4293.

Cui, J., Zhao, W., Huang, Z., Jarvis, E.D., Gilbert, M.T., Walker, P.J., Holmes, E.C., Zhang, G., 2014. Low frequency of paleoviral infiltration across the avian phylogeny. Genome Biol. 15, 539.

Ellinghaus, D., Kurtz, S., Willhoeft, U., 2008. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinforma. 9, 18.

Hayward, A., Cornwallis, C.K., Jern, P., 2015. Pan-vertebrate comparative genomics unmasks retrovirus macroevolution. Proc. Natl. Acad. Sci. USA 112, 464–469.

Henzy, J.E., Coffin, J.M., 2013. Betaretroviral envelope subunits are noncovalently associated and restricted to the mammalian class. J. Virol. 87, 1937–1946.

Henzy, J.E., Gifford, R.J., Johnson, W.E., Coffin, J.M., 2014. A novel recombinant retrovirus in the genomes of modern birds combines features of avian and mammalian retroviruses. J. Virol. 88, 2398–2405.

Johnson, W.E., Coffin, J.M., 1999. Constructing primate phylogenies from ancient retrovirus sequences. Proc. Natl. Acad. Sci. USA 96, 10254–10260.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Kaushik, N., Chowdhury, K., Pandey, V.N., Modak, M.J., 2000. Valine of the YVDD motif of moloney murine leukemia virus reverse transcriptase: role in the fidelity of DNA synthesis. Biochemistry 39, 5155–5165.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: detection and analysis of recombination patterns in virus genomes. Virus Evol. 1, vev003.

Niewiadomska, A.M., Gifford, R.J., 2013. The extraordinary evolutionary history of the reticuloendotheliosis viruses. PLoS Biol. 11, e1001642.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Xu, Z., Wang, H., 2007. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. 35, W265–268.

Yonezawa, T., Segawa, T., Mori, H., Campos, P.F., Hongoh, Y., Endo, H., Akiyoshi, A., Kohno, N., Nishida, S., Wu, J., Jin, H., Adachi, J., Kishino, H., Kurokawa, K., Nogi, Y., Tanabe, H., Mukoyama, H., Yoshida, K., Rasoamiaramanana, A., Yamagishi, S., Hayashi, Y., Yoshida, A., Koike, H., Akishinonomiya, F., Willerslev, E., Hasegawa, M., 2017. Phylogenomics and morphology of extinct Paleognaths reveal the origin and evolution of the ratites. Curr. Biol. 27, 68–77.

Zhuo, X., Feschotte, C., 2015. Cross-species transmission and differential fate of an endogenous retrovirus in three mammal lineages. PLoS Pathog. 11, e1005279.