# Coelacanth *SERINC2* Inhibits HIV-1 Infectivity and Is Counteracted by Envelope Glycoprotein from Foamy Virus

Pavitra Ramdas,[a] Vipin Bhardwaj,[a] Aman Singh,[a] Nagarjun Vijay,[b] ⬤ Ajit Chande[a]

[a]Molecular Virology Laboratory, Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, India
[b]Computational Evolutionary Genomics Lab, Department of Biological Sciences, Indian Institute of Science Education and Research, Bhopal, India

**ABSTRACT**   *SERINC5* restricts nef-defective HIV-1 by affecting early steps of the virus life cycle. Distantly related retroviruses with a wide host range encode virulent factors in response to challenge by *SERINC5*. However, the evolutionary origins of this antiretroviral activity, its prevalence among the paralogs, and its ability to target retroviruses remain understudied. In agreement with previous studies, we found that four human *SERINC* paralogs inhibit nef-defective HIV-1, with *SERINC*2 being an exception. Here, we demonstrate that this lack of activity in human *SERINC*2 is associated with its post-whole-genome duplication (post-WGD) divergence, as evidenced by the ability of pre-WGD orthologs from *Saccharomyces cerevisiae* and flies and a post-WGD-proximate *SERINC*2 from coelacanths to inhibit the virus. Intriguingly, Nef is unable to counter coelacanth *SERINC*2, indicating that such activity was directed toward other retroviruses found in coelacanths (like foamy viruses). However, foamy virus-derived vectors are intrinsically resistant to the action of *SERINC*2, and we show that the foamy virus envelope confers this resistance by affecting its steady-state levels. Our study highlights an ancient origin of antiretroviral activity in *SERINC*s and a hitherto-unknown interaction with a foamy virus.

**IMPORTANCE** *SERINC*5 constitutes a critical barrier to the propagation of retroviruses, as highlighted by parallel emergence of anti-SERINC5 activities among distant retroviral lineages. Therefore, understanding the origin and evolution of these host factors will provide key information about virus-host relationships that can be exploited for future drug development. Here, we show that *SERINC5*-mediated nef-defective HIV-1 infection inhibition is evolutionarily conserved. *SERINC*2 from coelacanth restricts HIV-1, and it was functionally adapted to target foamy viruses. Our findings provide insights into the evolutionary origin of antiretroviral activity in the *SERINC* gene family and uncover the role of SERINCs in shaping the long-term conflicts between retroviruses and their hosts.

**KEYWORDS**   SERINC, restriction factors, antiretroviral, whole-genome duplication, foamy viruses

Retroviruses exploit a wide host-range for their persistence, in response to which host species have continually evolved increasingly intricate antiviral defense strategies (1). As part of this ongoing arms race, while viruses have relied on the acquisition or fusion of diverse genes, the host defense mechanisms have been greatly strengthened by the functional divergence of gene copies following duplication of genes as well as whole genomes (2–6). Not just the host defense mechanism but the overall success of the vertebrate lineage is thought to have benefited immensely from the two whole-genome duplication events which have occurred over the course of evolution (2). While the study of gene and genome duplication has undoubtedly improved our understanding of how evolution works, it has also become a useful approach for obtaining important insights about the functional intricacies of restriction factors

(4–7). Tracking the evolutionary history of restriction factors based on their origin, loss, duplication, and increased rates of sequence evolution concomitant with the change in the pathogen repertoire has been used as a strategy to map the diversification of various restriction factors (8–10). The most prominent signature of this ongoing arms race is seen in the form of a strong, recurrent positive selection at functionally important residues of the restriction factor genes (11, 12). Restriction factors, being at the forefront of long-term host-virus conflict (13–15), show clear molecular signatures indicative of an arms race (11, 12). In fact, the presence of these signatures has been proposed as a hallmark of restriction factors (14, 16) and has been employed as a screening strategy to identify putative candidates with antiviral activities (4, 17).

In contrast, the antiretroviral host factors *SERINC*5 and *SERINC*3 display a comparatively uneventful evolutionary history (18, 19). This is counterintuitive, because distant retroviruses with a wide host range encode anti-*SERINC*5 virulent factors in response to challenge by *SERINC*5 (20–23). Hence, we sought to trace the evolutionary origins of the antiretroviral activity of *SERINC*5, its prevalence among the *SERINC* paralogs, and its relevance for retroviral inhibition. Our analysis to comprehend the evolutionary origins of the antiretroviral function of *SERINC*s identified an antiretroviral *SERINC*2 with a hitherto-unknown interaction with a foamy virus.

## RESULTS

**Antiretroviral activity among human *SERINC* paralogs and evolutionary trajectories.** Analysis of sequence similarity and gene structure conservation reveals that *SERINC*5 and *SERINC*4 share a recent ancestry (Fig. 1A). Similarly, *SERINC*3 and *SERINC*1 paralogs are most similar to each other (~60% identity). Despite having the lowest sequence similarity to either of the established antiviral *SERINC* paralogs, *SERINC*2 is relatively similar to *SERINC*3 (50%) and *SERINC*5 (37%). Given such levels of sequence similarity and conserved membrane topology, we evaluated the ability of other *SERINC* paralogs to inhibit nef-defective-HIV-1 infectivity. We transiently transfected HEK293T cells with the individual *SERINC*s along with nef-defective HIV-1 (NL4-3 isolate). The viruses were collected, and infectivity assays were performed as reported previously (20, 21). In agreement with earlier reports (24–26), we confirmed that the inhibition of nef-defective-HIV-1 infectivity is conserved among four human *SERINC*s; *SERINC*2 is the only paralog that shows no detectable activity against the virus (Fig. 1B).

Based on the ability of human *SERINC*3 and *SERINC*5 to restrict nef-defective HIV-1, they were proposed as restriction factors (21, 22). However, in contrast to other restriction factors, *SERINC*5 and *SERINC*3 genes have a rather uneventful history that is distinct from the traditional signatures of recurrent selection seen in genes that are part of an arms race (19). We investigated whether the absence of signatures indicative of positive selection was prevalent across all the *SERINC* paralogs as a general feature. To this end, we compared the evolutionary signatures of *SERINC* genes with those of previously identified restriction factors—mainly genes showing recurrent positive selection and a few functionally characterized genes which act as controls (Fig. 1C). Although the arms race signatures of a few restriction factors such as BST-2 and EIAF2AK2 were well correlated, the other newly identified restriction factors, including *SERINC*s, did not show any consistent pattern of clustering (Fig. 1C). The *SERINC*5 transcript is not upregulated upon interferon (IFN) treatment (27, 28). However, one of the prime features of restriction factors is their ability to be augmented upon interferon stimulation, and this feature has been linked to the arms race with viruses (29). While the genes which formed a cluster are indeed interferon responsive (IFN-responsive genes were obtained from reference 30), this property does not explain the overall pattern of clustering (Fig. 1C). Furthermore, arms race signatures are not prominent in several interferon-inducible genes and innate immune genes, including Toll-like receptors (TLRs) (Fig. 1C). Therefore, the lack of arms race signatures in *SERINC*s is probably not very surprising. The ability to restrict HIV-1 among human SERINC paralogs suggests that this feature has been evolutionarily
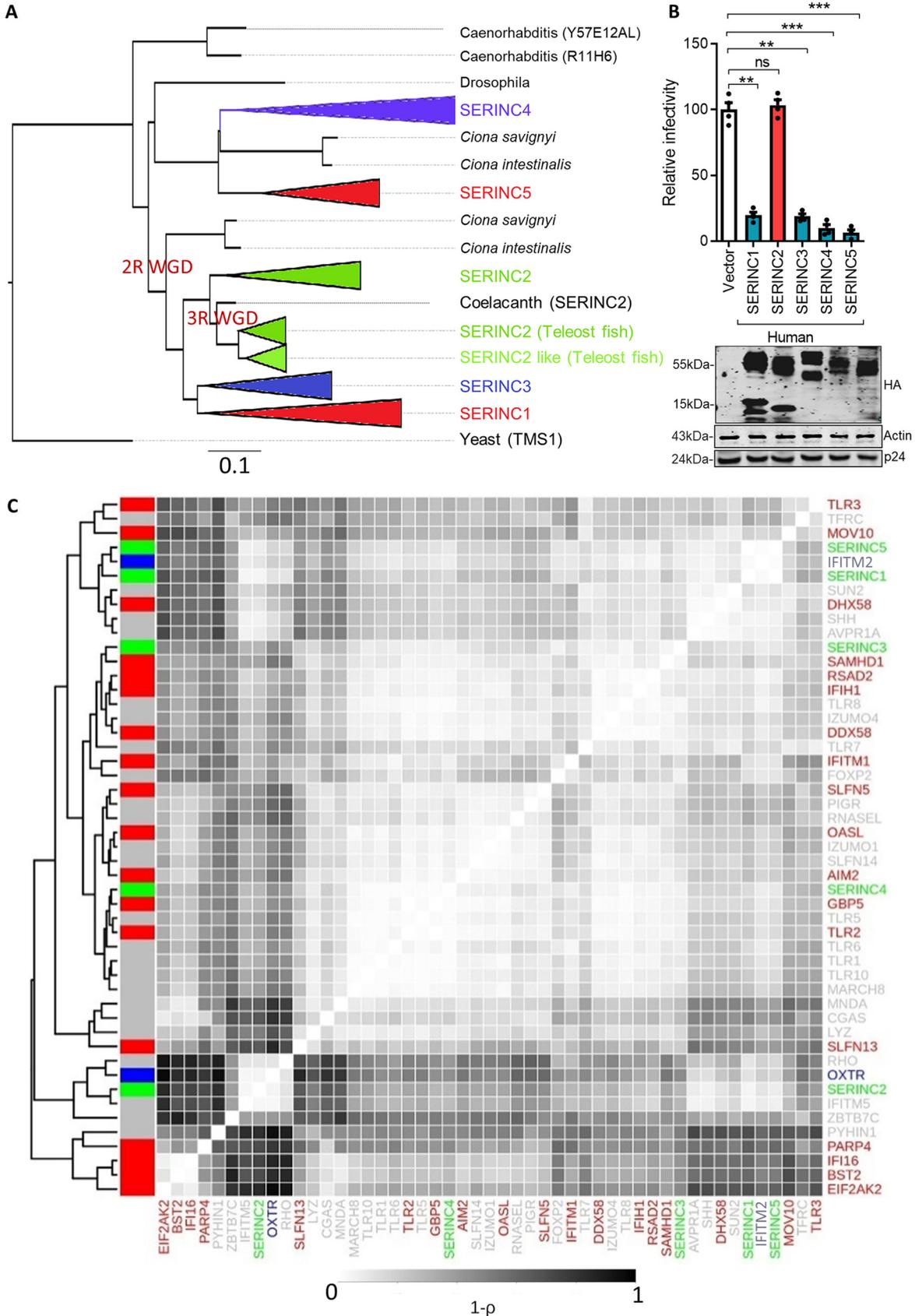
**FIG 1** The activity of human *SERINC* paralogs on HIV-1 infectivity and evolutionary history and arms race signatures of *SERINC* genes. (A) Phylogenetic analysis of *SERINC* paralogs. *S. cerevisiae, Caenorhabditis elegans,* and *D. melanogaster* have a single copy of this gene.

(Continued on next page)
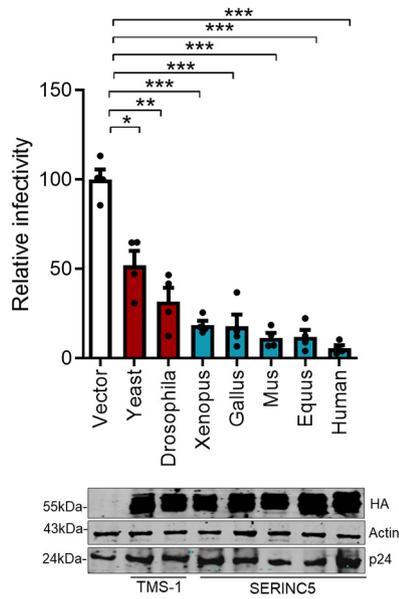
**FIG 2** Activity of *SERINC5* orthologs on HIV-1 infectivity. (Top) A single-cycle infectivity assay was performed using TZM-GFP target cells for nef-defective HIV-1 particles produced from HEK293T cells by transfecting 100 ng of indicated pre- and post-WGD SERINC5 orthologs encoded by pcDNA3.1($-$), human SERINC5 encoded by PBJ6, or the empty vector backbone along with pHXB2 envelope and NL4-3 Env($-$) Nef($-$). The infectivity was normalized to RT units using SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means ($n = 4$) and SD. An unpaired $t$ test was used. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. (Bottom) Western blot showing expression of C-terminally HA-tagged *SERINC5* orthologs and the corresponding $\beta$-actin and HIV-1 p24 from cell lysates (HEK293T).

conserved despite their early divergence. This observation prompted us further to investigate the evolutionary origins of this antiretroviral activity.

**Anti-HIV-1 activity is an ancient feature among *SERINC5* orthologs.** *SERINC*s plausibly shaped retroviral evolution, as indicated by the parallel emergence of anti-*SERINC*5 activity among diverse retroviral genomes (18, 23, 31, 32). Since the *SERINC* gene family is conserved across eukaryotic species, we investigated the extent to which the ability of *SERINC* orthologs to restrict HIV-1 is conserved. Unicellular eukaryotes and invertebrates have a single copy of the *SERINC* gene (*TMS*1); all these orthologs could restrict nef-defective HIV-1 (Fig. 2). However, the yeast (*Saccharomyces cerevisiae*) ortholog

**FIG 1** Legend (Continued)

Whole-genome duplication that occurred in an ancestor of mammals gave rise to multiple *SERINC* genes. The *SERINC* genes form two clusters consisting of *SERINC*1, *SERINC*2, and *SERINC*3 and of *SERINC*4 and *SERINC*5. The members of the first cluster, *SERINC*1, -2, and -3, are grouped as ohnologs based on a strict Q-score criterion by the http://ohnologs.curie.fr/ database. The second cluster, consisting of *SERINC*4 and -5, potentially reflects a subsequent duplication event that is shared by most of the post-WGD species. Colors denote different branches of each *SERINC* in chordates. The lengths of the triangles are proportional to the number of nucleotide substitutions that have taken place in a particular branch. The tree was generated using Ensembl. (B) (Top) A single-cycle infectivity assay was performed using TZM-GFP cells for nef-defective HIV-1 particles that were produced from HEK293T cells by transfecting 100 ng of indicated human SERINC paralogs encoded by pcDNA3.1($-$), SERINC5 encoded by PBJ6, or the empty vector backbone along with pHXB2 envelope and NL4-3 Env($-$) Nef($-$). The infectivity was normalized to RT units using an SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means ($n = 4$) and standard deviations (SD). An unpaired $t$ test was used to determine significance. **, $P < 0.01$; ***, $P < 0.001$; ns, not significant. (Bottom) Western blot showing expression of C-terminally HA-tagged *SERINC* and the corresponding $\beta$-actin and HIV-1 p24 from HEK293T cell lysate. (C) Hierarchical clustering of the arms race signatures in primate genes aligned using ClustalW aligner and used as input for FUBAR. Interferon-induced genes identified in reference 30 are color coded red (upregulated), blue (downregulated), green (*SERINC* paralogs), and gray (other selected genes). FUBAR quantifies evolutionary fingerprints of a gene by dividing the synonymous ($\alpha$) and nonsynonymous ($\beta$) rates into a finely discretized grid. Each point in this grid represents a pair of $\alpha$ and $\beta$ values, with the size of the circle representing the value $\hat{\theta}$ (the posterior mean), which provides a measure of how many of the sites in the multiple-sequence alignment contribute to that particular combination of $\alpha$ and $\beta$ values. The vector of 400 $\hat{\theta}$ values that correspond to each of the points on the 20-by-20 grid of $\alpha$ and $\beta$ values serves as the evolutionary fingerprint of a gene. Genes with canonical arms race signatures have a greater weight for grids with higher values of $\beta$ than genes that lack the arms race signature (19).

showed a relatively modest activity (2-fold) in comparison to all the other orthologs (inhibition between 4-fold and 36-fold). All these experiments were conducted under comparable transfection conditions. We also confirmed the expression of C-terminally hemagglutinin (HA)-tagged TMS-1 and other *SERINC*5 orthologs by Western blotting from cell lysates (Fig. 2, lower panel). Overall, these results indicate that anti-HIV activity is an ancient and evolutionarily conserved feature among *SERINC*s that predates the whole-genome duplication event between ~700 to 400 million years ago (MYA) (33).

**Coelacanth *SERINC*2 inhibits HIV-1.** Two rounds of whole-genome duplication (WGD) in the ancestor of chordates led to a substantial increase in the number of genes, thereby leading to the acquisition of new functions. This repertoire of genes also provides greater flexibility due to their mutually compensatory functions. Hence, post-WGD paralogous copies tend to diversify. A single copy of a *SERINC* ortholog (*TMS*1) is present in pre-WGD species, and post-WGD, the number of copies has increased to five in tetrapods and six in bony fishes (Fig. 1A). Our investigation of the pre-WGD ortholog of *SERINC* genes from yeasts (*Saccharomyces cerevisiae*) and flies (*Drosophila melanogaster*) *TMS1* found evidence of antiretroviral activity (Fig. 2). Functional data on the human paralogs that diverged early also suggest an ancestrally conserved activity of *SERINC*s in restricting HIV-1, with *SERINC*2 being the exception (Fig. 1B). Hence, the human *SERINC*2 paralog may have lost the ability to restrict retrovirus sometime after the WGD event (~700 to 400 MYA) (33). It is plausible that the species proximal to the WGD event might retain a version of *SERINC2* with antiviral activity (Fig. 3A). Therefore, we decided to systematically screen *SERINC2* orthologs from post-WGD species at various levels of sequence divergence from human *SERINC2*. To this end, we synthesized *SERINC*2 from various post-WGD species. The codons had to be optimized using the GeneArt portal for expression in human cells owing to their nonhuman origin. During virus production, coexpression of select *SERINC*2 orthologs and subsequent infectivity analysis revealed that coelacanth *SERINC*2 restricts nef-defective HIV-1 (~7-fold) under conditions in which *SERINC*2 isoforms from humans did not show any activity (Fig. 3B). A shorter isoform of human *SERINC*2 (Human-201) was found to be topologically similar (Fig. 3C) to that of coelacanth *SERINC*2 but lacked the activity to restrict nef-defective HIV-1 concurrent with the longer isoform (Fig. 3B). The cellular localization of coelacanth SERINC2 was assessed and found to be comparable to that of SERINC2 isoforms and SERINC5 from human (Fig. 3D). We saw a dose-dependent inhibition when coelacanth *SERINC*2 was expressed from plasmids carrying promoters of various strengths (PBJ6-, PBJ5-, and pcDNA-based expression vectors; the expression levels are indicated in the Western blots), suggesting that this is not an artifact of overexpression (Fig. 3E). We also verified the activity of coelacanth *SERINC*2 from a T-cell line that lacked endogenous *SERINC*3 and *SERINC*5 using the same constructs used in HEK293T cells and found consistent inhibition by coelacanth *SERINC*2 (Fig. 3F). Together, the experiments also suggested that the effect of coelacanth *SERINC*2 on nef-defective HIV-1 infectivity inhibition is not cell type specific.

**Gradual loss of antiretroviral activity in *SERINC*2.** Upon further assessment of the anti-HIV-1 activity of post-WGD *SERINC*2 orthologs, we found that while coelacanth *SERINC*2 reduced the infectivity by ~7-fold, *Xenopus SERINC*2 exhibited a modest inhibition (~2-fold). The activity is completely lost in the chicken *SERINC*2 ortholog and onwards (Fig. 3G). However, this lack of activity is persistent in mouse, horse, and human *SERINC*2 at comparable expression levels (Fig. 3G, bottom). Loss of human *SERINC*2 antiviral activity could have been associated with changes in pathogen repertoire or neofunctionalization. To experimentally test if this was in response to a change in the pathogen repertoire, we asked if the counteraction of human *SERINC*5 by known potent retroviral factors, HIV-1 Nef and murine leukemia virus (MLV) glycoGag, was analogous to that of coelacanth *SERINC*2. Under conditions in which Nef and glycoGag efficiently counteracted the restriction exerted by the human *SERINC*5 and the partial restriction of *Xenopus SERINC*2, counteraction of coelacanth *SERINC*2 restriction by these virulent factors was not apparent (Fig. 4A). We used a potent *SERINC*5 antagonist that we reported earlier (simian immunodeficiency virus SIVmac239-encoded Nef) and wondered if this was an allele-specific effect (21). Therefore, we next investigated
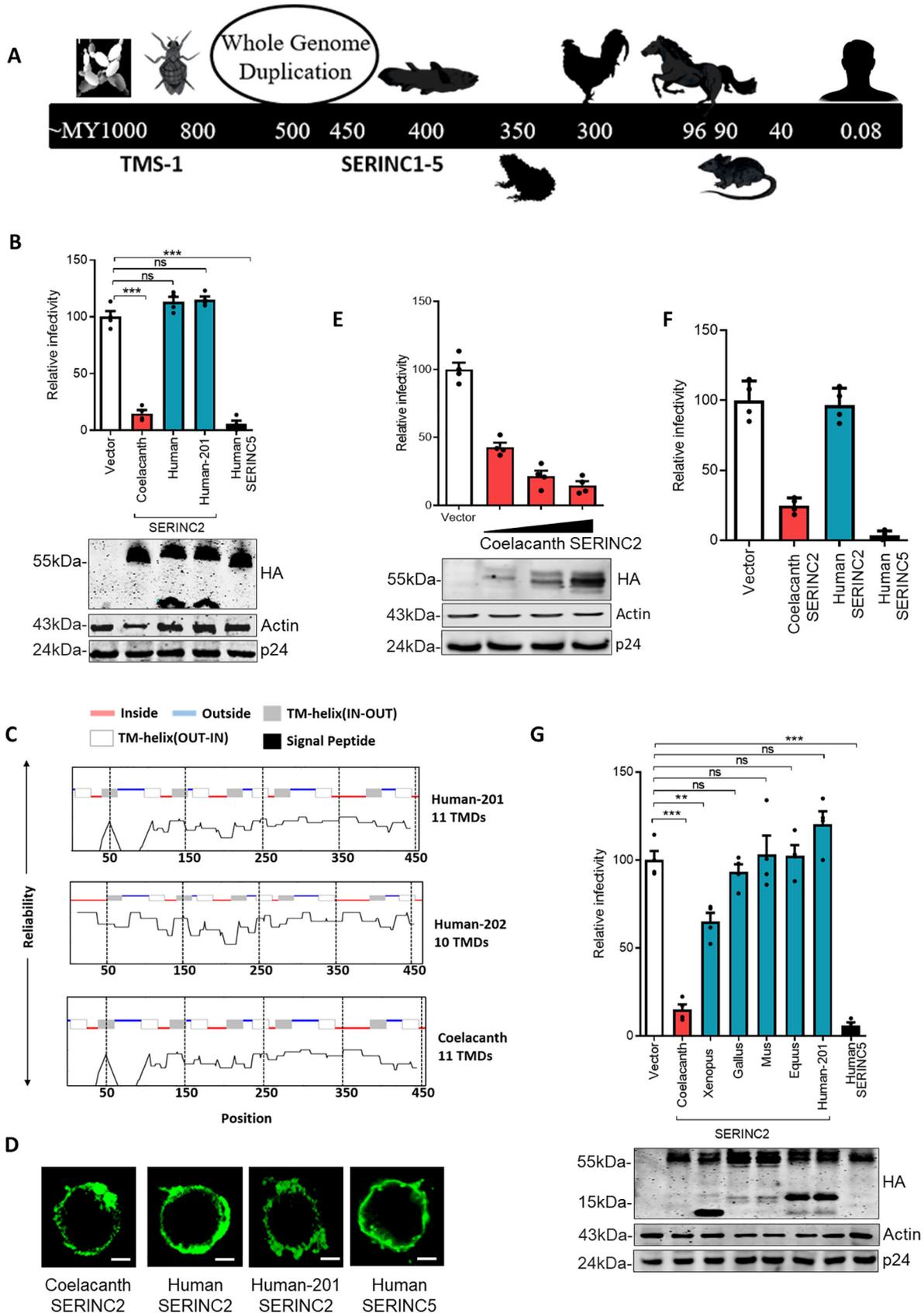
FIG 3 Coelacanth *SERINC*2 inhibits HIV-1 infectivity. (A) Schematic timeline depicting the sequence of events during the course of evolution, with the species (left to right) *S. cerevisiae*, *D. melanogaster*, *L. chalumnae*, *X. tropicalis*, *G. gallus*, *E. caballus*, *M. musculus*, and

(Continued on next page)

whether the representative *nef* alleles from human and nonhuman primate lentiviruses showed a similar phenotype and found that nef alleles did not rescue the infectivity in this case as it did for human *SERINC*5 (Fig. 4B). It was also earlier reported by us and others that the ability of *SERINC*5 to restrict HIV-1 inhibition varies with the envelope glycoproteins used for pseudotyping (20, 21). We checked if coelacanth *SERINC*2-mediated inhibition is dependent on the envelope glycoprotein usage. Coelacanth *SERINC*2 action indeed mirrored that of human *SERINC*5 in terms of the envelope sensitivity (Fig. 4C), further highlighting a functional similarity. However, Nef and glycoGag failed to rescue this analogous activity of coelacanth *SERINC*2, indicating a unique feature that perhaps made it invisible to these retroviral factors.

**Human foamy virus envelope counteracts coelacanth *SERINC*2.** Three distinct retroviruses were reported to have independently come up with antagonizing factors to elude the inhibition by *SERINC*5 (20–22, 35). The lack of activity in Nef and glycoGag against coelacanth *SERINC*2 (Fig. 4A and B) prompted us to check whether antiretroviral SERINC2 has evolved to inhibit other retroviruses. We learnt that the coelacanth has an endogenous foamy virus (36), the genome organization of which resembles that of the prototype foamy virus (FV) (Fig. 5A). Recent work has also shown the existence of endogenous foamy virus in reptiles, birds, and various mammals (37, 38). To experimentally test the presence of anti-FV activity in human *SERINC* paralogs as well as the *SERINC*2 orthologs, we first tested the activity of all human SERINCs against a human FV. Surprisingly, we find that FV is insensitive to any of the human *SERINC* paralogs tested as well as the coelacanth *SERINC*2, under conditions in which nef-defective HIV-1 was consistently inhibited (Fig. 5B). We argued that this might have been associated with the intrinsic ability of an FV-encoded factor to counter the inhibition. To delineate this, we coexpressed foamy virus genes individually to check their ability to rescue nef-defective HIV-1, so that the insensitivity of FV and the presence of antagonizing factor could be revealed. Surprisingly, with none of the FV components expressed in *trans*, we found an ability to rescue the inhibition exerted by human *SERINC*5. Under these conditions, however, Nef completely antagonized the *SERINC*5 restriction (Fig. 5C). The inhibition exerted by coelacanth *SERINC*2, interestingly, was antagonized by the FV envelope glycoprotein (Fig. 5D). The phenotype was also consistent when we expressed increasing amounts (50, 100, or 200 ng) of FV envelope (Fig. 5E). Due to the unavailability of commercial antibodies against FV vectors, we attempted to tag the FV envelope

**FIG 3** Legend (Continued)

*H. sapiens*. (B) (Top) A single-cycle infectivity assay was performed using TZM-GFP target cells for nef-defective HIV-1 particles that were produced from HEK293T cells by transfecting 100 ng of coelacanth SERINC2, human SERINC2 isoforms encoded by pcDNA3.1(−), human SERINC5 encoded by PBJ6, or the empty vector backbone along with pHXB2 envelope and NL4-3 Env(−) Nef(−). The infectivity was normalized to RT units using an SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means ($n = 4$) and SD. An unpaired *t* test was used. ***, $P < 0.001$; ns, not significant. (Bottom) Western blot showing expression of C-terminally HA-tagged *SERINC* expressors and the corresponding $\beta$-actin and HIV-1 P24 from cell lysate (HEK293T). (C) Topological features of human *SERINC*2 splice isoforms and coelacanth *SERINC*2 predicted using TOPCONS (80). (D) Immunofluorescence of HA-tagged coelacanth SERINC2, human SERINC2 isoforms, and human SERINC5 transfected in JTAg*SERINC*5/3$^{−/−}$ cells and visualized using mouse anti-HA primary antibody and mouse Alexa Fluor 488-conjugated secondary antibody. Bar, 10 $\mu$m. (E) (Top) A single-cycle infectivity assay was performed using TZM-GFP target cells for nef-defective HIV-1 particles that were produced from HEK293T cells, by transfecting coelacanth encoded by PBJ6, PBJ5, pcDNA3.1(−), or the empty vector backbone along with pHXB2 envelope and NL4-3 Env(−) Nef(−). The infectivity was normalized to RT units using an SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means ($n = 4$) and SD. (Bottom) Western blot showing expression of C-terminally HA-tagged coelacanth SERINC2 and the corresponding $\beta$-actin and p24 from cell lysates (HEK293T). (F) A single-cycle infectivity assay was performed using TZM-GFP target cells for nef-defective HIV-1 particles that were produced from JTAg SERINC5/3$^{−/−}$, by transfecting coelacanth SERINC2, human SERINC2, and human SERINC5 encoded by pcDNA3.1(−) or the empty vector backbone along with pHXB2 envelope and NL4-3 Env(−) Nef(−). The infectivity was normalized to RT units using an SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means ($n = 3$) and SD. (G) (Top) A single-cycle infectivity assay was performed using TZM-GFP target cells for nef-defective HIV-1 particles that were produced from HEK293T cells by transfecting 100 ng of SERINC2 orthologs (coelacanth, *Xenopus*, *Gallus*, *Mus*, *Equus*, and human) encoded by pcDNA3.1(−) and human SERINC5 encoded by PBJ6 or the empty vector backbone along with pHXB2 envelope and NL4-3 Env(−) Nef(−). The infectivity was normalized to RT units using an SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means ($n = 4$) and SD. An unpaired *t* test was used. **, $P < 0.01$; ***, $P < 0.001$; ns, not significant. (Bottom) Western blot showing expression of C-terminally HA-tagged *SERINC* expressors and the corresponding $\beta$-actin and HIV-1 p24 from cell lysate (HEK293T).
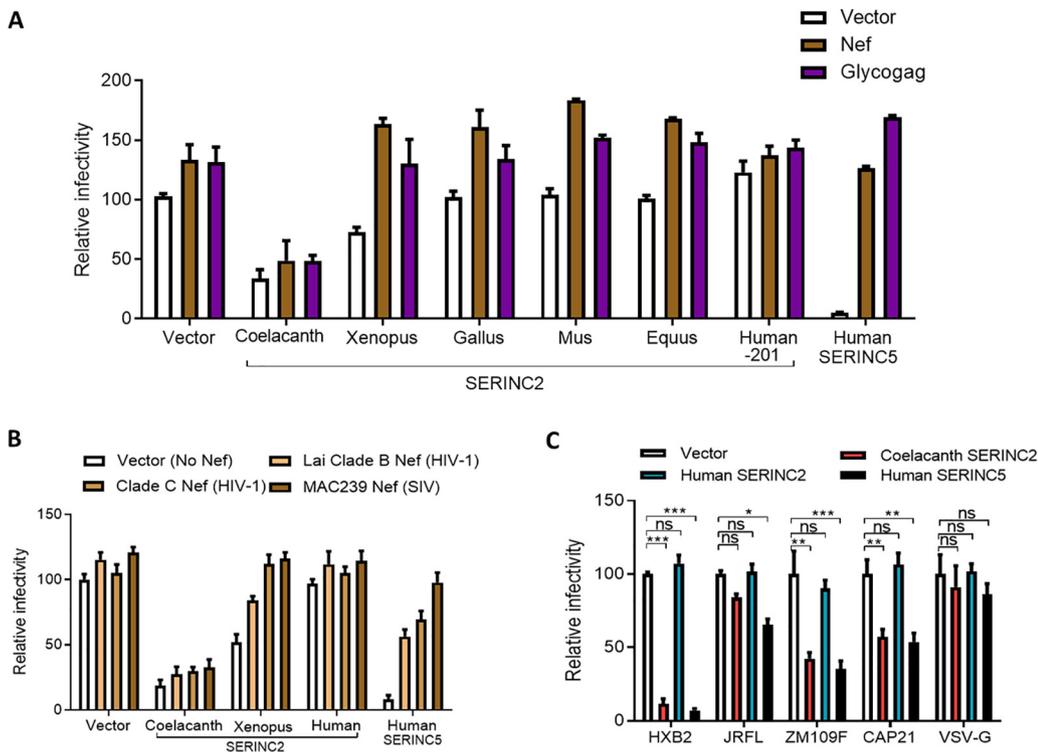
FIG 4 Ability of retroviral factors to antagonize *SERINC*s. (A) A single-cycle infectivity assay was performed using TZM-GFP target cells for HIV-1 particles that were produced from HEK293T cells either by transfecting Nef, glycoGag, or empty vector in in combination with 100 ng of SERINC2 orthologs (coelacanth, *Xenopus*, *Gallus*, *Mus*, *Equus*, and human) and human SERINC5 encoded by PBJ6 or the empty vector backbone along with pHXB2 envelope and NL4-3 Env(−) Nef(−). The infectivity was normalized to RT units using an SGPERT assay. Human *SERINC5* served as a control for Nef and glycoGag counteraction of the exerted restriction. The values obtained from the empty vector control was set to 100 for comparison with SERINC expressors. Values are means (*n* = 4) and SD. (B) A single-cycle infectivity assay was performed using TZM-GFP target cells for HIV-1 particles that were produced from HEK293T cells by transfecting the indicated Nef alleles or empty vector in in combination with 100 ng of SERINC2 orthologs (coelacanth, *Xenopus*, and human) and human SERINC5 encoded by PBJ6 or the empty vector backbone along with pHXB2 envelope and NL4-3 Env(−) Nef(−). The infectivity was normalized to RT units using an SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means (*n* = 4) and SD. (C) A single-cycle infectivity assay was performed using TZM-GFP target cells for nef-defective HIV-1 particles that were produced from HEK293T cells by transfecting 100 ng of coelacanth SERINC2, human SERINC2, and human SERINC5 encoded by PBJ6 or the empty vector backbone along with NL4-3 Env(−) Nef(−) with HIV-1 clade B (HXB2 and JR-FL), clade C (ZM109F and CAP210) and vesicular stomatitis virus envelope glycoprotein (VSV-G). The infectivity was normalized to RT units using an SGPERT assay. Human *SERINC5* served as control for Nef and glycoGag counteraction of the exerted restriction. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means (*n* = 4) and SD. An unpaired *t* test was used. *, *P* < 0.05; **, *P* < 0.01; ***, *P* < 0.001; ns, not significant.

glycoprotein with HA and FLAG for visualizing the expression by Western blotting. However, the C-terminal epitope tagging resulted in noninfectious FV particles (data not shown), suggesting interference of the tag with envelope function. We reasoned that the constructs were valid, as they consistently contributed to the production of functional virus particles in the infectivity assays, which was considered proof of expression for foamy components. Regardless, we confirmed the mRNA-level expression using specific primers for the indicated FV genes and found that they were expressed in these experimental settings (Fig. 5C and D, lower panels).

Having found an envelope counteracting the SERINC2 activity, we further investigated whether this rescue was an effect of cross-packaging of the HIV-1 core by the FV envelope. By using a pseudoparticle comprising an HIV-1 core complemented with the FV envelope, we showed that the effect was not due to cross-packaging (Fig. 5F), as the FV envelope incorporation into budding virions would yield infectious particles under these conditions, which was not the case. This rescue by the FV envelope glycoprotein,
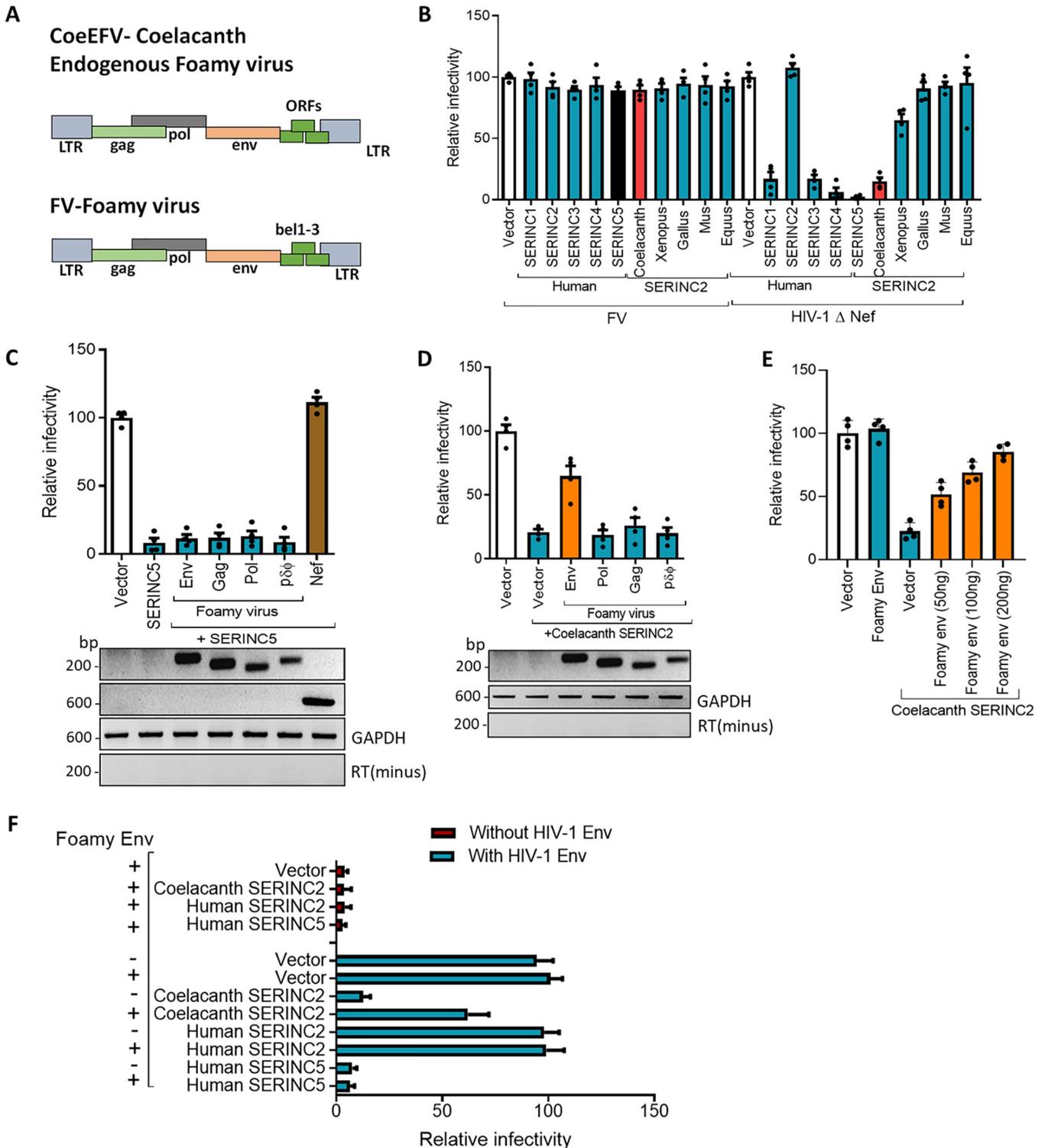
FIG 5 Effect of coelacanth *SERINC*2 on retrovirus infectivity and its antagonism by the envelope glycoprotein. (A) Genome organization of an endogenous foamy virus (CoeEFV) from coelacanth and the prototype foamy virus (FV). (B) A single-cycle infectivity assay was performed using TZM-GFP target cells for FV-derived vectors and nef-defective HIV-1 particles that were produced from HEK293T cells by transfecting 100 ng of SERINC2 orthologs (coelacanth, *Xenopus*, *Gallus*, *Mus*, *Equus*, and human) encoded by pcDNA3.1(−) and human SERINC5 encoded by PBJ6 or the empty vector backbone along with FV-derived vectors for FV and NL4-3 Env(−) Nef(−) and pHXB2 for HIV-1 infectivity. The infectivity was normalized to RT units using an SGPERT assay. Human *SERINC*5 served as a control for Nef and glycoGag counteraction of the exerted restriction. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means (*n* = 4) and SD. (C) (Top) A single-cycle infectivity assay was performed using TZM-GFP target cells for nef-defective HIV-1 particles that were produced from HEK293T cells by cotransfecting FV vectors (env, pol, gag, and transfer vector) and 100 ng of human SERINC5 encoded by PBJ6 or the empty vector backbone along with pHXB2 envelope and NL4-3 Env(−) Nef(−). The infectivity was normalized to RT units using an SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means (*n* = 4) and SD. (Bottom) RT-PCR for the indicated FV vectors and Nef from transfected cell lysate (HEK293T). GAPDH (glyceraldehyde-3-phosphate

therefore, appears to be specific to its ability to antagonize coelacanth *SERINC*2, and thereby, it restores most of the HIV-1 infectivity (Fig. 5D).

**FV envelope prevents *SERINC2* incorporation into virions.** We previously established that virulence factors like Nef, S2, and glycoGag promoted the exclusion of *SERINC*5 from virions to increase the particle infectivity (20, 21). We investigated whether the FV envelope glycoprotein acts similarly. Nef-defective particles were produced from HEK293T cells in the presence of coelacanth *SERINC*2 and FV envelope glycoprotein. Western blotting revealed that although coelacanth *SERINC*2 was readily detectable in the nef-defective virions, the expression of FV envelope diminished the signal to a level similar to the background (Fig. 6A). This experiment indicated that, like Nef, S2, and glycoGag, the FV envelope enhances the particle infectivity by preventing the virion incorporation of the coelacanth *SERINC*2.

Strikingly, we observed that when FV envelope was coexpressed, the steady-state level of coelacanth *SERINC*2 in the producer cell lysates was affected (Fig. 6A, Western blots from producer cells). This indicates that FV envelope affects the stability of coelacanth *SERINC*2 by promoting its degradation. To examine this further, we tested the effect of FV envelope in HEK293T cells by expressing increasing amounts of coelacanth *SERINC*2 (50, 100, and 200 ng) against a constant amount of FV envelope (100 ng) glycoprotein-encoding plasmid. Western blotting revealed that coexpression of FV envelope glycoprotein resulted in a decreased amount of coelacanth *SERINC*2 in cell lysates regardless of the amount of plasmid transfected (Fig. 6B). In contrast to the effect of FV envelope observed on coelacanth *SERINC*2, the steady-state level of the protein remained unperturbed when it was coexpressed with an HIV-1 envelope glycoprotein under similar experimental conditions (Fig. 6B, bottom).

Interestingly, we found that FV envelope decreases steady-state levels of coelacanth *SERINC*2 in a bafilomycin-sensitive manner, implying that coelacanth *SERINC*2 is targeted to lysosomes for destruction. Accordingly, the ability of FV env to rescue the infectivity is impaired upon challenge with bafilomycin during the virus production but not by the proteasomal inhibitor MG132 (Fig. 6C). Notably, we did not see any apparent change in the expression levels of human *SERINC5* or *SERINC2* upon FV envelope expression (Fig. 6C, bottom). Accordingly, the infectivity inhibition exerted by *SERINC5* is not counteracted by FV envelope glycoprotein, which is consistent with the data in Fig. 5C. Further, bafilomycin A1 (lysosomal acidification inhibitor) and MG132 (proteasomal inhibitor) did not affect the outcome of infectivity in the case of human *SERINC2*. Taken together, these experiments suggested that the effect of FV envelope on the steady-state levels of coelacanth *SERINC*2 is specific.

**Sensitivity of FV to coelacanth *SERINC2*.** Further, we questioned whether the foamy virus was at all sensitive to the effects of coelacanth *SERINC2*. To this end, we pseudotyped FV core with an envelope glycoprotein of HIV-1 lacking a C-terminal tail, because the native full-length HIV-1 envelope was not capable of packaging the core to produce infectious virions (Fig. 6D). Using a HIV-1 envelope with a C-terminal deletion, we showed that the FV core is sensitive to coelacanth *SERINC2* restriction (Fig. 6D).

**FIG 6** FV envelope targets coelacanth *SERINC*2 for degradation. (A) (Top) Western blot of HA-tagged coelacanth *SERINC*2 with the corresponding p24 from virions in the presence and absence of foamy virus envelope. (Bottom) Western blot of HA-tagged coelacanth *SERINC*2 with corresponding $\beta$-actin from HEK293T cell lysate. (B) (Top) Western blot of HA-tagged coelacanth *SERINC*2 at 50 ng, 100 ng, and 200 ng in the presence (+) and absence (−) of FV envelope (200 ng) and corresponding $\beta$-actin from HEK293T cell lysate from a 12-well plate. (Bottom) Western blot of HA-tagged coelacanth *SERINC*2 at 50 ng, 100 ng, and 200 ng in the presence (+) and absence (−) of HIV-1 envelope and corresponding $\beta$-actin and HIV-1 gp41 from HEK293T cell lysate from a 12-well plate. (C) (Top) A single-cycle infectivity assay was performed using TZM-GFP target cells for nef-defective HIV-1 particles that were produced from HEK293T cells by transfecting 100 ng of coelacanth *SERINC*2, human *SERINC*2 encoded by pcDNA3.1(−), human *SERINC*5 encoded by PBJ6, or the empty vector backbone along with pHXB2 envelope and NL4-3 Env (-) Nef(−) in the presence or absence of FV env. In the presence of FV env, the cells were treated with either DMSO control, bafilomycin A1 (100 nM), or MG132 (20 $\mu$M) for 15 h. The infectivity was normalized to RT units using an SGPERT assay. The values obtained from the empty vector control were set to 100 for comparison with SERINC expressors. Values are means ($n = 4$) and SD. (Bottom) Western blot showing expression of C-terminally HA-tagged *SERINC* expressors for indicated conditions and the corresponding $\beta$-actin and HIV-1 p24 from cell lysate (HEK293T). (D) A single-cycle infectivity assay was performed using TZM-GFP target cells for FV particles produced from HEK293T cells. The foamy core was pseudotyped with either FV envelope, HIV-1 envelope, or HIV-1 envelope lacking a C-terminal tail (ΔCt), and the indicated coelacanth SERINC2, human SERINC2 encoded by pcDNA3.1(−), human SERINC5 encoded by PBJ6, or the empty vector backbone was coexpressed to check the ability of individual *SERINC* to restrict foamy virus. Values are means ($n = 4$) and SD.

Altogether, this indicates that the FV core is susceptible to the action of coelacanth *SERINC*2, and the virus possesses a mechanism to antagonize the inhibitory effect.

## DISCUSSION

The use of comparative evolutionary genetic approaches has continued to enrich our understanding of restriction factor biology for more than a decade (7, 39–42). Reciprocal loss of duplicated genes in different species has been shown to contribute

**A**



**B**



**C**



**FIG 7** Sequence divergence among SERINC2 orthologs and structural analysis. (A) Multiple-sequence alignment (MSA) of *SERINC*2 orthologs using ClustalW; highlighted regions are the sites of sequence divergence. (B) Cryo-EM structure of TMS-1 from *D. melanogaster* reported in reference 56, visualized using

toward species-specific differences in susceptibility to pathogens (43–45). We found that, similar to reciprocal gene loss, *SERINC*2 and *SERINC*5 show reciprocal functional adaptation against divergent retroviruses. While the spectrum of activity among human paralogs we tested here is in agreement with previous findings (21, 22, 24, 25), the ability of *SERINC* orthologs to restrict HIV-1 is also remarkable. Since exogenous retroviruses have not yet been reported from pre-WGD species that we considered, the range of targets that a pre-WGD *SERINC* ortholog (TMS-1) will have remains to be ascertained. While genes such as CCR4 and DHH1 that physically interact with TMS-1 have been implicated in yeast retrotransposon activity (46, 47), the mechanism that TMS-1 would manifest on retroelements, if any, remains to be investigated. TMS-1 in flies may still be required to regulate the mobility of gypsy retroelements, as the gypsy envelope has been used to pseudotype Moloney murine leukemia virus-based vectors to efficiently infect fly cells (48, 49).

Moreover, the analogous mechanism that Rous sarcoma virus uses to produce Pr180gag-pol (as found in yeast Ty-1 transposons [50]) makes the existence of retroelement interactions for other host factors like TMS-1 more probable. Post-WGD, *SERINC* has five copies, and this may have been associated with tackling an increasing diversity of pathogens during speciation while simultaneously retaining core functionality. As demonstrated here with cross-packaging studies, we show that a *SERINC*2 ortholog—which was thought to be inactive—may have constituted a critical barrier to foamy viruses early on, plausibly leading to their divergence (36, 51). Loss of activity in other *SERINC*2 orthologs may have been associated with neofunctionalization, as suggested by localized sequence divergence (Fig. 7A and B; also, see Videos S1 and S2 in the supplemental material), the presence of HNF4alpha binding enhancer (Fig. 7C), and changes in tissue-specific expression patterns (see Data Set S1) exemplified by expression of *SERINC*2 in the livers of primates. For the first time, we observed a naturally occurring antiretroviral *SERINC2* from coelacanth. Future studies exploring the sequence level differences shown in Fig. 7B (also see Videos S1 and S2) among functionally distinct *SERINC*2 orthologs may help in dissecting the antiretroviral function from its putative tissue-specific function. Interestingly, coelacanth *SERINC*2 inhibited HIV-1 and remained invisible to the most potent retroviral virulent factors, Nef (SIVmac239) and the glycoGag from MLV. Further studies can provide more insights into the lack of activity in these retroviral factors and require relevant models to study physiological relevance.

Foamy viruses codiverged with their hosts (36, 38, 52), and this interrelation is visible in their ability to introduce variations in envelope glycoproteins, among other documented genomic alterations. The FV envelope counteracting *SERINC*2 is the first evidence of a canonical gene product being employed to evade *SERINC*2 restriction. The FV envelope's ability to engage the endolysosomal machinery of the host to selectively target coelacanth *SERINC*2, but not human *SERINC*2 and *SERINC*5, for degradation in a heterologous host is impressive, as was the case with other retroviral factors (53, 54). The counteraction phenotype associated with the FV envelope demonstrated here is also in agreement with the ability of *SERINC*5 to restrict the virus in an envelope-dependent manner (20, 21, 25, 31, 55, 56), indicating the envelope as a determinant for *SERINC* sensitivity to particle restriction. Therefore, the finding of FV envelope-dependent activity against coelacanth SERINC2 is in agreement with previous reports (55, 56). We acknowledge that for the study of foamy viruses and their interaction with coelacanth *SERINC*2, the absence of a native experimental model was a limitation. The availability of such experimental models would have strengthened the findings further to understand the role of *SERINC*s in shaping the evolution of retroviruses other than HIV-1. Further, no exogenous retrovirus is known to infect coelacanths (36). A better

**FIG 7** Legend (Continued)

ChimeraX software. The regions highlighted in the MSA in panel A are in white (first region), yellow (second region), and green (third region). (Top) Monomer; (middle) hexamer front view; (bottom) hexamer rear view. (C) Evidence for the presence of an HNF4a binding site in the intron of the human SERINC2 gene.

understanding of SERINC activity against foamy viruses therefore awaits discovery of new exogenous foamy viruses.

Similarly, the role of *SERINC*5 from coelacanths in inhibiting retrovirus remains to be established due to the unavailability of reliable genomic sequence information to reconstruct a functional protein-coding gene. An early challenge by *SERINC*2 may have led to subsequent divergence of the FV envelope toward insensitivity to *SERINC*2 restriction. Different *SERINC* paralogs, therefore, might have specialized in restraining specific retroviruses, leading to the coevolution in response to specific paralogs (32, 57).

We foresee that weaker signatures, exemplified by *SERINC*s despite the constant challenge from viruses, could be due to the native functions of such transmembrane proteins, where adaptation through diversification in response to the pathogen would result in the loss of a core function (58). The poor signatures could also be because the antiretroviral activity is spread out over multiple host genes; for instance, a recent report shows that a Nef-sensitive TIM1 activity is potentiated by *SERINC*5 (59). The host, therefore, can afford such redundancy without having to diversify much. Another example is the TLRs, where the extracellular domain shows signatures of recurrent positive selection in contrast to the conserved membrane-spanning region (15). Intriguingly, virus-specific TLRs are under stronger purifying selection than nonviral TLRs (60), potentially due to the larger number of pathogen-associated molecular patterns (PAMPs) associated with nonviral pathogens. This constraint, however, may be more pronounced in *SERINC*s, as they are multipass transmembrane proteins known to inhibit only retroviruses. While the core function of *SERINC*s in eukaryotes awaits independent observations (61), our study indicates more *SERINC*-like restriction factors that display poor signatures of arms race but are functionally active against viruses.

In conclusion, evolution-guided analysis for tracking the origin of antiviral activity in the *SERINC* genes and the dynamics following whole-genome duplication have identified the presence of antiretroviral activity in the only *SERINC* thought to be deficient. The antiretroviral activity among *SERINC*s may have shaped the evolution of distant retroviruses, and the presence of an evasion strategy in a spumavirus representative to target coelacanth *SERINC*2 implies a fundamental role of these host factors in shaping retrovirus evolution.

## MATERIALS AND METHODS

**Plasmids and reagents.** The plasmids and reagents, including antibodies used in the current study, are presented in Tables SI and SII, respectively.

**Viruses and infectivity assay.** Viruses were produced from HEK293T (ECACC) cells (21) plated in 10-cm$^2$ dishes by calcium phosphate transfection and were limited to single-cycle replication. Seven micrograms of NL4-3 (env defective and Nef defective), 1 $\mu$g env-expressing plasmid, and 100 ng of plasmids expressing *SERINC* genes (Table SI) or the equivalent corresponding empty vector and PBJ6 *SERINC*5 HA were used for virus production. Foamy viruses were produced by transient transfection using pCIES (0.736 $\mu$g), pCIPS (1.5 $\mu$g), pCIGS (11.84 $\mu$g), p$\Delta\Phi$ (11.84 $\mu$g) (from reference 62; plasmid details are provided in Table SI) and pcDNA (100 ng) vectors expressing *SERINC* genes or an equivalent amount of control vector in HEK293T cells. The viral particles were collected as cell culture supernatant 48 h posttransfection, centrifuged at 300 $\times$ *g* for 5 min, and quantified using an SG-PERT reverse transcription assay (63, 64). Following this, viruses were diluted for infection in TZM-GFP reporter cells (21) that were seeded in a 96-well plate, 24 h prior to infection. HIV-1 infectivity was quantified by scoring the green nuclei using a SpectraMax MiniMax 300 imaging cytometer (Molecular Devices, USA). Foamy virus infectivity was examined by determining the number of green fluorescent protein (GFP)-positive cells expressed from the transducing vector carrying a GFP expression cassette (p$\Delta\Phi$), indicating the fraction of the cell population transduced. The acquired values were normalized to the reverse transcriptase units obtained from the SG-PERT assay as previously described (20). Results are expressed as percentages of the value for a vector control, normalized to 100.

**Immunofluorescence.** For electroporation of JTAg$^{SERINC3/5KO}$ (20), cells in exponential growth phase were harvested (10$^7$ cells/sample) at 300 $\times$ *g* for 10 min. Prior to addition of Opti-MEM, the cells were washed with 1$\times$ phosphate-buffered saline (PBS [1$\times$], pH 7.0) to remove residual serum and cell debris. Each sample was resuspended in 200 $\mu$l warm Opti-MEM. Five micrograms of constructs expressing HA-tagged *SERINC*2 orthologs or equivalent control vector was then mixed into the suspended cells. The cells and DNA mixture were added to a 2-mm-gap electroporation cuvette (Bio-Rad, USA). The cells were pulsed at 140 V and 1,000 $\mu$F with exponential decay on a Bio-Rad GenePulserXcell module. Warm RPMI (600 $\mu$l) with 20% fetal bovine serum (FBS) was immediately added to the electroporated cells, which were then transferred to a 6-well plate containing RPMI with 10% FBS. Forty-eight hours

posttransfection, cells were spun down at 300 × *g* for 5 min, resuspended in 100 μl RPMI, laid on poly-L-lysine-coated glass coverslips (Genetix), and fixed with 4% paraformaldehyde (PFA). After fixation, the cells were washed twice with 1× PBS. The cells were then permeabilized using BD Perm/Wash followed by detection of HA tag with purified anti-HA.11 epitope tag antibody (1:200) and Alexa Fluor 488-tagged secondary antibody (1:500). The coverslips were transferred to a glass slide and mounted using ProLong glass antifade reagent. Images were acquired after 12 h with a Zeiss confocal microscope (LSM740).

**Incorporation of SERINCs into virions.** Viruses were produced by transfecting HEK293T cells using calcium phosphate transfection reagent in a 10-cm² plate, as mentioned above. The medium was replaced with Dulbecco's modified Eagle medium (DMEM) containing 2% FBS after 12 to 15 h of transfection. After 48 h, the virus-containing supernatant was collected and centrifuged at 500 × *g* for 10 min to exclude any cell debris. Following this, the viruses were filtered using a 0.22-μm syringe filter (Cole Parmer). The suspension was overlaid on a 25% sucrose (in 1× PBS) cushion and concentrated at 100,000 × *g* for 2 h at 4°C using a Beckman-Coulter ultracentrifuge. After the spin, the supernatant was aspirated, and the pellet was suspended in Laemmli buffer containing 50 mM Tris(2-carboxyethyl) phosphine hydrochloride (TCEP).

**Western blotting.** After the collection of viruses from producer cells (HEK293T), the cells were harvested in ice-cold PBS. Cells were washed twice at 500 × *g* for 5 min. The PBS was aspirated until the pellet was completely dry. The pellets were either processed for lysis or stored in a −80°C freezer until further application. The cell pellets were lysed in DDM lysis buffer (100 mM NaCl, 10 mM HEPES [pH 7.5], 50 mM TCEP, 1% *n*-dodecyl-β-D-maltoside [DDM]) and EDTA-free protease inhibitor cocktail and rocked on ice for 30 min. Following this, the lysates were clarified by centrifugation at 10,000 × *g* for 15 min, and the supernatant was collected and mixed with 4× Laemmli buffer with 50 mM TCEP.

SDS-Tricine-PAGE was used for resolving cell pellets and virions for analysis by Western blotting. The proteins from the gel were electrotransferred onto a low-fluorescence polyvinylidene difluoride (PVDF) membrane using a semidry transfer unit (TE77; Hoefer, USA) for 75 min with a 125-A constant current with the voltage set at a maximum of 20 V. The membrane containing proteins was then blocked with Odyssey blocking buffer. After 20 min of blocking, anti-HA.11 epitope tag antibody (mouse monoclonal) and anti-actin (rabbit monoclonal) (diluted 1:4,000 in blocking buffer) were used to probe the membrane to detect the HA-tagged proteins. After incubation with the primary antibody for 1 h, the membrane was washed three times with 1× Tris-buffered saline–Tween (TBS-T) for 5 min. Goat anti-mouse immunoglobulin (680RD) and goat anti-rabbit immunoglobulin (800CW; Li-Cor) antibodies were used to detect infrared dye signal from the membrane. The Western blot signal was acquired and analyzed on the Odyssey imager system (Li-Cor Biosciences) and is presented in greyscale format.

**RT-PCR.** For reverse transcription-PCR (RT-PCR), HEK293T cells were cotransfected with FV vectors (env, pol, gag, and transfer vector) and 100 ng of human SERINC5 or coelacanth SERINC2 along with pHXB2 envelope and NL4-3 Env(−) Nef(−) for production of nef-defective HIV-1. The producer cells were then collected for RNA isolation using TRIzol followed by DNase treatment to remove plasmid DNA contamination. The RNA was then concentrated using an RNA cleanup and concentrator kit (Thermo Scientific). From this, first-strand cDNA synthesis was performed using reverse transcriptase (Thermo Scientific). The cDNA was then diluted and used as a template for the amplification of FV vectors using the following primers: FV env forward, 5′-CCGGAACCCATAGTGGTGAA-3′; FV env reverse, 5′-CTTAGGCCACGGTTTGGGT-3′; FV pol forward, 5′-CTGGGATTCAGGGGCAACAA-3′; FV pol reverse, 5′-CCAAGGAACATCTGTTGGCG-3′; FV gag forward, 5′-CTTCCTGCTCCTGTACCGTC -3′; FV gag reverse, 5′-AGGAGCATTTCGTCCTAGCC-3′; FV transfer vector (pdeltaphi) forward, 5′- GACGGTATCGATGGTACCGG -3′; FV transfer vector (pdeltaphi) reverse, 5′- TTAAGAACCTTGTGTCTCTC -3′; Nef (SIVmac239) forward, 5′- ATGGGTGGAGCTATTTCCAT -3′; Nef (SIVmac239) reverse, 5′- GCGAGTTTCCTTCTTGTCAG -3′; GAPDH forward 5′- TGGAGAAGGCTGGGGCTCATTTGCA -3′; GAPDH reverse 5′- CATACCAGGAAATGAGCTTGACAA-3′.

The PCR conditions were 95°C for 2 min (initial denaturation), followed by 95°C for 30 s (final denaturation), 55°C for 30 s (annealing), and 74°C for 1 min (extension) for 30 cycles, and 74°C for 10 min (final extension).

**Genome correction and quality check.** It has been shown that the genomic sequences for many high-quality primate genomes provided in databases such as NCBI have artifactual single-base-pair-level errors (65). For example, the genome assembly sequence available in the database might have the incorrect base "A" at a particular site. However, the actual base present in the DNA of the organism might be the base "T." Such artifacts in the genomic sequence get introduced during the sequencing error correction step of the genome assembly process. Despite the various biases (66, 67) that result from such poor-quality sequences, the error-prone sequences continue to be a source of incorrect and misleading results (68).

To rule out the possibility of artifacts in the evolutionary analysis that arise from errors in the sequences of genome assemblies (true even for high-quality reference genomes found on Ensembl [68]), we systematically assessed the quality of each of the primate genomes used in this study. Genome assemblies of 15 primate species were downloaded from Ensembl release 98 through the FTP (file transfer protocol) site. Whole-genome-sequencing data sets corresponding to each of these species were obtained from the Short Read Archive (SRA) with the criterion of ≥30× coverage. Details of the genome assemblies used and the corresponding raw read data from each species are provided in Table SIII. The raw read data were mapped to the corresponding genomes using the bwa mem read mapper (69) with default settings. The alignment files obtained from the mapping step were used to generate the genotype likelihood estimates using the program angsd (70). The genotype likelihood estimates were provided to the program referee (71) to assign quality scores and perform genome correction. Overall, we found that in all the primate genomes considered, less than 1% of the bases were corrected by the

program referee. The sequencing data used for performing genome correction are not from the same individual that was used for genome assembly. Hence, it is possible that many of the corrected positions are merely nucleotide polymorphisms. These quality control steps ensured that none of the arms race signatures detected in the genes considered in our study is the result of incorrect genomic sequences.

**Manual curation and multiple-sequence alignment.** The manually curated open reading frame multi-fasta files consisting of 70 genes from ~15 primate species were collected from Ensembl. The 15 primate species from the infraorder Simiiformes used to compare the arms race signatures consists of olive baboon (*Papio anubis*), common marmoset (*Callithrix jacchus*), African green monkey (*Chlorocebus sabaeus*), gorilla (*Gorilla gorilla gorilla*), human (*Homo sapiens*), rhesus macaque (*Macaca mulatta*), crab-eating macaque (*Macaca fascicularis*), bonobo (*Pan paniscus*), chimpanzee (*Pan troglodytes*), golden snub-nosed monkey (*Rhinopithecus roxellana*), squirrel monkey (*Saimiri boliviensis boliviensis*), drill (*Mandrillus leucophaeus*), sooty mangabey (*Cercocebus atys*), Nancy Ma's night monkey (*Aotus nancymaae*), and Panamanian white-faced capuchin (*Cebus capucinus imitator*). The species name and Ensembl transcript ID for each of the genes used to look for arms race signatures are provided in Table SIV. We extended our previous multiple-sequence alignment strategies (72) by including additional multiple-sequence alignment programs to generate eight independent alignments for each gene. Several multiple-sequence alignment tools were used to ensure that the inferred patterns of sequence evolution were not restricted to the alignment strategy used. The choice of the actual multiple-sequence alignment tools used was based on the performance-based classification of algorithms (73).

**Use of FUBAR to find evolutionary fingerprints.** Traditional approaches that endeavor to find arms race signatures in genes look for the recurrent occurrence of positive selection in the same gene in different evolutionary lineages. However, when arms race signatures need to be estimated and compared for large numbers of species, such lineage-specific tests are time-consuming and result in reduced statistical power. A more recent approach has been to use the full joint distribution of synonymous ($\alpha$) and nonsynonymous ($\beta$) rates as an evolutionary fingerprint of a gene (19). The program FUBAR (74) is available as part of the HyPhy package (75). FUBAR quantifies evolutionary fingerprints of a gene by dividing the synonymous ($\alpha$) and nonsynonymous ($\beta$) rates into a finely discretized grid. Each point in this grid represents a pair of $\alpha$ and $\beta$ values, with the size of the circle representing the value $\hat{\theta}$. The value $\theta$ is the posterior mean, which provides a measure of how many of the sites in the multiple-sequence alignment contribute to that particular combination of $\alpha$ and $\beta$ values. The vector of 400 $\hat{\theta}$ values that correspond to each of the points on the 20-by-20 grid of $\alpha$ and $\beta$ values serves as the evolutionary fingerprint of a gene. Genes with canonical arms race signatures have a greater weight for grids with higher values of $\beta$ than genes that lack the arms race signature (19). The use of such signatures to quantify arms race signatures is convenient for comparing signatures of several genes. Previous comparisons of these signatures in primates have shown that the well-known arms race genes BST-2, APOBEC3F, APOBEC3G, and TRIM5-alpha have similar signatures of arms race that are missing in SERINC3 and SERINC5 (19). However, the genes SERINC1, SERINC2, and SERINC4 were not used in the previous study. To better understand the similarities and differences in the evolutionary signatures among the SERINC genes with respect to other known restriction factors, we used the hierarchical clustering of the arms race signatures of all these genes (Fig. 1C). We calculated the distance measure defined by Murrell et al. (19) to perform the hierarchical clustering of the selection signatures obtained from FUBAR.

**Rationale for using coelacanth SERINC2.** More than 95% of all extant fish species belong to the infraclass Teleostei (teleost fish), and abundant fish species that are commercially important or serve as model organisms (such as the zebrafish) belong to this group (76). These teleost fish belong to the class Actinopterygii (ray-finned fishes); they actually contain six copies of the SERINC genes and are known to have had a third round of whole-genome duplication (3R-WGD) (77). Phylogenetic studies have consistently found that the coelacanths and lungfish belong to the clade Sarcopterygii (lobe-finned fish) and share a more recent common ancestor with tetrapods than Actinopterygii (78). Importantly, the coelacanths share the two rounds of whole-genome duplication (2R-WGD) found in other tetrapods and lack the third round (3R-WGD) of whole-genome duplication found in teleost fish (77). Among the currently available high-quality genome assemblies, the coelacanth is the only organism from the clade Actinopterygii, and it has five copies of the SERINC genes. Therefore, we selected the coelacanth for our studies on SERINC2.

**Visualization of TMS-1 structure.** Most recently, the structure of pre-WGD SERINC (TMS-1) was resolved by cryo-electron microscopy (cryo-EM) and was found to consist of six monomers forming a hexamer structure (56). We downloaded this structure from PDB (Protein Data Bank; accession number 6SP2) and visualized the broad regions (Fig. 7B) corresponding to these three clusters of sequence divergence (Fig. 7A) using ChimeraX software (79). The first region, corresponding to the *D. melanogaster* TMS-1 protein sequence MPFCINSTSSYSSGALSAVSGGSLQ, is highlighted in white; the second region, with the sequence TWSAVANNPEKECNPGMFGMMEGFGNATTTAAPPTHTTRVTFDTTNIIG, is highlighted in yellow; and the third region, CISSAVEVSKISHDNSEKRDTEAGTDGSGKPSTDTETEGVTYSWS, is highlighted in green.

**GenBank accession numbers for genes used in experimental studies.** The GenBank numbers for genes used here are as follows: *Saccharomyces cerevisiae* TMS-1, NM_001180413; *Drosophila melanogaster* TMS-1, NM_140636.3; *Xenopus tropicalis* SERINC5, XM_002940195, and SERINC2, NM_001016562.2; *Gallus gallus* SERINC5, XM_424762.6, and SERINC2, NM_001030890.1; *Mus musculus* SERINC5, NM_172588.2, and SERINC2, NM_172702.3; *Equus caballus* SERINC5, XM_001503874.4, and SERINC2, XM_023634619.1; *Homo sapiens* SERINC1, NM_020755.4, SERINC2, NM_001199038.2, SERINC2 short isoform, NM_178865.5, SERINC3, NM_006811.4, SERINC4, NM_001258031.1, and SERINC5, NM_001174072.3; and coelacanth (*Latimeria chalumnae*) SERINC2, XM_005993129.2,

**Statistical analysis.** The significance of results was statistically analyzed using Student's unpaired *t* test in GraphPad Prism 9. The relative differences were considered statistically significant when *P* was <0.05.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, PDF file, 0.7 MB.
**SUPPLEMENTAL FILE 2**, MP4 file, 1.2 MB.
**SUPPLEMENTAL FILE 3**, MP4 file, 5.3 MB.

## REFERENCES

1. Simon V, Bloch N, Landau NR. 2015. Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. Nat Immunol 16:546–553. https://doi.org/10.1038/ni.3156.
2. Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3:e314. https://doi.org/10.1371/journal.pbio.0030314.
3. Kasahara M. 2010. Genome duplication and T cell immunity. Prog Mol Biol Transl Sci 92:7–36. https://doi.org/10.1016/S1877-1173(10)92002-4.
4. Malfavon-Borja R, Sawyer SL, Wu LI, Emerman M, Malik HS. 2013. An evolutionary screen highlights canonical and noncanonical candidate antiviral genes within the primate TRIM gene family. Genome Biol Evol 5:2141–2154. https://doi.org/10.1093/gbe/evt163.
5. Paparisto E, Woods MW, Coleman MD, Moghadasi SA, Kochar DS, Tom SK, Kohio HP, Gibson RM, Rohringer TJ, Hunt NR, Di Gravio EJ, Zhang JY, Tian M, Gao Y, Arts EJ, Barr SD. 2018. Evolution-guided structural and functional analyses of the HERC family reveal an ancient marine origin and determinants of antiviral activity. J Virol 92:e00528-18. https://doi.org/10.1128/JVI.00528-18.
6. Strebel K. 2013. HIV accessory proteins versus host restriction factors. Curr Opin Virol 3:692–699. https://doi.org/10.1016/j.coviro.2013.08.004.
7. Blanco-Melo D, Venkatesh S, Bieniasz PD. 2016. Origins and evolution of tetherin, an orphan antiviral gene. Cell Host Microbe 20:189–201. https://doi.org/10.1016/j.chom.2016.06.007.
8. Alcaide M, Edwards SV. 2011. Molecular evolution of the toll-like receptor multigene family in birds. Mol Biol Evol 28:1703–1715. https://doi.org/10.1093/molbev/msq351.
9. Heusinger E, Kluge SF, Kirchhoff F, Sauter D. 2015. Early vertebrate evolution of the host restriction factor tetherin. J Virol 89:12154–12165. https://doi.org/10.1128/jvi.02149-15.
10. Velová H, Gutowska-Ding MW, Burt DW, Vinkler M. 2018. Toll-like receptor evolution in birds: gene duplication, pseudogenization, and diversifying selection. Mol Biol Evol 35:2170–2184. https://doi.org/10.1093/molbev/msy119.
11. Boso G, Buckler-White A, Kozak CA. 2018. Ancient evolutionary origin and positive selection of the retroviral restriction factor Fv1 in muroid rodents. J Virol 92:e00850-18. https://doi.org/10.1128/JVI.00850-18.
12. Compton AA, Malik HS, Emerman M. 2013. Host gene evolution traces the evolutionary history of ancient primate lentiviruses. Philos Trans R Soc Lond B Biol Sci 368:20120496. https://doi.org/10.1098/rstb.2012.0496.
13. Daugherty MD, Young JM, Kerns JA, Malik HS. 2014. Rapid evolution of PARP genes suggests a broad role for ADP-ribosylation in host-virus conflicts. PLoS Genet 10:e1004403. https://doi.org/10.1371/journal.pgen.1004403.
14. Duggal NK, Emerman M. 2012. Evolutionary conflicts between viruses and restriction factors shape immunity. Nat Rev Immunol 12:687–695. https://doi.org/10.1038/nri3295.
15. Wlasiuk G, Khan S, Switzer WM, Nachman MW. 2009. A history of recurrent positive selection at the toll-like receptor 5 in primates. Mol Biol Evol 26:937–949. https://doi.org/10.1093/molbev/msp018.
16. Harris RS, Hultquist JF, Evans DT. 2012. The restriction factors of human immunodeficiency virus. J Biol Chem 287:40875–40883. https://doi.org/10.1074/jbc.R112.416925.
17. McLaren PJ, Gawanbacht A, Pyndiah N, Krapp C, Hotter D, Kluge SF, Götz N, Heilmann J, Mack K, Sauter D, Thompson D, Perreaud J, Rausell A, Munoz M, Ciuffi A, Kirchhoff F, Telenti A. 2015. Identification of potential HIV restriction factors by combining evolutionary genomic signatures with functional analyses. Retrovirology 12:41. https://doi.org/10.1186/s12977-015-0165-5.
18. Firrito C, Bertelli C, Vanzo T, Chande A, Pizzato M. 2018. SERINC5 as a new restriction factor for human immunodeficiency virus and murine leukemia virus. Annu Rev Virol 5:323–340. https://doi.org/10.1146/annurev-virology-092917-043308.
19. Murrell B, Vollbrecht T, Guatelli J, Wertheim JO. 2016. The evolutionary histories of antiretroviral proteins SERINC3 and SERINC5 do not support an evolutionary arms race in primates. J Virol 90:8085–8089. https://doi.org/10.1128/JVI.00972-16.
20. Chande A, Cuccurullo EC, Rosa A, Ziglio S, Carpenter S, Pizzato M. 2016. S2 from equine infectious anemia virus is an infectivity factor which counteracts the retroviral inhibitors SERINC5 and SERINC3. Proc Natl Acad Sci U S A 113:13197–13202. https://doi.org/10.1073/pnas.1612044113.
21. Rosa A, Chande A, Ziglio S, De Sanctis V, Bertorelli R, Goh SLSL, McCauley SMSM, Nowosielska A, Antonarakis SE, Luban J, Santoni FA, Pizzato M. 2015. HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation. Nature 526:212–217. https://doi.org/10.1038/nature15399.
22. Usami Y, Wu Y, Göttlinger HG. 2015. SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef. Nature 526:218–223. https://doi.org/10.1038/nature15400.
23. Heigele A, Kmiec D, Regensburger K, Langer S, Peiffer L, Stürzel CM, Sauter D, Peeters M, Pizzato M, Learn GH, Hahn BH, Kirchhoff F. 2016. The potency of Nef-mediated SERINC5 antagonism correlates with the prevalence of primate lentiviruses in the wild. Cell Host Microbe 20:381–391. https://doi.org/10.1016/j.chom.2016.08.004.
24. Schulte B, Selyutina A, Opp S, Herschhorn A, Sodroski JG, Pizzato M, Diaz-Griffero F. 2018. Localization to detergent-resistant membranes and HIV-

1 core entry inhibition correlate with HIV-1 restriction by SERINC5. Virology 515:52–65. https://doi.org/10.1016/j.virol.2017.12.005.

25. Sood C, Marin M, Chande A, Pizzato M, Melikyan GB. 2017. SERINC5 protein inhibits HIV-1 fusion pore formation by promoting functional inactivation of envelope glycoproteins. J Biol Chem 292:6014–6026. https://doi.org/10.1074/jbc.M117.777714.

26. Qiu X, Eke IE, Johnson SF, Ding C, Zheng Y-H. 2020. Proteasomal degradation of human SERINC4: a potent host anti-HIV-1 factor that is antagonized by nef. Curr Res Virol Sci 1:100002. https://doi.org/10.1016/j.crviro.2020.100002.

27. Zutz A, Schölz C, Schneider S, Pierini V, Münchhoff M, Sutter K, Wittmann G, Dittmer U, Draenert R, Bogner JR, Fackler OT, Keppler OT. 2020. SERINC5 is an unconventional HIV restriction factor that is upregulated during myeloid cell differentiation. J Innate Immun 12:399–409. https://doi.org/10.1159/000504888.

28. Passos V, Zillinger T, Casartelli N, Wachs AS, Xu S, Malassa A, Steppich K, Schilling H, Franz S, Todt D, Steinmann E, Sutter K, Dittmer U, Bohne J, Schwartz O, Barchet W, Goffinet C. 2019. Characterization of endogenous SERINC5 protein as anti-HIV-1 factor. J Virol 93:e01221-19. https://doi.org/10.1128/JVI.01221-19.

29. Hoffmann HH, Schneider WM, Rice CM. 2015. Interferons and viruses: an evolutionary arms race of molecular interactions. Trends Immunol 36:124–138. https://doi.org/10.1016/j.it.2015.01.004.

30. Shaw AE, Hughes J, Gu Q, Behdenna A, Singer JB, Dennis T, Orton RJ, Varela M, Gifford RJ, Wilson SJ, Palmarini M. 2017. Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. PLoS Biol 15:e2004086. https://doi.org/10.1371/journal.pbio.2004086.

31. Beitari S, Pan Q, Finzi A, Liang C. 2020. Differential pressures of SERINC5 and IFITM3 on HIV-1 envelope glycoprotein over the course of HIV-1 infection. J Virol 94:e00514-20. https://doi.org/10.1128/JVI.00514-20.

32. Kmiec D, Akbil B, Ananth S, Hotter D, Sparrer KMJ, Stürzel CM, Trautz B, Ayouba A, Peeters M, Yao Z, Stagljar I, Passos V, Zillinger T, Goffinet C, Sauter D, Fackler OT, Kirchhoff F. 2018. SIVcol Nef counteracts SERINC5 by promoting its proteasomal degradation but does not efficiently enhance HIV-1 replication in human CD4+ T cells and lymphoid tissue. PLoS Pathog 14:e1007269. https://doi.org/10.1371/journal.ppat.1007269.

33. Lynch M. 1999. The age and relationships of the major animal phyla. Evolution 53:319–325. https://doi.org/10.2307/2640769.

34. Reference deleted.

35. Ahmad I, Li S, Li R, Chai Q, Zhang L, Wang B, Yu C, Zheng YH. 2019. The retroviral accessory proteins S2, Nef, and glycoMA use similar mechanisms for antagonizing the host restriction factor SERINC5. J Biol Chem 294:7013–7024. https://doi.org/10.1074/jbc.RA119.007662.

36. Han GZ, Worobey M. 2012. An endogenous foamy-like viral element in the coelacanth genome. PLoS Pathog 8:e1002790. https://doi.org/10.1371/journal.ppat.1002790.

37. Katzourakis A, Gifford RJ, Tristem M, Thomas M, Gilbert P, Pybus OG. 2009. Macroevolution of complex retroviruses. Science 325:1512–1512. https://doi.org/10.1126/science.1174149.

38. Wei X, Chen Y, Duan G, Holmes EC, Cui J. 2019. A reptilian endogenous foamy virus sheds light on the early evolution of retroviruses. Virus Evol 5:vez001. https://doi.org/10.1093/ve/vez001.

39. Johnson WE. 2013. Rapid adversarial co-evolution of viruses and cellular restriction factors. Curr Top Microbiol Immunol 371:123–125. . https://doi.org/10.1007/978-3-642-37765-5-5.

40. McLaughlin RN, Gable JT, Wittkopp CJ, Emerman M, Malik HS. 2016. Conservation and innovation of APOBEC3A restriction functions during primate evolution. Mol Biol Evol 33:1889–1901. https://doi.org/10.1093/molbev/msw070.

41. Newman RM, Hall L, Connole M, Chen G-L, Sato S, Yuste E, Diehl W, Hunter E, Kaur A, Miller GM, Johnson WE. 2006. Balancing selection and the evolution of functional polymorphism in Old World monkey TRIM5. Proc Natl Acad Sci U S A 103:19134–19139. https://doi.org/10.1073/pnas.0605838103.

42. Sawyer SL, Wu LI, Emerman M, Malik HS. 2005. Positive selection of primate TRIM5 identifies a critical species-specific retroviral restriction domain. Proc Natl Acad Sci U S A 102:2832–2837. https://doi.org/10.1073/pnas.0409853102.

43. Browne EP, Littman DR. 2008. Species-specific restriction of Apobec3-mediated hypermutation. J Virol 82:1305–1313. https://doi.org/10.1128/JVI.01371-07.

44. Meng X, Zhang F, Yan B, Si C, Honda H, Nagamachi A, Sun LZ, Xiang Y. 2018. A paralogous pair of mammalian host restriction factors form a critical host barrier against poxvirus infection. PLoS Pathog 14:e1006884. https://doi.org/10.1371/journal.ppat.1006884.

45. Mitchell PS, Patzina C, Emerman M, Haller O, Malik HS, Kochs G. 2012. Evolution-guided identification of antiviral specificity determinants in the broadly acting interferon-induced innate immunity factor MxA. Cell Host Microbe 12:598–604. https://doi.org/10.1016/j.chom.2012.09.005.

46. Irwin B, Aye M, Baldi P, Beliakova-Bethell N, Cheng H, Dou Y, Liou W, Sandmeyer S. 2005. Retroviruses and yeast retrotransposons use overlapping sets of host genes. Genome Res 15:641–654. https://doi.org/10.1101/gr.3739005.

47. Miller JE, Zhang L, Jiang H, Li Y, Pugh BF, Reese JC. 2018. Genome-wide mapping of decay factor-mRNA interactions in yeast identifies nutrient-responsive transcripts as targets of the deadenylase Ccr4. G3 (Bethesda) 8:315–330. https://doi.org/10.1534/g3.117.300415.

48. Llorens JV, Clark JB, Martínez-Garay I, Soriano S, De Frutos R, Martínez-Sebastián MJ. 2008. Gypsy endogenous retrovirus maintains potential infectivity in several species of Drosophilids. BMC Evol Biol 8:302. https://doi.org/10.1186/1471-2148-8-302.

49. Teysset L, Burns JC, Shike H, Sullivan BL, Bucheton A, Terzian C. 1998. A Moloney murine leukemia virus-based retroviral vector pseudotyped by the insect retroviral gypsy envelope can infect Drosophila cells. J Virol 72:853–856. https://doi.org/10.1128/JVI.72.1.853-856.1998.

50. Mellor J, Fulton SM, Dobson MJ, Wilson W, Kingsman SM, Kingsman AJ. 1985. A retrovirus-like strategy for expression of a fusion protein encoded by yeast transposon Ty1. Nature 313:243–246. https://doi.org/10.1038/313243a0.

51. Aiewsakun P, Katzourakis A. 2017. Marine origin of retroviruses in the early Palaeozoic Era. Nat Commun 8:13954. https://doi.org/10.1038/ncomms13954.

52. Katzourakis A, Aiewsakun P, Jia H, Wolfe ND, LeBreton M, Yoder AD, Switzer WM. 2014. Discovery of prosimian and afrotherian foamy viruses and potential cross species transmissions amidst stable and ancient mammalian co-evolution. Retrovirology 11:61. https://doi.org/10.1186/1742-4690-11-61.

53. Shi J, Xiong R, Zhou T, Su P, Zhang X, Qiu X, Li H, Li S, Yu C, Wang B, Ding C, Smithgall TE, Zheng Y-H. 2018. HIV-1 Nef antagonizes SERINC5 restriction by downregulation of SERINC5 via the endosome/lysosome system. J Virol 92:e00196-18. https://doi.org/10.1128/JVI.00196-18.

54. Li S, Ahmad I, Shi J, Wang B, Yu C, Zhang L, Zheng Y-H. 2018. Murine leukemia virus glycosylated Gag reduces murine SERINC5 protein expression at steady-state levels via the endosome/lysosome pathway to counteract SERINC5 antiretroviral activity. J Virol 93:e01651-18. https://doi.org/10.1128/JVI.01651-18.

55. Beitari S, Ding S, Pan Q, Finzi A, Liang C. 2017. The effect of HIV-1 Env on SERINC5 antagonism. J Virol 91:e02214-16. https://doi.org/10.1128/JVI.02214-16.

56. Pye VE, Rosa A, Bertelli C, Struwe WB, Maslen SL, Corey R, Liko I, Hassall M, Mattiuzzo G, Ballandras-Colas A, Nans A, Takeuchi Y, Stansfeld PJ, Skehel JM, Robinson CV, Pizzato M, Cherepanov P. 2020. A bipartite structural organization defines the SERINC family of HIV-1 restriction factors. Nat Struct Mol Biol 27:78–83. https://doi.org/10.1038/s41594-019-0357-0.

57. Dai W, Usami Y, Wu Y, Göttlinger H. 2018. A long cytoplasmic loop governs the sensitivity of the anti-viral host protein SERINC5 to HIV-1 Nef. Cell Rep 22:869–875. https://doi.org/10.1016/j.celrep.2017.12.082.

58. Sojo V, Dessimoz C, Pomiankowski A, Lane N. 2016. Membrane proteins are dramatically less conserved than water-soluble proteins across the tree of life. Mol Biol Evol 33:2874–2884. https://doi.org/10.1093/molbev/msw164.

59. Li M, Waheed AA, Yu J, Zeng C, Chen HY, Zheng YM, Feizpour A, Reinhard BM, Gummuluru S, Lin S, Freed EO, Liu SL. 2019. TIM-mediated inhibition of HIV-1 release is antagonized by Nef but potentiated by SERINC proteins. Proc Natl Acad Sci U S A 116:5705–5714. https://doi.org/10.1073/pnas.1819475116.

60. Liu G, Zhang H, Zhao C, Zhang H. 2020. Evolutionary history of the Toll-like receptor gene family across vertebrates. Genome Biol Evol 12:3615–3634. https://doi.org/10.1093/gbe/evz266.

61. Trautz B, Wiedemann H, Luchtenborg C, Pierini V, Kranich J, Glass B, Krausslich H-G, Brocker T, Pizzato M, Ruggieri A, Brugger B, Fackler OT. 2017. The host-cell restriction factor SERINC5 restricts HIV-1 infectivity without altering the lipid composition and organization of viral particles. J Biol Chem 292:13702–13713. https://doi.org/10.1074/jbc.M117.797332.

62. Trobridge G, Josephson N, Vassilopoulos G, Mac J, Russell DW. 2002. Improved foamy virus vectors with minimal viral sequences. Mol Ther 6:321–328. https://doi.org/10.1006/mthe.2002.0672.

63. Pizzato M, Erlwein O, Bonsall D, Kaye S, Muir D, McClure MO. 2009. A one-step SYBR Green I-based product-enhanced reverse transcriptase assay for the quantitation of retroviruses in cell culture supernatants. J Virol Methods 156:1–7. https://doi.org/10.1016/j.jviromet.2008.10.012.

64. Vermeire J, Naessens E, Vanderstraeten H, Landi A, Iannucci V, van Nuffel A, Taghon T, Pizzato M, Verhasselt B. 2012. Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors. PLoS One 7:e50859. https://doi.org/10.1371/journal.pone.0050859.

65. Hubisz MJ, Lin MF, Kellis M, Siepel A. 2011. Error and error mitigation in low-coverage genome assemblies. PLoS One 6:e17034. https://doi.org/10.1371/journal.pone.0017034.

66. Schneider A, Souvorov A, Sabath N, Landan G, Gonnet GH, Graur D. 2009. Estimates of positive Darwinian selection are inflated by errors in sequencing, annotation, and alignment. Genome Biol Evol 1:114–118. https://doi.org/10.1093/gbe/evp012.

67. Mallick S, Gnerre S, Muller P, Reich D. 2009. The difficulty of avoiding false positives in genome scans for natural selection. Genome Res 19:922–933. https://doi.org/10.1101/gr.086512.108.

68. Mittal P, Jaiswal SK, Vijay N, Saxena R, Sharma VK. 2019. Comparative analysis of corrected tiger genome provides clues to its neuronal evolution. Sci Rep 9:18459. https://doi.org/10.1038/s41598-019-54838-z.

69. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv

70. Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15:356. https://doi.org/10.1186/s12859-014-0356-4.

71. Thomas GWC, Hahn MW. 2019. Referee: reference assembly quality scores. Genome Biol Evol 11:1483–1486. https://doi.org/10.1093/gbe/evz088.

72. Shinde SS, Teekas L, Sharma S, Vijay N. 2019. Signatures of relaxed selection in the CYP8B1 gene of birds and mammals. J Mol Evol 87:209–220. https://doi.org/10.1007/s00239-019-09903-6.

73. Blackburne BP, Whelan S. 2013. Class of multiple sequence alignment algorithm affects genomic analysis. Mol Biol Evol 30:642–653. https://doi.org/10.1093/molbev/mss256.

74. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K. 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol Biol Evol 30:1196–1205. https://doi.org/10.1093/molbev/mst030.

75. Pond SLK, Frost SDW, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21:676–679. https://doi.org/10.1093/bioinformatics/bti079.

76. Near TJ, Eytan RI, Dornburg A, Kuhn KL, Moore JA, Davis MP, Wainwright PC, Friedman M, Smith WL. 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proc Natl Acad Sci U S A 109:13698–13703. https://doi.org/10.1073/pnas.1206625109.

77. Singh PP, Isambert H. 2020. OHNOLOGS v2: a comprehensive resource for the genes retained from whole genome duplication in vertebrates. Nucleic Acids Res 48:D724–D730. https://doi.org/10.1093/nar/gkz909.

78. Takezaki N, Figueroa F, Zaleska-Rutczynska Z, Takahata N, Klein J. 2004. The phylogenetic relationship of tetrapod, coelacanth, and lungfish revealed by the sequences of forty-four nuclear genes. Mol Biol Evol 21:1512–1524. https://doi.org/10.1093/molbev/msh150.

79. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612. https://doi.org/10.1002/jcc.20084.

80. Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res 43:W401–W407. https://doi.org/10.1093/nar/gkv485.