New environment, new invaders - repeated horizontal transfer of LINEs to sea snakes

James D. Galbraith[1], Alastair J. Ludington[1], Alexander Suh[2,3], Kate L. Sanders[1], David L. Adelson*[1]

1) School of Biological Sciences, University of Adelaide, Adelaide, SA 5005, Australia

2) Department of Ecology and Genetics - Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, SE-752 36 Uppsala, Sweden

3) Department of Organismal Biology - Systematic Biology, Evolutionary Biology Centre, Uppsala University, SE-752 36 Uppsala, Sweden

* David L. Adelson[1]

Email: david.adelson@adelaide.edu.au

James: https://orcid.org/0000-0002-1871-2108
Alexander: https://orcid.org/0000-0002-8979-9992
Kate: https://orcid.org/0000-0002-9581-268X
David: https://orcid.org/0000-0003-2404-5636

Classification

Biological Sciences: Evolution

Keywords

horizontal transfer, transposable element, Serpentes

Author Contributions

J.D.G., A.L., A.S and D.L.A. designed research; K.L.S. and A.L. provided olive sea snake genome assembly; A.L. provided olive sea snake genome transcriptome; J.D.G. and A.L. performed research; and J.D.G., K.L.S. and D.L.A. wrote the paper with input from A.L. and A.S.

This PDF file includes:

Main Text
Figures 1 to 5
Table 1

Abstract

While numerous studies have found horizontal transposon transfer (HTT) to be widespread across metazoans, few have focused on HTT in marine ecosystems. To investigate potential recent HTTs into marine species we searched for novel repetitive elements in sea snakes, a group of elapids which transitioned to a marine habitat at most 18 Mya. Our analysis uncovered repeated HTTs into sea snakes following their marine transition. Such major shifts in habitat should require significant genomic changes.The six subfamilies of LINE retrotransposons identified in the olive sea snake (*Aipysurus laevis*) are transcribed, and hence are likely still active and expanding across the genome. A search of 600 metazoan genomes found all six were absent from other amniotes, including terrestrial elapids, with the most similar transposons present in fish and marine invertebrates. The one exception was a similar transposon found in sea kraits, a lineage of amphibious elapids which independently transitioned to a marine environment following their divergence from terrestrial species 25 Mya. Our finding of repeated horizontal transfer events into separate lineages of marine snakes greatly expands past findings of frequent horizontal transfer in the marine environment, suggesting it is ideal for the transfer of transposons.Transposons are drivers of evolution as sources of genomic sequence and hence genomic novelty. This provides evidence of the environment influencing evolution of metazoans not only through specific selection pressures, but also by contributing novel genomic material.

Significance Statement

Recent research has found horizontal transfer (HT) of transposons between marine animals. We analyzed the olive sea snake (*Aipysurus laevis*) genome, uncovering HT of six novel retrotransposons into sea snakes since their marine transition within the last 18 Mya. All six are absent from terrestrial animals and are most similar to retrotransposons found in fish, corals and the independently marine sea kraits. All six retrotransposons are likely still active and expanding across the genome in *A. laevis*. Our findings suggest the marine environment is ideal for the HT of transposons; and provide evidence that changing environments can influence evolution not only through novel selective pressures, but also by contributing novel genomic material.

Main Text

Introduction

Transposons are a major component of metazoan genomes, making up 8 to 28% of the typical amniote genome (1–4). Transposons are split into two classes: Class I containing LINEs (long interspersed elements) and LTR (long terminal repeat) retrotransposons; and Class II containing DNA transposons (5). Whilst transposons are normally vertically transmitted (parent to offspring) there have been many instances of horizontal transfer of transposons (HTT) observed between distantly related species. While HTT of DNA transposons and LTR retrotransposons appears to be more common, many examples of HTT of non-LTR retrotransposons (LINEs) have been described (6). These include transfers of RTE-BovBs between ticks and distant vertebrate lineages (7), of AviRTEs between birds and parasitic nematodes (8), and of Rex1 elements between teleost fish (9). As transposons proliferate throughout a genome they can contribute novel coding sequences, alter gene regulatory networks, modify coding regions and lead to gene copy number variation (10–13). Within a lifetime most insertions will be neutral and some may be deleterious; however, on an evolutionary time scale, some TE insertions constitute a key source of genomic innovation as organisms adapt to new and changing environments (14, 15). Hydrophiinae (Elapidae) is a prolific radiation of more than 100 terrestrial snakes plus ~70 aquatic species. The aquatic species form two separate lineages which independently transitioned to a

marine habitat: the fully marine sea snakes and the amphibious sea kraits (*Laticauda*) (16). Sea snakes are phylogenetically nested inside the terrestrial hydrophiine radiation and appeared ~6-18 Mya, while sea kraits form the sister lineage to all other Hydrophiinae and diverged 25 Mya (16, 17). Sea snakes include >60 species in two major clades, *Hydrophis* and *Aipysurus-Emydocephalus*, which shared a semi-aquatic common ancestor ~6-18 Mya and exhibit highly contrasting evolutionary histories since their transitions to a fully marine lifestyle (16, 18, 19). Both of these lineages have independently developed adaptations to the aquatic environment including valve-like nostrils allowing for full closure when underwater and tail paddles for efficient underwater movement (20). However, the *Aipysurus-Emydocephalus* lineage has continued to evolve at the same rate as terrestrial lineages of Hydrophiinae, diverging into 9 species, while the Hydrophis lineage has rapidly radiated into 48 species (21).
Following major changes in habitat ecology, such as sea snakes transition from a terrestrial to a marine habitat, organisms must adapt to their new environment, with transposons potentially contributing to adaptations (22, 23). Here we report an analysis of transposons in sea snake genomes, where the marine environment appears to have fostered the repeated, independent acquisition of these transposons through horizontal transfer of transposons (HTT). The repeated HTT suggests that direct effects of the environment on genome structure may be an important but overlooked driver of evolutionary change during major ecological transitions.

Results

Annotation of additional sea snake transposons

We previously performed ab initio repeat annotation of the olive sea snake (*Aipysurus laevis*) genome using CARP (4)and RepeatModeler (24) to compare its repetitive content to that of its terrestrial relatives *Notechis scutatus* (tiger snake) and *Pseudonaja textilis* (eastern brown snake). Most repetitive sequences identified by CARP were not well classified because both CARP and RepeatModeler rely on homology to reference sequences from Repbase (25), a database of repeats from highly studied species that are evolutionarily distant to Hydrophiinae. This reliance on sequence homology for *ab initio* repeat annotation of newly sequenced species often results in the incorrect annotation of repeats (26). Because LINEs contain endonuclease (EN) and reverse transcriptase (RT) protein domains, we used a structural homology approach based on the presence of EN and RT domains in these poorly annotated repeats. We identified four additional full-length LINE subfamilies present in the *A. laevis* genome (Figure 1) but absent from *P. textilis* and *N. scutatus*. Two of these subfamilies, Rex1-Snek_1 (five full-length copies found) and Rex1-Snek_2 (three full-length copies found) belong to the CR1/Jockey superfamily but share less than 100 bp nucleotide sequence homology. Manual viewing of a multiple sequence alignment of the five full-length copies identified by CARP revealed Rex1-Snek_1 to be three subfamilies; henceforth named Rex1-Snek_1, Rex1-Snek_3 and Rex1-Snek_4. Rex1-Snek_3 and Rex1-Snek_4 have 90% and 89% pairwise identity with Rex1-Snek_1 respectively. The other two subfamilies, RTE-Snek (three full-length sequences found) and Proto2-Snek (one full-length sequence found) belong to the RTE superfamily but have no significant nucleotide sequence homology. In addition to the full-length sequences, we identified hundreds of highly similar copies with 5' truncation patterns characteristic of recently active LINEs (Figure 2). Specifically, coverage plots of the RTE-Snek and Proto2-Snek families are typical of LINEs, with a clear pattern of 5' truncated insertions (27). All six LINE subfamilies were most similar to Repbase TE reference sequences from a marine annelid worm and teleost fishes (25) (see Table 1, SI Dataset 1).
The absence of these recently active LINE subfamilies from terrestrial snakes that diverged within the last approximately 18 Mya, combined with the finding that they were most similar to transposons from distantly related aquatic organisms, suggested HTT as a likely explanation. There are three diagnostic features of HTT: 1) the sporadic presence of a TE family within a set

of closely related species, 2) a higher than expected degree of sequence identity in long diverged species and 3) discordant topologies for the phylogenies of transposons and their host species (28).

Presence/absence in closely related species

As mentioned above, the six LINEs are absent from close terrestrial relatives of *A. laevis*. To test if the LINEs have a sporadic distribution in near relatives we performed reciprocal BLAST+ nucleotide searches for their presence in two closely related sea snake genome assemblies, *Hydrophis melanocephalus* (slender-necked sea snake) and *Emydocephalus ijimae* (Ijima's turtleheaded sea snake); the two closely related terrestrial species, *N. scutatus* and *P. textilis*; an independently aquatic species, *Laticauda colubrina* (yellow-lipped sea krait); and a distant terrestrial relative, *Ophiophagus hannah* (king cobra). The reciprocal search for RTE-Snek revealed a similar yet distinct RTE subfamily present in *L. colubrina*, henceforth referred to as RTE-Kret. From these searches, we found RTE-Snek was restricted to *A. laevis* and RTE-Kret to be restricted to *L. colubrina*. In addition to being present in *A. laevis,* Proto2-Snek was also present in *E. ijimae*, Rex1-Snek_2 in *E. ijimae* and *H. melanocephalus*, and Rex1-Snek_3 and Rex1-Snek_4 in *H. melanocephalus*. This reciprocal search confirmed all six LINEs to be absent from both terrestrial (*N. scutatus*, *P. textilis* and *O. hannah*) and aquatic (*L. colubrina*) outgroups, and RTE-Kret to be restricted to *L. colubrina* (Fig. 3, SI Fig. 3-9).

As an independent verification of presence/absence and to look for potential current activity of the LINEs, we searched transcriptomes of a variety of tissues from three sea snakes - *A. laevis, A. tenuis* and *Hydrophis major* (SI Dataset 2). We identified high-identity fragments of all four Rex1-Sneks in at least one of the *A. laevis*, *A. tenuis* and *H. major* transcriptomes. High identity fragments of the RTE-Snek and Proto2-Snek were present in *A. laevis* and *A. tenuis*, yet largely absent from *H. major*, with only one small RTE-Snek-like fragment present in a *H. major* testis transcriptome. The presence of LINE transcripts both confirmed the presence of specific LINEs and indicated potential ongoing retrotransposition of these elements. The presence/absence of all six LINEs identified in *A. laevis* across other sea snakes and their close terrestrial relatives is indicative of multiple, independent HTT events (Fig. 3).

Search for HTT donor species

In order to identify potential donor taxa for our six LINEs transferred into sea snakes, we searched for and curated similar LINEs in more than 600 metazoan genomes (SI Dataset 3). Our manual curation found homologous Rex1s in fish and squamates, Proto2s in fish, and RTEs widespread across a variety of marine organisms including fish, echinoderms, corals and sea kraits (see Fig. 4, SI Dataset 4). We then aligned our original LINE sequences against a database containing both our curated repeats and Repbase repeats. All six of our original LINEs were most similar to curated LINEs found in marine species (Table 1) with pairwise identity for all closest hits between 75-85%. Rex1-Snek_1, Rex1-Snek_2, Rex1-Snek_3 and Rex1-Snek_4 were most similar to Rex1s curated from a variety of fish genomes. Proto2-Snek was most similar to a Proto2 from the European carp (*Cyprinus carpio*) genome and RTE-Snek most similar to RTE-Kret from *L. colubrina*. However, we were unable to identify plausible donor species as none of the cross species hits was greater than 87% nucleotide sequence identity (Table 1).

Discordant phylogenies of RTEs and of Rex1s compared to host species.

As extreme discordance between repeat and species phylogenies is evidence of HTT, we compared the tree topology of all RTEs, Proto2s and Rex1s, using both Repbase sequences and our curated sequences, to the species tree topology. As illustrated in figure 5, the species and repeat phylogenies of all six sea snake LINEs and the *Laticauda* LINE are highly discordant, evidenced by their clustering with teleost fishes, confirming likely HTT events from marine organisms into sea snakes and sea kraits. The presence/absence pattern observed, illustrated in figure 3, suggests 6 to 8 transfers into sea snakes and 1 into sea kraits.

Discussion

We provide strong evidence that the six LINEs identified in *A. laevis* were horizontally transferred to sea snakes since their transition to a marine habitat, from marine species. This is based on the absence of all six LINEs from terrestrial relatives and the discordance of the species and LINE phylogenies (Figs. 3 and 5). While all six are currently expressed based on transcriptome data, the number of large, near-identical fragments of RTE-Snek and Proto2-Snek found within the *A. laevis* genome is larger than for Rex1s and indicates potentially greater replication since the original insertion events that occurred at most 12 Mya.

As all seven of the HTT LINEs are most similar to LINEs in distantly related marine metazoans, the donor species for each is likely a fish or marine invertebrate. However, the degree of sequence divergence between RTE-Kret and the six LINEs from *A. laevis* from the most similar LINEs from aquatic species means we cannot identify a donor species. As very few species with ranges overlapping that of *Laticauda* and *Aipysurus* have been sequenced and the range of *Aipysurus* spans highly biodiverse habitats, it is unlikely we will further narrow the donor of any of these six LINEs without significant additional genome sequence data from Indo-Pacific tropical marine species.

While we were unable to determine the donor species, our finding of HTT between marine species is in line with multiple past studies finding HTT within and across marine phyla. HTT is prolific and particularly well described in aquatic microbial communities (reviewed in-depth in Sobecky and Hazen, 2009 (29)). HTT of LINEs, LTR retrotransposons and DNA transposons has been reported in marine metazoans, with past studies describing the transfer of Rex1s and Rex3s between teleost fishes (9, 30), *Steamer*-like LTR retrotransposons both within and across phyla (31), and Mariner DNA transposons between diverse crustaceans (32). What sets our findings apart is that HTTs have occurred multiple times as a result of the recent terrestrial to marine transition of the *Aipysurus*/*Hydrophis* common ancestor. The transfer of all six transposons occurred <18 Mya from aquatical animal donor species that diverged from snakes >400 Mya (33, 34). As illustrated in figure 3, the varying presence/absence of the six LINEs across the three species of sea snakes is indicative of multiple HTT events as opposed to a single event. The timing of HTT into marine squamates is not specific to sea snakes, as we found transfer of an RTE-Kret to the sea kraits which underwent an independent transition to the same habitat. These repeated invasions suggest the marine environment potentially fosters HTT, with more examples likely to be revealed by additional genome sequences from marine species.

The likely ongoing replication of all 6 *A. laevis* LINEs, as evidenced by both the presence of insertions and transcripts with near 100% identity, continues to contribute genetic material to the evolution of *Aipysurus*. Previous investigators have reported entire genes, exons, regulatory sequences and noncoding RNAs in vertebrates derived from transposons, as well as TE insertions leading to genomic rearrangement (reviewed in-depth in Warren et al. (35)). For example, the insertion of CR1 fragments near phospholipase A2 venom genes in vipers led to non-allelic homologous recombination, in turn causing duplication of these genes (34). Rapid genomic innovation would have been necessary for *Aipysurus* to adapt to the marine environment, with the independent evolution of paddle-like tails, salt excretion glands and dermal photoreception following their divergence from their most recent common ancestor with *Hydrophis* (36–38). Other adaptations are likely to have occurred or are occurring for sea snakes to better adapt to their marine habitat, as evolutionary transitions from terrestrial to marine habits entail massive phenotypic changes spanning metabolic, sensory, locomotor, and communication-related traits. Future research will examine the association between HTT-derived sequences and genome regions identified as containing signatures of selection.

Conclusions

Our findings reveal repeated HTT of LINEs into a fully marine lineage following their transition from a terrestrial environment, while Australian terrestrial elapids show no evidence of HTT of LINEs. In addition, we discovered one homologous HTT event to have occurred in a lineage of snakes which has independently transitioned to a semi-aquatic environment, providing more evidence that the marine environment promotes HTT. The continued expression of all six transposons also suggests ongoing impact on the evolution of *Aipysurus*. Combined this supports a likely role for habitat transitions making a direct contribution to the evolution of metazoan genome content, rather than genomes evolving solely in response to selection imposed by changing environmental conditions. We can view the ancestral genome and novel selection from a new habitat as the two "parents" that give rise to new species, but our data indicate that HTT from the new environment may act as a "third parent", with this more likely in some habitats.

Materials and Methods

Identification and classification of repetitive sequences in *Aipysurus laevis*

We identified repetitive sequences present in the *A. laevis* assembly using CARP (4) and RepeatModeler (24). Using RPSTBLASTN 2.7.1+ (39) and the CDD and Pfam databases (40, 41) we identified protein domains present in all repetitive sequences over 2300 bp in length which had been either classified by CARP as "Chimeric", "PartitalAnnotation" or "Retrovirus_like", or classified as "Unknown" by RepeatModeler. We treated all consensus sequences containing over 80% of both an exo-endonuclease domain and a reverse transcriptase domain as potential LINEs. We used CENSOR 4.2.29 (42) to classify the consensus sequences. To reduce redundancy, we aligned all potential LINEs to all other potential LINEs using BLASTN 2.7.1+ (43, 44) with default parameters and constructed clusters based on pairwise identity (97% or higher). From each cluster the longest sequence was treated as the representative sequence.
To create a better consensus for each LINE subfamily, we manually curated new consensus sequences using a "search, extend, align, trim" method. Using the largest consensus from a subfamily we used BLASTN 2.7.1+ (43, 44) with default parameters to search for the repeat within the *A. laevis* genome. We selected the best thirty hits over 1000 bp based on bitscore and extended the coordinates of these sequences by 1000 bp at each end of the hit. We constructed multiple sequence alignments (MSAs) of the extended sequences using MAFFT v7.310 (45). Where multiple full length sequences showing significant non-homology were present the LINE subfamily was split into multiple subfamilies. Finally, we manually edited the extended sequences in Geneious Prime 2020.0.2 to remove non-homologous regions and created a new consensus sequence. If only one full length copy of a subfamily was present in the genome it was used as the consensus sequence. We used PCOILs (46) and HHpred (47) searches of the translated ORFs against the CDD and Pfam databases (40, 41) to identify any additional protein domains or structures present in the six LINEs.

Search for LINEs in closely related species genomes and transcriptomes

To determine if the six LINEs were present in closely related species we used BLASTN 2.7.1+ (43, 44) to perform a reciprocal nucleotide search of appropriate elapid snake genomes downloaded from GenBank (48). After discovering and curating RTE-Kret in *L. colubrina* the process was repeated. Similarly, we used BLASTN 2.7.1+ (43, 44) to perform reciprocal searches for the six LINEs in transcriptomes from various tissues of *A. laevis*, *A. tenuis* and *H. major* from Crowe-Riddell et al. (38).

Search for and curation of similar LINEs in other metazoan genomes

To identify other species containing the six LINEs, we used BLASTN 2.7.1+ (43, 44) to search over 600 metazoan genomes downloaded from Genbank (44) using relaxed parameters (-evalue

0.00002 -reward 3 -penalty -4 -xdrop_ungap 80 -xdrop_gap 130 -xdrop_gap_final 150 -word_size 10 -dust yes -gapopen 30 -gapextend 6). We treated species containing a hit of at least 1000 bp as potentially containing a similar LINE. In species potentially containing a similar LINE we attempted to manually curate the LINE using a variant of the "search, extend, align, trim" method described above. We used a consensus of the initial hits within the species as the query for the BLASTN search of the genome, and extended hits by 3000 bp in each direction. If an MSA appeared to contain multiple LINE families, the MSA was split into the families and consensuses constructed for each individual family. To identify LINEs in *A. laevis* similar to those identified in other elapid snakes we used the same approach as above, using LINEs curated from *N. scutatus* as the query. The same method was used to curate RTE-Kret from the *L. colubrina* assembly.

Characterising divergence patterns in the HT repeats across Hydrophiinae
To identify fragments of the six HTT LINEs identified in *A. laevis* and determine their divergence from the consensus sequences we performed a reciprocal best hit search using BLASTN 2.7.1+ (43, 44) on the *A. laevis, E. ijimae, H. cyanocinctus, H. melanocephalus, N. scutatus, P. textilis, L. colubrina* and *O. hannah* assemblies . Following the identification and curation of RTE-Kret this was repeated.

Repeat phylogeny construction
For constructing repeat phylogenies we created two libraries; one containing all curated and Repbase Rex1s and another containing all curated and Repbase RTE-like (Proto2, RTE and BovB) LINEs. In addition, each library contained an outgroup LINE based on the Eickbush and Malik (49) phylogeny of LINEs. We removed all sequences not containing at least 80% of both the endonuclease and reverse transcriptase domains from each library based on RPSBLAST (39) searches against the NCBI CDD (40).
We created MSAs of each library of LINEs using MAFFT (43) and removed poorly aligned regions using Gblocks (50). Finally we constructed phylogenies from the trimmed MSA using RAxML (51) with 20 maximum likelihood trees and 500 bootstraps.

Species phylogeny construction
We used TimeTree (52) to infer species phylogenies presented in Figure 4. In cases in which a species of interest was not present in the TimeTree database, where possible we used an appropriate species from the same clade in its place and corrected the species names on the resulting tree.

Acknowledgments

We thank Dan Kortschak (University of Adelaide), Jesper Boman (Uppsala University), Valentina Poena (Uppsala University) and Atma Ivancevic (University of Colorado) for helpful suggestions and discussions.

References

1. R. N. Platt, M. W. Vandewege, D. A. Ray, Mammalian transposable elements and their impacts on genome evolution. *Chromosome Res.* **26**, 25–43 (2018).
2. G. Zhang, *et al.*, Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320 (2014).
3. R. E. Green, *et al.*, Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science* **346** (2014).
4. L. Zeng, R. D. Kortschak, J. M. Raison, T. Bertozzi, D. L. Adelson, Superior ab initio identification, annotation and characterisation of TEs and segmental duplications from

genome assemblies. *PLOS ONE* **13**, e0193588 (2018).
5. T. Wicker, *et al.*, A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **8**, 973–982 (2007).
6. J. Peccoud, R. Cordaux, C. Gilbert, Analyzing Horizontal Transfer of Transposable Elements on a Large Scale: Challenges and Prospects. *BioEssays* **40**, 1700177 (2018).
7. A. M. Ivancevic, R. D. Kortschak, T. Bertozzi, D. L. Adelson, Horizontal transfer of BovB and L1 retrotransposons in eukaryotes. *Genome Biol.* **19**, 85 (2018).
8. A. Suh, *et al.*, Ancient horizontal transfers of retrotransposons between birds and ancestors of human pathogenic nematodes. *Nat. Commun.* **7**, 11396 (2016).
9. J.-N. Volff, C. Körting, M. Schartl, Multiple lineages of the non-LTR retrotransposon Rex1 with varying success in invading fish genomes. *Mol. Biol. Evol.* **17**, 1673–1684 (2000).
10. R. Rebollo, M. T. Romanish, D. L. Mager, Transposable Elements: An Abundant and Natural Source of Regulatory Sequences for Host Genes. *Annu. Rev. Genet.* **46**, 21–42 (2012).
11. E. B. Chuong, N. C. Elde, C. Feschotte, Regulatory activities of transposable elements: From conflicts to benefits. *Nat. Rev. Genet.* **18**, 71–86 (2017).
12. S. Cerbin, N. Jiang, Duplication of host genes by transposable elements. *Curr. Opin. Genet. Dev.* **49**, 63–69 (2018).
13. L. Schrader, J. Schmitz, The impact of transposable elements in adaptive evolution. *Mol. Ecol.* **28**, 1537–1549 (2019).
14. E. Casacuberta, J. González, The impact of transposable elements in environmental adaptation. *Mol. Ecol.* **22**, 1503–1517 (2013).
15. J. Salces-Ortiz, C. Vargas-Chavez, L. Guio, G. E. Rech, J. González, Transposable elements contribute to the genomic response to insecticides in Drosophila melanogaster. *Philos. Trans. R. Soc. B Biol. Sci.* **375**, 20190341 (2020).
16. M. S. Y. Lee, K. L. Sanders, B. King, A. Palci, Diversification rates and phenotypic evolution in venomous snakes (Elapidae). *R. Soc. Open Sci.* (2016) https:/doi.org/10.1098/RSOS.150277 (May 6, 2019).
17. K. L. Sanders, M. S. Y. Lee, R. Leys, R. Foster, J. Scott Keogh, Molecular phylogeny and divergence dates for Australasian elapids and sea snakes (hydrophiinae): Evidence from seven genes for rapid evolutionary radiations. *J. Evol. Biol.* **21**, 682–695 (2008).
18. K. L. Sanders, M. S. Y. Lee, Mumpuni, T. Bertozzi, A. R. Rasmussen, Multilocus phylogeny and recent rapid radiation of the viviparous sea snakes (Elapidae: Hydrophiinae). *Mol. Phylogenet. Evol.* **66**, 575–591 (2013).
19. C. R. Nitschke, M. Hourston, V. Udyawer, K. L. Sanders, Rates of population differentiation and speciation are decoupled in sea snakes. *Biol. Lett.* **14**, 20180563 (2018).
20. H. B. Lillywhite, *How snakes work : structure, function and behavior of the world's snakes* (Oxford University Press, 2014) (June 10, 2019).
21. K. L. Sanders, Mumpuni, M. S. Y. Lee, Uncoupling ecological innovation and speciation in sea snakes (Elapidae, Hydrophiinae, Hydrophiini). *J. Evol. Biol.* **23**, 2685–2693 (2010).
22. C. Schlötterer, R. Kofler, E. Versace, R. Tobler, S. U. Franssen, Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity* **114**, 431–440 (2015).
23. D. A. Marques, F. C. Jones, F. Di Palma, D. M. Kingsley, T. E. Reimchen, Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. *Nat. Ecol. Evol.* **2**, 1128–1138 (2018).
24. A. Smit, R. Hubley, *RepeatModeler*.
25. W. Bao, K. K. Kojima, O. Kohany, Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **6**, 11 (2015).
26. R. N. Platt, L. Blanco-Berdugo, D. A. Ray, Accurate Transposable Element Annotation Is Vital When Analyzing New Genome Assemblies. *Genome Biol. Evol.* **8**, 403–410 (2016).
27. D. D. Luan, M. H. Korman, J. L. Jakubczak, T. H. Eickbush, Reverse transcription of R2Bm

RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).
28. J. C. Silva, E. L. Loreto, J. B. Clark, Factors that Affect the Horizontal Transfer of Transposable Elements. *Curr. Issues Mol. Biol.* **6**, 57–72 (2004).
29. P. A. Sobecky, T. H. Hazen, "Horizontal Gene Transfer and Mobile Genetic Elements in Marine Systems" in *Horizontal Gene Transfer: Genomes in Flux*, Methods in Molecular Biology., M. B. Gogarten, J. P. Gogarten, L. C. Olendzenski, Eds. (Humana Press, 2009), pp. 435–453.
30. F. Carducci, *et al.*, Rex Retroelements and Teleost Genomes: An Overview. *Int. J. Mol. Sci.* **19**, 3653 (2018).
31. M. J. Metzger, A. N. Paynter, M. E. Siddall, S. P. Goff, Horizontal transfer of retrotransposons between bivalves and other aquatic species of multiple phyla. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E4227–E4235 (2018).
32. N. Casse, *et al.*, Species sympatry and horizontal transfers of Mariner transposons in marine crustacean genomes. *Mol. Phylogenet. Evol.* **40**, 609–619 (2006).
33. R. E. Broughton, R. Betancur-R, C. Li, G. Arratia, G. Ortí, Multi-locus phylogenetic analysis reveals the pattern and tempo of bony fish evolution. *PLOS Curr. Tree Life* (2013) https:/doi.org/10.1371/currents.tol.2ca8041495ffafd0c92756e75247483e (January 7, 2020).
34. L. C. Hughes, *et al.*, Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proc. Natl. Acad. Sci.* **115**, 6249–6254 (2018).
35. I. A. Warren, *et al.*, Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. *Chromosome Res.* **23**, 505–531 (2015).
36. K. L. Sanders, A. R. Rasmussen, J. Elmberg, Independent Innovation in the Evolution of Paddle-Shaped Tails in Viviparous Sea Snakes (Elapidae: Hydrophiinae). *Integr. Comp. Biol.* **52**, 311–320 (2012).
37. F. Brischoux, R. Tingley, R. Shine, H. B. Lillywhite, Salinity influences the distribution of marine snakes: implications for evolutionary transitions to marine life. *Ecography* **35**, 994–1003 (2012).
38. J. M. Crowe-Riddell, *et al.*, Phototactic tails: Evolution and molecular basis of a novel sensory trait in sea snakes. *Mol. Ecol.* **28**, 2013–2028 (2019).
39. A. Marchler-Bauer, S. H. Bryant, CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* **32**, W327–W331 (2004).
40. A. Marchler-Bauer, *et al.*, CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* **45**, D200–D203 (2017).
41. R. D. Finn, *et al.*, The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).
42. O. Kohany, A. J. Gentles, L. Hankus, J. Jurka, Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* **7**, 474 (2006).
43. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
44. C. Camacho, *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
45. K. Katoh, D. M. Standley, MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
46. M. Gruber, J. Söding, A. N. Lupas, Comparative analysis of coiled-coil prediction methods. *J. Struct. Biol.* **155**, 140–145 (2006).
47. L. Zimmermann, *et al.*, A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J. Mol. Biol.* **430**, 2237–2243 (2018).
48. D. A. Benson, *et al.*, GenBank. *Nucleic Acids Res.* **45**, D37–D42 (2017).
49. T. H. Eickbush, H. S. Malik, "Origins and evolution of retrotransposons" in *Mobile DNA II*, N. L. Craig, R. Craigie, M. Gellert, A. M. Lambowitz, Eds. (American Society of Microbiology

Press, 2002), pp. 1111–1144.
50. G. Talavera, J. Castresana, Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Syst. Biol.* **56**, 564–577 (2007).
51. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
52. S. B. Hedges, J. Dudley, S. Kumar, TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* **22**, 2971–2972 (2006).
53. H. Wickham, ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **3**, 180–185 (2011).
54. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).

Figures and Tables

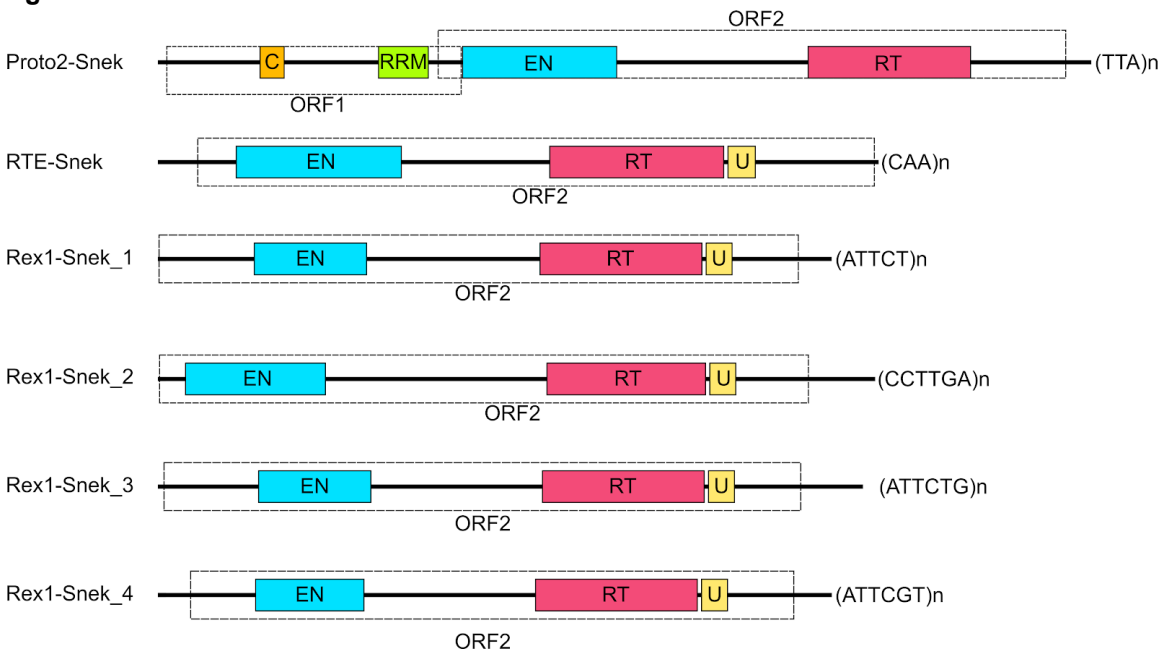

Figure 1. Structure of the seven horizontally transferred LINEs - Proto2-Snek, Rex1-Snek_1, Rex1-Snek_2, Rex1-Snek_3, Rex1-Snek_4 and RTE-Snek. Cyan represents endonuclease (EN), red reverse transcriptase (RT), orange coiled coil (CC), green RNA-recognition motif (RRM), and yellow domain of unknown function 1891 (U). Protein domains were identified using RPSBLAST (40) and HHpred (47) searches against CDD and Pfam (40, 41) databases and the coiled coil domain was identified using PCOILS (46).

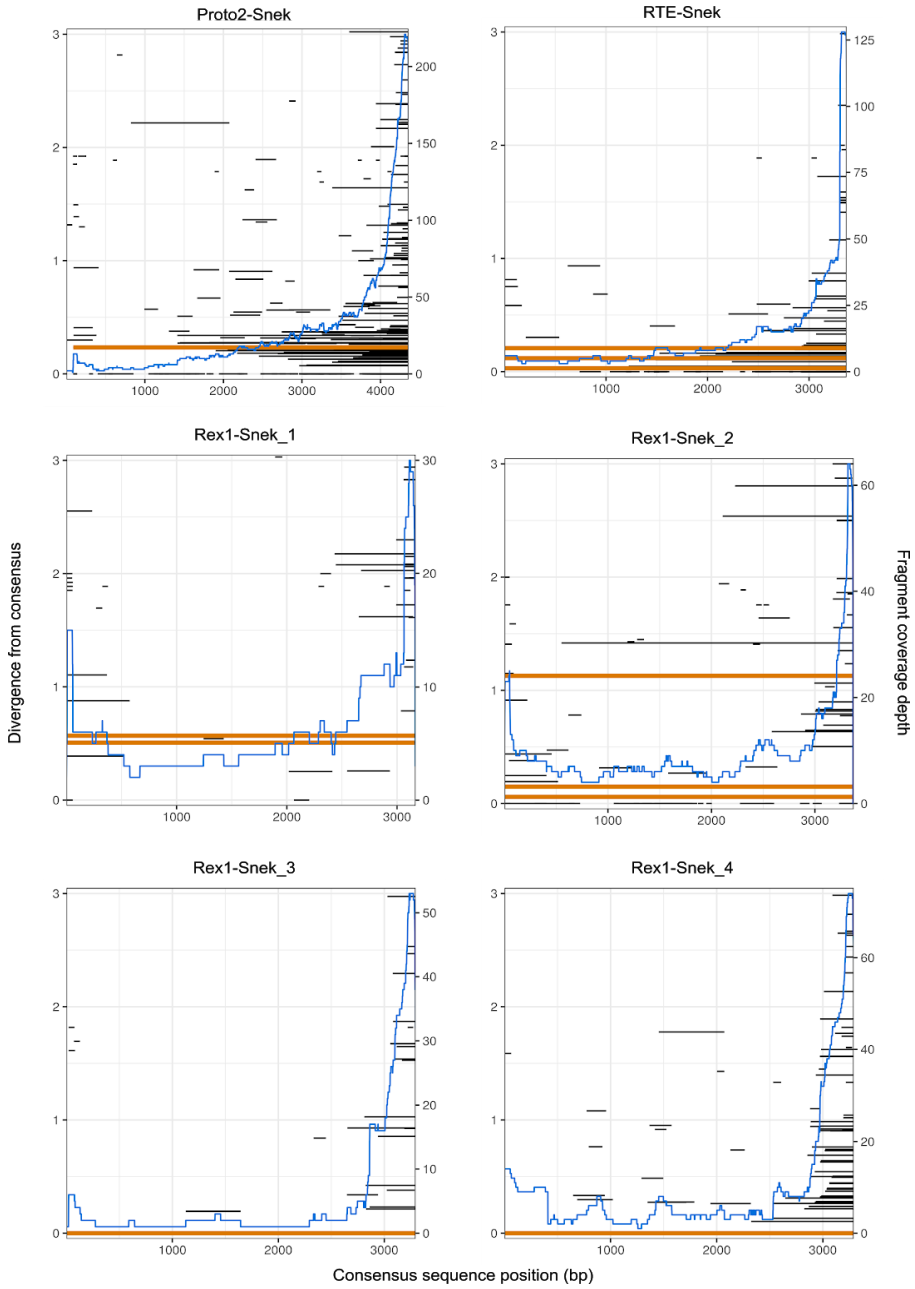

Figure 2. Coverage and divergence from consensus of the six additional LINEs identified in the *Aipysurus laevis* genome. LINE fragments were identified with BLASTN (43, 44) and plotted using ggplot2 (53) using the consensus2genome script (https://github.com/clemgoub/consensus2genome). The blue line represents the depth of coverage of fragments aligned to the subfamily consensus sequence (shown on right hand Y-axis). Each horizontal line represents the divergence of a fragment and its position mapped to the repeat consensus (position shown on X-axis); red shows full length repeats and black shows repeat fragments. The divergence from consensus of the repeats is shown on left hand Y-axis.

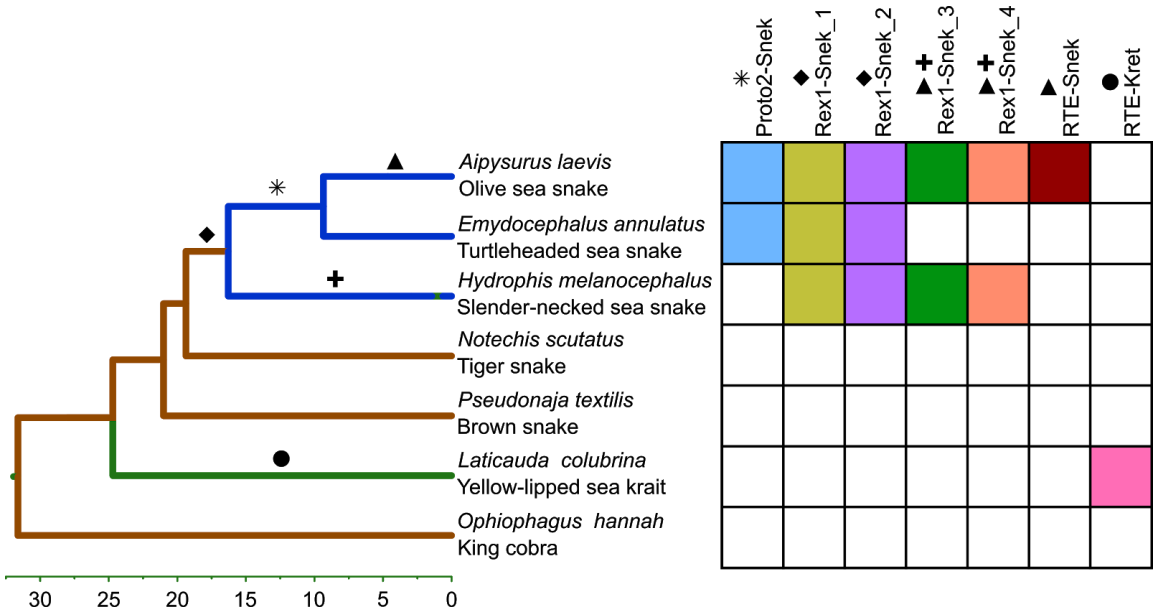

Figure 3. Presence of the seven HTT LINE subfamilies across the phylogeny of elapid snakes (adapted from Lee et al, 2016 (16)). Colour of lineage represents habitat - marine species are blue, terrestrial brown and amphibious green. Each symbol represents the likely timing of horizontal transfers. Presence/absence determined using reciprocal BLASTN search (43, 44) using default parameters.

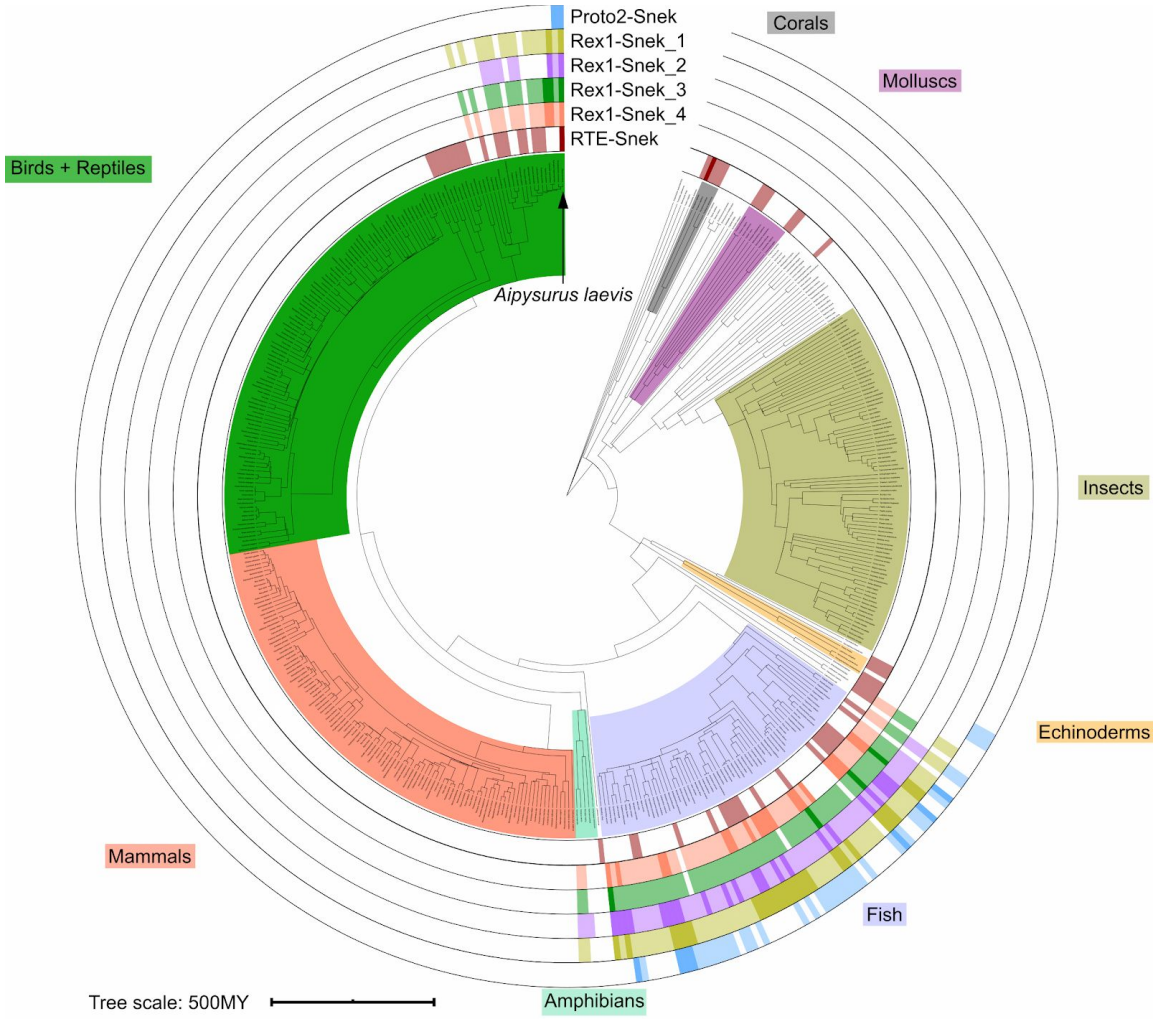

Figure 4. Presence of the six HTT LINEs across 458 Metazoa. In each ring darker shading represents the presence of at least one sequence over 1000 bp in length showing 75% or higher pairwise identity to the LINE, lighter shading represents the presence of at least one sequence over 1000 bp with less than 75% pairwise identity, and white represents the complete absence of similar sequences. Presence of LINEs identified using BLASTN (43, 44) and plotted in iToL (54). Species tree generated using TimeTree (52), manually edited to correct elapid phylogeny to fit (16). Interactive tree available at https://itol.embl.de/shared/jamesdgalbraith.

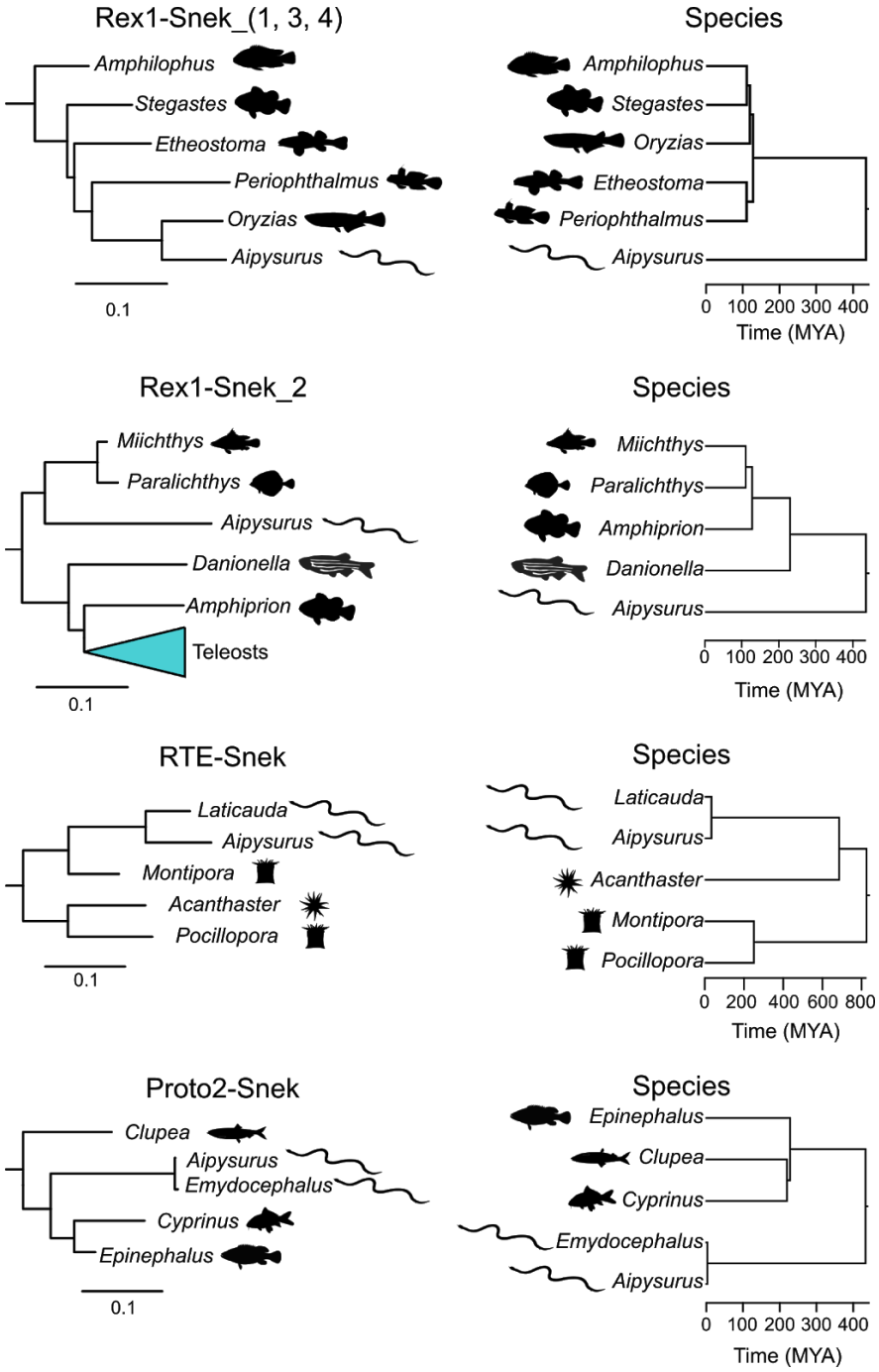

Figure 5. Excerpts from phylogenies of all intact RTEs, Proto2-Snek-like RTEs and RTE-Snek-like RTEs compared to host species phylogeny. The blue triangle on the left represents a condensed, very large subtree of Rex1 sequences. TE phylogeny scale bar represents substitutions per site. RTE, Proto2 and Rex1 phylogenies are extracts from larger phylogenies constructed using RAxML (51) based on MAFFT (45) alignments trimmed with Gblocks (50) (for full phylogenies see SI Appendix, Fig. S1 and Fig. S2). Species trees constructed with TimeTree (52).

Most similar RepBase sequences			
Repeat (query)	**Species (target repeat)**	**Percent identity**	**Hit length (bp)**
Rex-Snek_1	*Petromyzon marinus* (Rex1-1_PM)	67.5	1,359
Rex-Snek_2	*Cyprinus carpio* (Rex1-1_CCa)	75.9	2,795
Rex-Snek_3	*Petromyzon marinus* (Rex1-1_PM)	66.7	1,359
Rex-Snek_4	*Petromyzon marinus* (Rex1-1_PM)	64.2	2,796
RTE-Snek	*Petromyzon marinus* (RTE-2_PM)	62.9	3,100
Proto2-Snek	*Oryzias latipes* (Proto2-1_OL)	65.6	666
Most similar curated repeats			
Rex-Snek_1	*Oryzias latipes*	85.3	2,991
Rex-Snek_2	*Cyprinus carpio*	76.9	2,758
Rex-Snek_3	*Oryzias latipes*	82.5	2,980
Rex-Snek_4	*Oryzias latipes*	82.2	2,976
RTE-Snek	*Laticauda colubrina*	84.8	3,235
Proto2-Snek	*Cyprinus carpio*	75.9	2,752

Table 1.
Most similar RepBase and curated repeats for each repeat family in species outside of snakes. RepBase was searched using the six consensus *Aipysurus laevis* repeats using relaxed BLASTN parameters (see Methods). A database of our curated repeats from all searched species (see Methods) was searched using the six consensus *Aipysurus laevis* repeats using default BLASTN parameters.