



Phylogenetic relatedness rather than aquatic habitat fosters horizontal transfer of transposable elements in animals

Héloïse Muller, Rosina Savisaar, Jean Peccoud, et al.

Genome Res. 2025 35: 2011-2022 originally published online August 8, 2025

Access the most recent version at doi:[10.1101/gr.280432.125](https://doi.org/10.1101/gr.280432.125)

References This article cites 67 articles, 8 of which can be accessed free at:
<http://genome.cshlp.org/content/35/9/2011.full.html#ref-list-1>

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

© 2025 Muller et al.; Published by Cold Spring Harbor Laboratory Press

Research

Phylogenetic relatedness rather than aquatic habitat fosters horizontal transfer of transposable elements in animals

Héloïse Muller,^{1,2} Rosina Savisaar,³ Jean Peccoud,⁴ Sylvain Charlat,^{5,6,7} and Clément Gilbert^{1,7}

¹Université Paris-Saclay, CNRS, IRD, UMR Évolution, Génomes, Comportement et Écologie, 91198 Gif-sur-Yvette, France;

²Department of Biological Sciences, University at Albany, State University of New York, Albany, New York 12222, USA; ³Mondego Science, 11000 Carcassonne, France; ⁴UMR CNRS 7267 Écologie et Biologie des Interactions, Équipe Écologie Évolution Symbiose, Université de Poitiers, 86073 Poitiers, France; ⁵Laboratoire de Biométrie et Biologie Évolutive, Université de Lyon, Université Lyon 1, CNRS, UMR 5558, 69622 Villeurbanne, France; ⁶Laboratoire PIMIT (Processus Infectieux en Milieu Insulaire Tropical), CNRS, UMR 9192, 97490 Sainte Clotilde, La Réunion, France

Horizontal transfer of transposable elements (HTT) is an important driver of genome evolution, yet the factors conditioning this phenomenon remain poorly characterized. Here, we screen 247 animal genomes from four phyla (annelids, arthropods, mollusks, chordates), spanning 19 independent transitions between aquatic and terrestrial lifestyles, to evaluate the suspected positive effects of aquatic habitat and of phylogenetic relatedness on HTT. Among the 6043 independent HTT events recovered, the vast majority (>85%) involve DNA transposons, of which Mariner-like and hAT-like elements have the highest rates of horizontal transfer and of intragenomic amplification. Using a novel approach that circumvents putative biases linked to phylogenetic inertia and taxon sampling, we find that HTT rates positively correlate with similarity in habitat type but are not significantly higher in aquatic than in terrestrial animals. However, modeling the number of HTT events as a function of divergence time in a Bayesian framework reveals a clear positive effect of phylogenetic relatedness on HTT rates in most of the animal species studied (162 out of 247). The effect is very pronounced: A typical species is expected to show 10 times more transfers with a species it diverged from 250 million years (My) ago than with a species it diverged from 650 My ago. Overall, our study underscores the pervasiveness of HTT throughout animals and the impact of evolutionary relatedness on its dynamics.

[Supplemental material is available for this article.]

Transposable elements (TEs), like any other genome component, are primarily transmitted vertically, from parents to their descendants through reproduction. The ability of TEs to move and multiply within genomes also makes them particularly prone to horizontal transfer (Silva et al. 2004). Horizontal transfer of TEs (HTT) allows these elements to invade new genomes, enhancing their persistence over evolutionary time (Schaack et al. 2010). This process, therefore, deeply influences genome evolution, given that TEs compose a major fraction of the chromosomes of most eukaryotic taxa (Gilbert and Feschotte 2018). Whereas their evolutionary impacts are well identified, the mechanisms by which TEs circulate between organisms—whether they use vectors such as viruses and extracellular vesicles (Gilbert and Cordaux 2017; Ono et al. 2019; Widen et al. 2023) or transfer more directly during parasitism or predation (Dupuy et al. 2011; Kambayashi et al. 2022; Heisserer et al. 2023)—remain mostly open to speculation. This limitation stems from the lack of direct observations of ongoing horizontal transfers, which are only inferred through the traces they leave in genomes. Such traces have been continuously reported for more than 30 years, and their analysis revealed horizontal transfers among closely and distantly related species of fungi,

plants, and animals (Daniels et al. 1990; Bartolomé et al. 2009; Gilbert et al. 2010; Wallau et al. 2012; Scarpa et al. 2024). These investigations, however, did not recover enough transfer events to clarify the evolutionary and ecological factors governing HTT. Hence, they also could not elucidate its causes and underlying mechanisms (Carvalho et al. 2023).

Recent large-scale studies using comparative genomics have revealed that HTT is pervasive in flowering plants (Baidouri et al. 2014), insects (Wallau et al. 2016; Peccoud et al. 2017; Reiss et al. 2019; de Melo and Wallau 2020), and several vertebrate groups (Ivancevic et al. 2018; Zhang et al. 2020; Paulat et al. 2023). Among the trends that emerged from the hundreds to thousands of transfers reported in each analysis, a global positive effect of phylogenetic relatedness on HTT rates was found in insects and plants (Bartolomé et al. 2009; Baidouri et al. 2014; Peccoud et al. 2017; de Melo and Wallau 2020). In other words, a higher number of HTT events were found between closely related species than between more distant ones. Shared biogeographical origin was also found to correlate with a higher rate of recent HTT in insects overall (Peccoud et al. 2017) (but not among mosquitoes [de Melo and Wallau 2020]). Finally, a strong effect of the taxon hosting TEs was

⁷These authors contributed equally to this work.

Corresponding author: clement.gilbert@cnrs.fr

Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.280432.125>.

© 2025 Muller et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <https://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

shown in two studies (Reiss et al. 2019; Zhang et al. 2020), which identified lepidopterans and ray-finned fishes as hotspots of HTT among insects and vertebrates, respectively. The large excess of HTT found in ray-finned fishes (Zhang et al. 2020), as well as multiple individual reports of HTT in other aquatic organisms, have led to the suggestion that TEs may more frequently transfer in aquatic than in terrestrial environments (Wang and Liu 2016; Metzger et al. 2018; Galbraith et al. 2020; Hassan et al. 2024). The hypothesis of a positive effect of the aquatic medium on HTT has also been put forth to explain the over-representation of introns derived from intron TE in aquatic species (Gozashti et al. 2022, 2025). The idea that aquatic environments may foster horizontal transfer stems from the observation that it contains large amounts of circulating DNA, either free or in membrane vesicles, that may be readily available for uptake by many organisms sharing the same body of water (Abe et al. 2020; Pérez-Etayo et al. 2020; Brann et al. 2024). Beyond these hints, the effect of aquatic environment, or of any other ecological trait, on HTT has not yet been formally tested. Likewise, the influence of phylogenetic relatedness has not yet been quantitatively assessed in terms of the strength of the effect or its prevalence across species. These shortcomings partly reflect the difficulty of statistically analyzing HTT events in large data sets due to the biases related to taxon sampling. As a horizontal transfer is generally inferred from the sharing of similar TEs by several species (Peccoud et al. 2018), its detection depends on the species selection (a transfer may not be called if too few species share similar TEs). The data set composition may also affect estimates of correlations between observed transfer rates and ecological factors, as the choice of species defines the network of possible ecological interactions. An evolutionary analysis of phenotypic/ecological traits also faces the nonindependence of trait states among related species, referred to as phylogenetic inertia.

Here, we tackled these difficulties by developing a new approach to investigate the global impact of two factors on HTT at the scale of the metazoan phylogeny. First, we specifically tested whether aquatic habitats are more conducive to HTT than terrestrial habitats. To do so, we inferred HTT events among genomes from aquatic and terrestrial metazoan species spanning multiple independent habitat transitions, accounting for possible correlations between habitat and other traits along the host phylogeny. Secondly, we precisely characterized the influence of phylogenetic relatedness on horizontal transfers. To this end, we developed Bayesian models to quantify this effect for each species separately and assessed its variation across animals.

Results

Massive horizontal transfer of transposable elements in metazoans

Testing the effect of factors like habitat or phylogenetic relatedness on HTT rates while accounting for phylogenetic inertia requires a careful selection of taxa that cover multiple independent state changes over large phylogenetic distances. To fulfill these requirements, we selected genome assemblies from 126 terrestrial and 121 aquatic species (either fully or semiaquatic) distributed across 19 taxonomic groups, each of which has undergone at least one habitat transition during the evolution of animals (Supplemental Dataset S1). These groups include the transition from water to land at the origin of amniotes, five secondary transitions from land to water among amniotes, transitions from water to land in amphibians, the transition from land to water in hemipteran in-

sects, at least seven transitions from land to semiaquatic lifestyle in insects, and transitions from water to land in crustaceans, chelicerates, gastropods, and annelids.

We identified TEs in all species via a de novo discovery using the RepeatModeler pipeline (Flynn et al. 2020). Consensus sequences of TE families generated by this step were gathered with those from Repbase (Bao et al. 2015) and from Peccoud et al. (2017) and Zhang et al. (2020) into a shared library (284,643 sequences longer than 300 base pairs [bp]), which was used to annotate individual TE copies within each genome via the RepeatMasker program (Fig. 1; Smit et al. 2013). This search yielded a total of 111,102,018 TE copies (or fragments of copies) larger than 300 bp in the 247 genomes (Figs. 1, 2A). As expected (Peccoud et al. 2017; Zhang et al. 2020), Class 1 TEs (retrotransposons) proved the most abundant in the analyzed genomes: 56.3% of the TE copies included in the present study are LINE elements and 21.4% are LTR elements. Class 2 elements (DNA transposons) make up 22.3% of the total number of copies included in the present study.

We then identified TEs that have horizontally transferred between species by implementing an established criterion

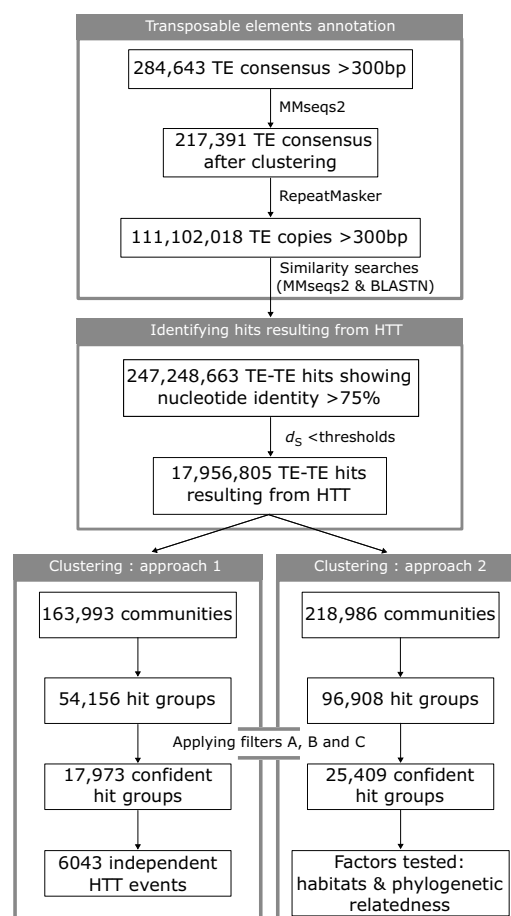


Figure 1. Graphical summary of the method used to identify HTT. A more detailed figure detailing all steps of the method can be found in Supplemental Figure S1. Two independent clustering approaches were performed. Approach 1 was applied in a two-step procedure as described in Zhang et al. (2020). Approach 2 was applied on each species pair separately and comprised only one step. Filters A, B, and C were applied to avoid false positives; they are described in Supplemental Figure S2 and in Supplemental Methods.

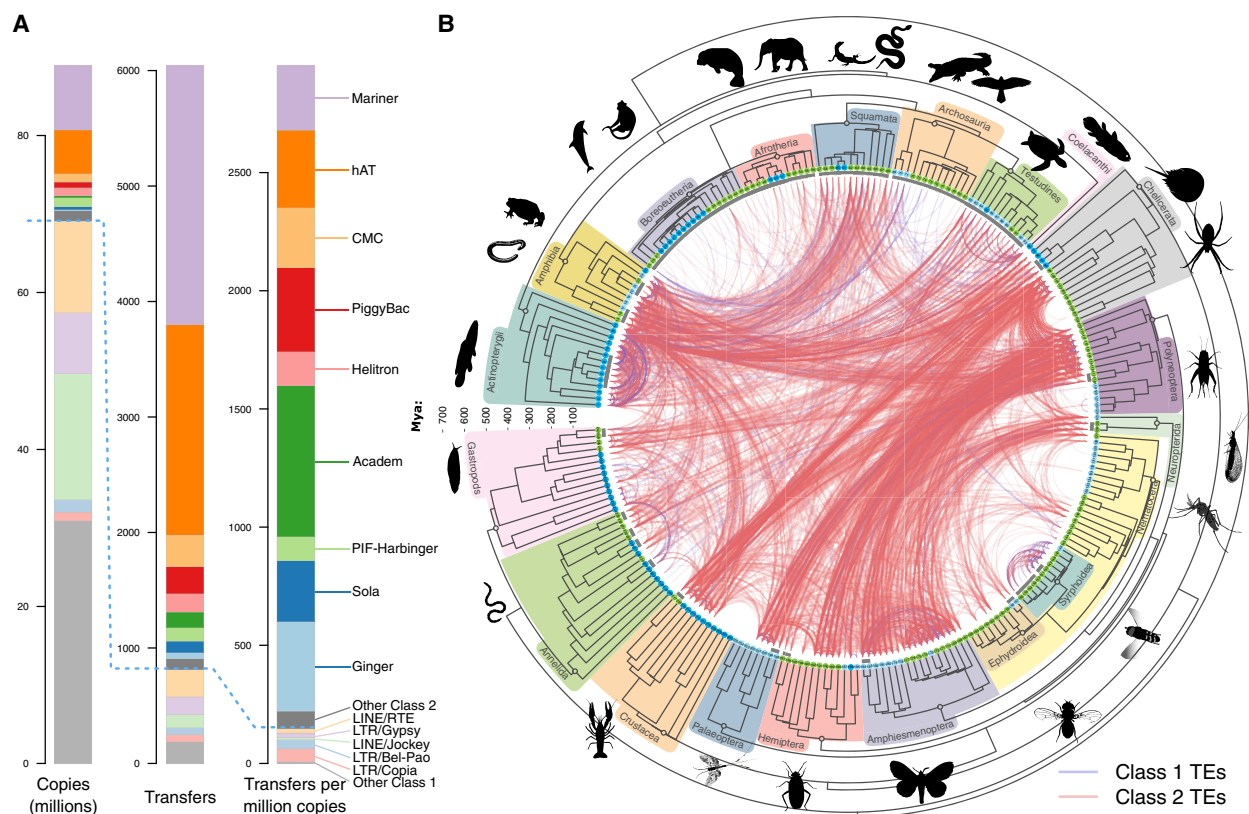


Figure 2. Horizontal transfer of transposable elements in animals. (A) Statistics on TE superfamilies. Bar plots show the total number of copies of each TE superfamily, the total number of independent transfers recovered for each superfamily, and the total number of independent transfers per million copies (see also Supplemental Fig. S3A). Superfamilies above the blue dashed line comprise Class 2 elements (DNA transposons); those below comprise Class 1 elements (retrotransposons). Superfamilies involved in fewer than 50 transfer events are grouped in the “Other” category for each class. (B) Detected horizontal transfer events. The concave time tree represents the phylogeny of the 247 species analyzed (tree in Newick format in Supplemental Dataset S2). Species are numbered on tree tips according to Supplemental Dataset S1. Species living in terrestrial, semiaquatic, or fully aquatic habitat have their numbers circled in green, light blue, and dark blue, respectively. The 19 taxonomic groups are highlighted in different colors (silhouettes of representative species were obtained from PhyloPic.org and cleanpng.com). Concave gray lines at the branch tips encompass “species units” within which HTT was not studied (due to the difficulty of distinguishing HTT from vertical inheritance of TEs; see Supplemental Methods; Supplemental Fig. S2A). Each curve represents one of the 6043 independent HTT events recovered and connects the two species involved in the hit of the highest sequence identity in the transfer. Blue curves represent HTT of Class 1 TEs and red curves HTT of Class 2 TEs.

(Bartolomé et al. 2009), according to which transferred elements should show greater sequence similarity than would be expected if they were vertically inherited from the species’ last common ancestor (see Methods). To this aim, we extracted TE copies from all 247 genomes and performed reciprocal similarity searches between these copies within the 30,313 pairs of species sharing a last common ancestor older than 40 million years, that is, ignoring transfers (of any age) between closely related species. Species that diverged after that time were considered too closely related for a confident distinction between horizontal transfer and vertical inheritance from the species’ last common ancestor. Among the resulting hits between TE copies (hereafter referred to as “TE-TE hits”), we retained those showing $\geq 75\%$ nucleotide sequence identity (Fig. 1). Focusing on highly similar TE-TE hits ensures that we recovered recent HTT events only, which would have mostly occurred after the evolutionary transitions between aquatic and terrestrial habitats. Therefore, the habitat type of the analyzed species may be extrapolated to the ancestors or relatives that emitted or acquired transferred TEs. Among these TE-TE hits, we then selected those involving TE copies that were more similar than expected under the hypothesis of strict vertical inheritance. Following earli-

er studies (Reiss et al. 2019; Zhang et al. 2020), we compared the synonymous distance (d_s) between TE copies to that of core genes that have been vertically inherited by the corresponding lineages. We rely on synonymous mutations only to account for the fact that genes and TEs evolve under different selection regimes. Specifically, we considered that a TE-TE hit resulted from HT when its associated d_s was lower than at least 99.5% of the d_s values measured for these core genes. In addition, we required the d_s value to be below 0.5 (Zhang et al. 2020) to further ensure the recency of the HTT events. Doing so, we retained 17,956,805 TE-TE hits that are expected to result from relatively recent HTT (Fig. 1). Examining these TE-TE hits, we determined that we could not confidently distinguish HT from vertical inheritance between clades separated by a median d_s of BUSCO genes lower than 0.85 (see Supplemental Methods; Supplemental Fig. S2A). Therefore, we discarded TE-TE hits between these clades and defined such clades as “species unit” (Fig. 2B).

To estimate the minimum number of independent HTT events that could explain these patterns of elevated interspecific TE similarity, we applied the approach developed in Zhang et al. (2020) (Fig. 1). The method deals with the fact that a single HTT

event can lead to many TE-TE hits, due to TE replication within genomes, and that fragmentation of TE copies into shorter sequences may lead to overestimating HTT events. It also removes redundancy in HTT counts by clustering likely identical HTT events recovered in multiple species (Supplemental Fig. S4). This approach yielded a minimum of 6043 independent HTT events among the 247 species (Fig. 2B; Supplemental Dataset S3). We stress that the studied species may be only indirectly involved in these transfers, as they likely inherited copies of TEs that were acquired or emitted by more or less distant relatives or ancestors (Supplemental Fig. S4). Transfers involve all 19 taxonomic groups and 85% of the investigated pairs of taxonomic groups, confirming that TEs have undergone transfers throughout the animal kingdom, between lineages separated by various divergence times.

Despite their high abundance in animal genomes, Class 1 TEs only make up 13.66% of the total number of HTT events (Fig. 2), as previously observed (Peccoud et al. 2017; Reiss et al. 2019; Zhang et al. 2020). Among Class 2 TEs, Mariner-like and hAT-like TEs make up the highest proportion of transfers (respectively, 37.12% and 30.18% of total transfers). Such a high number of transfers for hAT-like TEs contrasts with earlier studies in insects and vertebrates, in which Mariner-like TEs appeared to transfer horizontally at a much higher rate than other TEs (Peccoud et al. 2017; Zhang et al. 2020). We believe that the higher number of transfers for hAT-like TEs reported in the present study reflects better TE annotation methods rather than differences in taxa composition (Supplemental Methods; Supplemental Fig. S5). Because the number of TE copies in a genome is the product of incoming TEs through horizontal transfer, persistence rate, and within-genome transposition, we assessed whether different TE superfamilies differ in terms of their per-copy rate of transfer. Once normalized (Fig. 2A), the per-copy transfer rates of Mariner-like and hAT-like TEs (273.5 and 328.5 transfers per million copies, respectively) are, in fact, lower than those of other TE superfamilies, such as *Academ*, *Ginger*, and *PiggyBac* (638.9, 380.7, and 355.6 transfers per million copies, respectively). Thus, the higher absolute number of transfers observed for Mariner-like and hAT-like TEs may not reflect a higher propensity of their copies to horizontally transfer but rather their abundance within genomes.

No evidence that aquatic habitats foster horizontal transfer of transposable elements

Investigating factors that may affect HTT rates requires counting transfers in a manner that is not affected by the number of sampled taxa. The clustering method we used to estimate the minimal number of horizontal transfer events in the whole data set does not fulfill this requirement as it is more likely to detect HTT in clades containing more species. To circumvent this issue, the 17,956,805 TE-TE hits that we selected as resulting from HTT (see previous section) were clustered for each pair of species (see Methods; Supplemental Dataset S4). This approach yielded a total of 25,409 transfers summed over all pairs (Fig. 1). This number exceeds the one reported in the previous section because a same event of horizontal transfer can be counted several times in different pairs of related species (Supplemental Fig. S4). On the other hand, HTT counts in a given species pair are not affected by HTT detection in other pairs and are therefore also not biased by the composition of sampled taxa. The detected transfers are distributed across 6054 species pairs, representing ~25% of all the pairs included in our search (Supplemental Dataset S5). Twenty-eight

species, belonging to seven taxonomic groups, were not involved in a single HTT event (Supplemental Fig. S6).

To avoid any further bias that could be caused by our selection of different numbers of aquatic versus terrestrial species, we developed a repeated random sampling approach. At each sampling, we drew one species per taxonomic group and habitat type, and we counted only the transfers detected among the sampled species (Fig. 3A). Here, semiaquatic species were assigned to the aquatic habitat because they develop in, or partially occupy, the medium suspected to facilitate HTT. The difference between the number of transfers in the aquatic and in the terrestrial species was then computed for each of the 19 taxonomic groups, and the median of this difference was calculated over the groups as an estimator of the overall effect of the habitat. Because these groups represent independent habitat transitions, the estimator is not affected by phylogenetic inertia. Thus, this median should only be affected by differences in habitats, not by biological features specific to certain lineages (e.g., TE composition and genome size, which may influence the arrival and retention rates of TEs). To derive a *P*-value, this median was then compared to values of the equivalent statistic generated through permutations under the null hypothesis of absence of effect of the aquatic habitat (see Supplemental Methods). We performed 1000 samplings of species and calculated a *P*-value separately for each sampling. The distribution of the median difference in HTT rates (aquatic—terrestrial) obtained for the 1000 samplings was centered near zero (Fig. 3A), and only seven of the 1000 samplings resulted in *P* < 0.05 for more HTT in aquatic habitats (Supplemental Fig. S8A). There is therefore no systematic evidence nor even a trend suggesting more HTT in aquatic species (Fig. 3A). We also obtained similar results when semiaquatic species were removed (*P* < 0.05 in none of the samplings across the 11 taxonomic groups) (Supplemental Fig. S8B).

The absence of a trend for higher HTT rates in aquatic habitats contrasts with the high rate of transfers measured in certain aquatic taxa (Zhang et al. 2020). To investigate this apparent contradiction, we checked our ability to measure the previously identified excess of HTT in actinopterygians (ray-finned fishes) compared to other vertebrates (Zhang et al. 2020). Applying the procedure used in Zhang et al. (2020) to our data did reveal a similar excess (Supplemental Fig. S9). Importantly, this excess does not constitute evidence for a positive effect of the aquatic habitat on HTT, because actinopterygians constitute a single clade subject to confounding factors and phylogenetic inertia. Furthermore, we assessed whether the absence of higher HTT rates in aquatic versus terrestrial habitats could reflect a higher rate of TE loss in aquatic species but found no evidence for such an effect (Supplemental Methods). We also tested our ability to detect the much-expected positive influence of habitat similarity on horizontal transfer rates. Such a correlation should naturally arise from the greater probability of contact between organisms occupying similar habitats. For this analysis, all aquatic habitats were considered “similar” to each other and “different” from terrestrial ones (themselves considered similar to each other). We then compared the mean number of transfers per pair of species occupying similar habitats to that measured per pair of species occupying different habitats. To account for the confounding effect of species relatedness on transfer rates, comparisons were restricted to pairs of species that diverged from the same most recent common ancestor (MRCAs). Importantly, transfers involving species pairs having different MRCAs should not result from the same underlying events (Zhang et al. 2020) and can be considered independent in that regard. Furthermore, mean transfer rates are not computed per

species but per species pair in which HTT was studied independently. Consequently, the analysis should not be affected by taxon sampling. Among the 47 MRCAs that were amenable to this comparison, 32 (68%) showed a higher mean rate of HTT between species occupying similar habitats (Fig. 3B). These MRCAs are

scattered across the phylogeny (Supplemental Fig. S7). Their proportion significantly exceeds 0.5 ($p \sim 0.009$, one-tailed test), assuming independence of HTT rates measured in different MRCAs. As opposed to the aquatic medium, habitat similarity thus appears to generally foster horizontal transfers in animals.

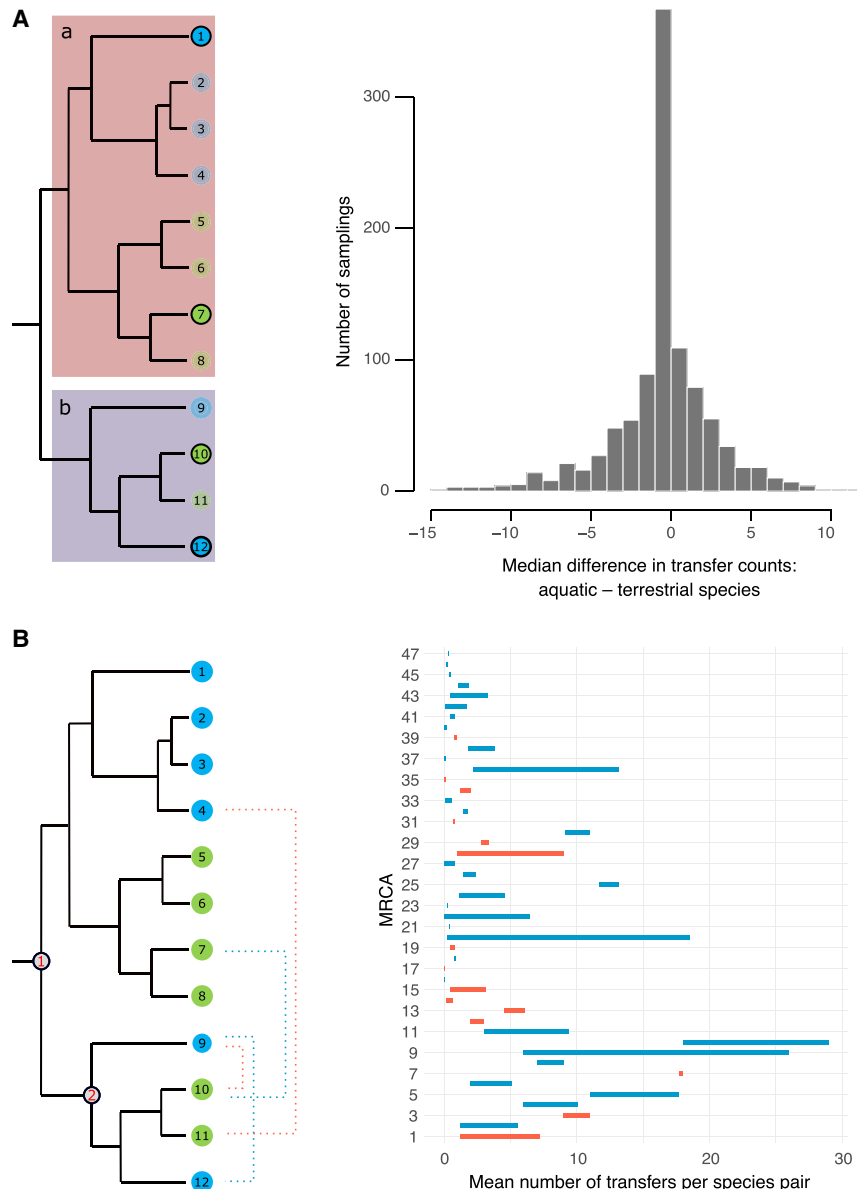


Figure 3. The effect of habitat on horizontal transfer of transposable elements. (Left) Principles of the approaches explained using dummy phylogenetic trees composed of 12 species belonging to two taxonomic groups (a,b). Aquatic and terrestrial species are indicated with blue and green circles, respectively. (Right) Results of analyses. (A) The effect of aquatic habitat. For each taxonomic group and habitat, one species (circled in black) is randomly sampled and its total number of HTT with the other sampled species (here, three species) is calculated. The median of the difference of HTT counts between the aquatic and the terrestrial species is then computed across all taxonomic groups. This sampling procedure was repeated 1000 times to obtain a distribution of this median (right-hand histogram). (B) The effect of habitat similarity. For species that diverged from the same common ancestor (MRCA), the mean number of HTT events for a pair of species occupying similar habitats (represented by blue dashed lines) is compared to that measured for a pair of species occupying different habitats (red dashed lines). This comparison can only be made for the MRCAs numbered in red on the tree. Segments on the right-hand plot connect both means, computed for 47 MRCAs numbered on the y-axis (see Supplemental Fig. S7 for their locations on the time tree). For clarity, only even MRCA numbers are shown. Blue and red segments denote higher means for species pairs occupying similar and different habitats, respectively.

Phylogenetic relatedness is a major factor shaping horizontal transfer of transposable elements

To globally assess how phylogenetic relatedness promoted HTT among metazoans, we first grouped species pairs into classes of additive divergence time (“a.d.t.”, which is twice the age of their last common ancestors). We then computed the per-class proportion of species pairs involved in at least one HTT event. The result of this approach suggests a strong impact of phylogenetic distance: the proportion of species pairs not involved in HTT showed a clear positive correlation with divergence time (weighted Pearson’s $r=0.69$) (Fig. 4A). Whereas roughly 60% of the pairs of species separated by about 250 million years (My) a.d.t. are not involved in any HTT, the proportion reaches 90% for the deepest times. Moreover, after removing species pairs not involved in HTT, the median number of HTT appeared strongly negatively correlated with divergence time (weighted Pearson’s $r=-0.72$) (Fig. 4A). These results suggest that the tendency to transfer more among closely related species may be widespread among animals.

We next studied each species separately in order to more precisely estimate the prevalence of the effect across the phylogeny. We adopted a modeling approach because it allowed us to explicitly quantify the strength of the effect in each species, rather than simply testing whether an effect was present. The models were formulated in the Bayesian framework, setting a regularizing prior on the effect of divergence time to avoid overfitting in species that are involved in few transfers (see Methods). For each species, we fit a negative binomial regression model of the number of transfers as a function of the a.d.t. to each other species (Fig. 5; Supplemental Fig. S6). As a second covariate, our models also included habitat similarity, as defined in the previous section, to help mitigate a potential confounding effect.

Out of the 219 species involved in transfers, 162 (~74%) showed a confident negative effect, that is, a higher estimated HTT rate with closely related

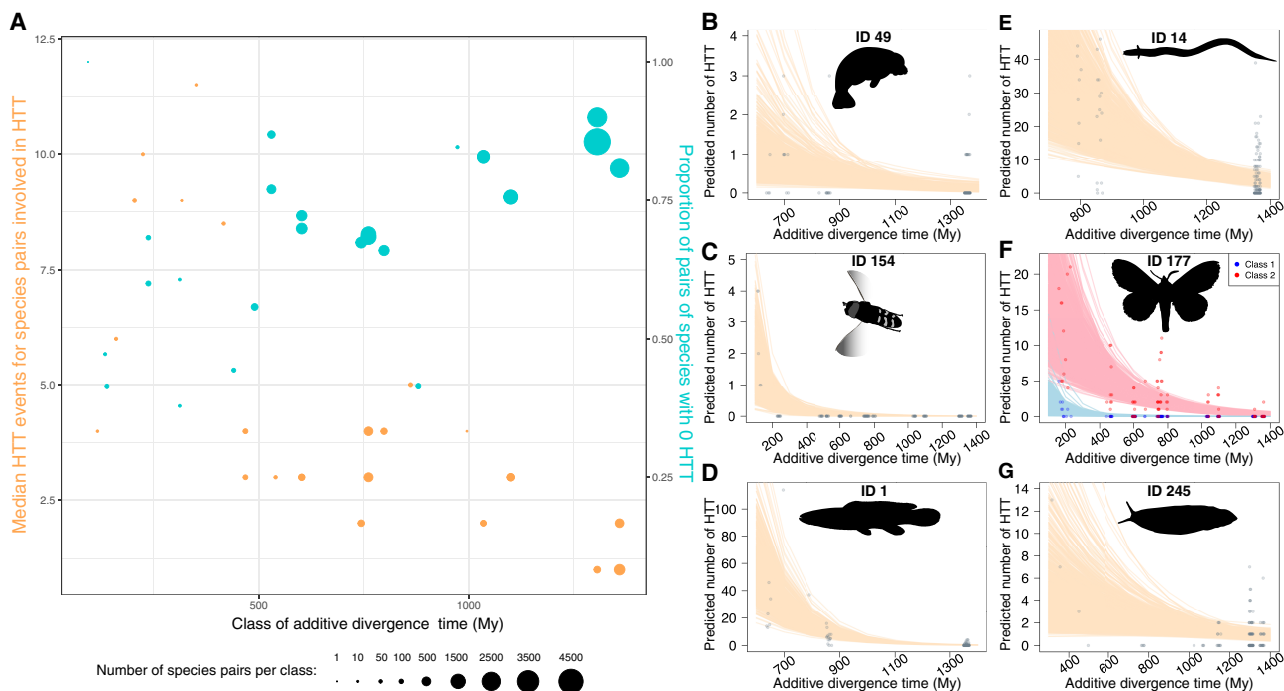


Figure 4. Phylogenetic relatedness favors horizontal transfer of transposable elements. (A) Correlation between divergence time and (i) the median number of HTT events per pair of species, excluding pairs not involved in HTT (orange circles, *left y-axis*, $r = 0.69$), and (ii) the proportion of pairs of species in which no HTT was inferred (turquoise circles, *right y-axis*, $r = -0.72$). Both (i) and (ii) were computed per class of divergence time, which are 100 My a.d.t. wide and overlap by 50 My a.d.t., using one species per species unit (Fig. 2B). Both correlation coefficients are weighted using the natural logarithm of the number of species pairs composing each class, to account for more precise estimates when many species pairs compose the median or proportion (here, the higher the a.d.t., the more species pairs are in the class of additive divergence time). Circles are horizontally located at the median divergence times within classes. (B–G) Estimates from a negative binomial regression model of the number of HTT events in a focal species as a function of the divergence time from “partner” species. IDs of the focal species are indicated in each panel; they correspond to tips in Figure 2B and their corresponding species names can be found in Supplemental Dataset S1. The model estimates are provided with the assumption of shared habitat. In F, Class 1 and Class 2 TEs are modeled separately. Semitransparent lines correspond to the expected number of HTT estimated from each of 36,000 Markov chain Monte Carlo draws. Points represent observed data and are slightly randomly shifted on the *x-axis* to minimize overlapping.

species (Fig. 5; Supplemental Fig. S6). An effect is called “confident” when the 95% highest posterior density interval (HPDI) of the regression coefficient of divergence time does not comprise 0. In contrast, only two species showed a confident effect in the opposite direction, and the remaining species did not show any confident effect. We note that the models are not mutually independent, as the same HTT event could be counted in several closely related species and because of potential phylogenetic inertia. However, Figure 5 shows that the effect is widespread among almost all the metazoan groups studied, underlining its pervasiveness across animals. The nonindependence of the models also prevents us from combining them into a single global model to estimate the average effect size. However, the results can be summarized by characterizing the strength of the effect in a typical species. The manatee *Trichechus manatus* (ID 49) sits at the median across species regarding the strength of the phylogenetic effect and is thus representative of the typical effect size (Fig. 4B). This species is expected to present 9.63 times as many HTT events at 500 My a.d.t. than at 1300 My a.d.t. (a range of divergence times that encompasses most of the observed events).

The effect of phylogenetic relatedness is the strongest in hoverflies (*Syrphidae*), in which the detected transfers were restricted to this family (Figs. 2B, 4C). Numerous other species showed strong effects (Fig. 4B–G). For example, for the actinopterygian *Amia calva* (ID 1), the model predicted, on average, 22.21–92.02 transfers (95% HPDI) with a species separated by 600 My a.d.t., but only

0.14–0.41 transfers at 1400 My a.d.t. (Fig. 4D). Note that for all predictions, we fixed the habitat to be similar in each transfer to make sure that the estimates reflected divergence time alone. The species involved in the highest number of HTT events ($n = 1675$) is the actinopterygian fish *Erpetoichthys calabaricus* (ID 14). Despite no closely related species in the data set, as the sole polypterid (ID 14 on Figs. 2, 5), it is subject to a strong effect of phylogenetic relatedness (Fig. 4E).

Importantly, the phylogenetic relatedness effect is widespread for both Class 1 and Class 2 elements (Supplemental Methods; Supplemental Fig. S6). Nonetheless, in many species, Class 1 TEs seem to show a stronger effect (see Supplemental Methods; Supplemental Fig. S6). For example, in the moth *Dendrolimus kikuchii* (ID 177), no HTT is detected past 216 My a.d.t. for Class 1 elements, whereas Class 2 elements transfer between lineages separated by a much wider range of divergence times (Fig. 4F). However, the low number of transfers of Class 1 TEs limits the power of our modeling approach, so that statistically solid conclusions cannot be drawn regarding the relative strength of the effect in the two classes.

A benefit of the modeling approach is its ability to predict the number of transfers while keeping the divergence time fixed at a value of our choosing. This enables fair comparisons of the HTT rate between species that differ with regard to how many close relatives they have in the data set. Given the strength of the phylogenetic relatedness effect, we used this approach to retest the effect of

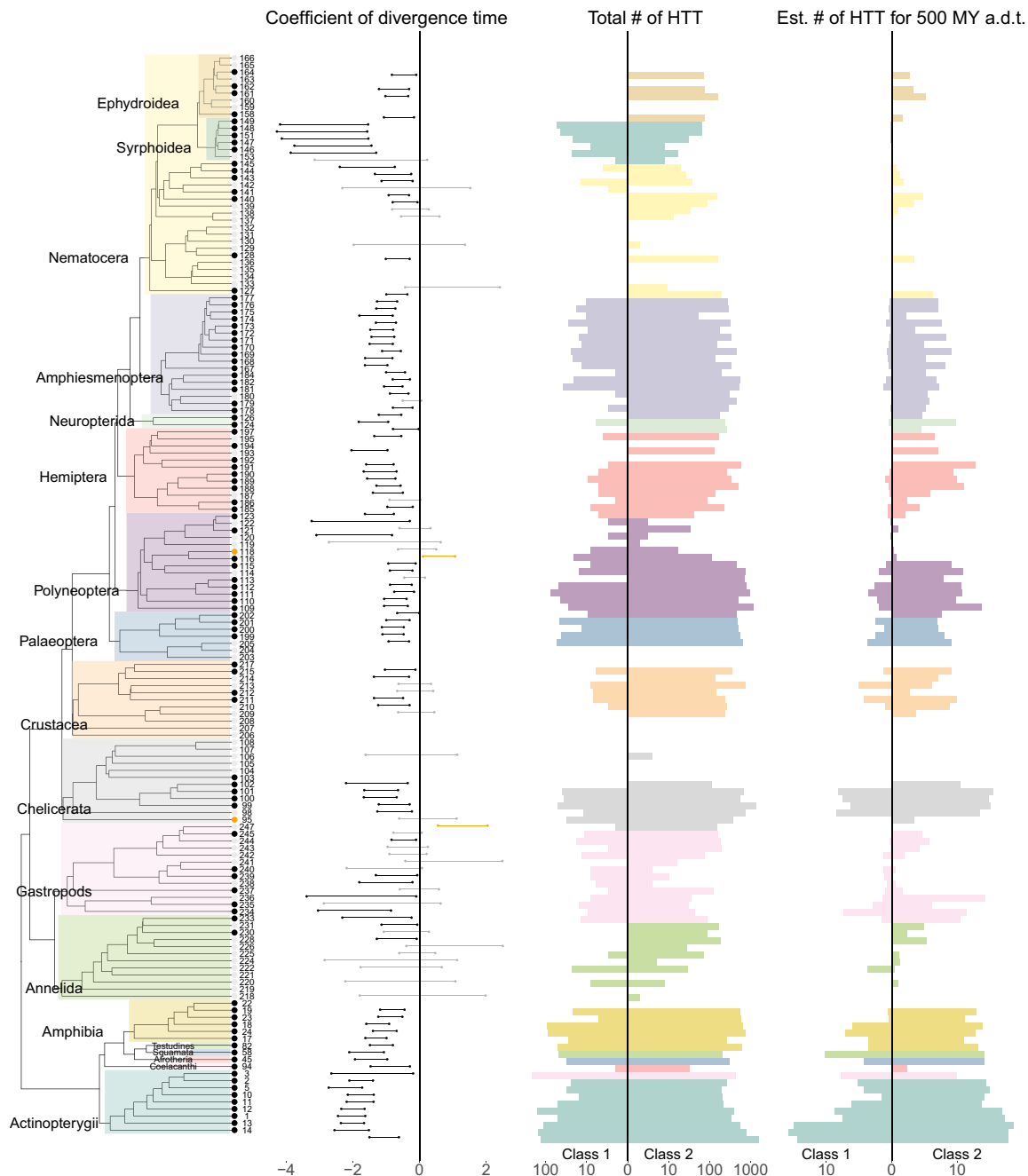


Figure 5. Modeling the effect of phylogenetic relatedness. The 95% highest posterior density interval (HPDI) of the effect of divergence time on horizontal transfer rate, computed for each species, is represented by a horizontal segment next to the corresponding tip of the tree (species are numbered as in Fig. 2). To save space, only one species per “species unit” was randomly chosen (see Supplemental Fig. S6 for the complete figure). Black, orange, and gray colors of segments and tree tips signify that the HPDI is below, above, or comprises zero, respectively, denoting a confidently negative, confidently positive, or nonconfident effect. Colors behind tree branches and in bar plots correspond to taxonomic groups, as in Figure 2. Absence of the HPDI segment indicates that no HTT events were observed for the species in question. The *left-hand* bar plot shows the observed \log_{10} count of HTT events involving each focal species, separately for Class 1 (*left-hand* bars) and Class 2 (*right-hand* bars) TEs. The *right-hand* bar plot is equivalent but shows the \log_{10} number of HTT events expected by the model when the time of divergence with the partner species is fixed at 500 My a.d.t. and a shared habitat is assumed (based on the median of the posterior distribution). Values of both bar plots can be recovered in Supplemental Dataset S6.

aquatic habitat, this time controlling for species relatedness. For each species, we estimated the expected number of HTT given a partner species that shares the same habitat and is separated by 500 My a.d.t. (Supplemental Fig. S10). We then calculated the median predicted number of HTT for each habitat type in each taxo-

nomic group (as detailed in the Methods section). No systematic tendency for a higher HTT rate in either habitat was observed at this phylogenetic distance ($P=0.29$, two-tailed paired *t*-test on the natural logarithm of the counts, sample size = 17 pairs of taxonomic groups). To summarize, neither the Bayesian models

accounting for the effect of phylogenetic relatedness nor the previous approach accounting for phylogenetic inertia and sampling bias could detect any evidence that aquatic habitat promotes HTT between metazoans.

Finally, our model also included habitat similarity as a predictor (Supplemental Fig. S11). In accordance with our previous results (Fig. 3B), a trend for greater HTT rates within similar habitats emerged: Among the 219 species involved in HTT, 73% showed higher HTT rates with species occupying similar habitats (i.e., the median of the posterior distribution of the coefficient was positive). However, the effect can be confidently called in only 59 species (95% HPDI entirely above 0). The statistical evidence for a widespread effect of habitat similarity is therefore weaker than the evidence for the effect of phylogenetic relatedness.

Discussion

We have estimated that the 247 metazoan species included in our study were involved in at least 6043 independent transfers of transposable elements. This number is three to six times higher than those reported in previous large-scale studies, despite including a similar number of species (2248 transfers among 195 insects [Peccoud et al. 2017] and 995 among 307 vertebrates [Zhang et al. 2020]). Part of the difference could be attributed to the broader phylogenetic scale of this study, which reduced the proportion of pairs of species which are too related to identify HTT. Another factor could be the improvements we achieved in HTT detection (Supplemental Methods). One of these improvements consisted in using the same, very large TE consensus library to annotate TE copies in all genomes, instead of a smaller, specific library for each species, as was done previously. Among other benefits, this large database allowed annotating many more hAT-like elements compared to our previous study (Supplemental Fig. S5; Zhang et al. 2020), leading to the discovery that these TEs were horizontally transferred at a rate similar to Mariner-like elements. When expressing HTT numbers relative to TE copy numbers, Class 2 TEs remain much more frequently transferred among metazoans than Class 1 elements, as already observed in insects and vertebrates (Peccoud et al. 2017; Reiss et al. 2019; Zhang et al. 2020). Among Class 2 TEs, however, Mariner-like and hAT-Like TEs appear to transfer at rates that are similar to other superfamilies. The overall evolutionary success of these two families among Class 2 TEs, as estimated by their number of copies, would thus result from a higher capacity to multiply and persist in genomes, not from greater rates of horizontal transfer.

Our results also highlight the importance of quantifying HTT in a way that is amenable to hypothesis testing, which has historically been challenging. Rates of gene HT are elevated in marine bacteria (McDaniel et al. 2010) and in marine unicellular eukaryotes (Van Etten and Bhattacharya 2020), and aquatic habitats have been repeatedly proposed as facilitating HT among eukaryotes, based on ad hoc interpretations (Wang and Liu 2016; Metzger et al. 2018; Galbraith et al. 2020; Gozashti et al. 2022, 2025; Hassan et al. 2024). There is also evidence supporting transfers of natterin-like toxin encoding genes involving marine animals (Gacesa et al. 2020), and the only known case of gene HT between vertebrates involves marine teleost fishes (Graham and Davies 2021; Han et al. 2023). However, our study on 19 taxonomic groups, representing independent transitions between aquatic and terrestrial habitats along the evolutionary history of metazoans, revealed no evidence that animal-borne TEs transfer more fre-

quently within compared to outside water. Because the sampling procedure picked only one species per habitat and taxonomic group at each iteration, it did not consider transfers within groups, hence missing transfers between more closely related species. However, we do not expect this limitation to affect our conclusions, unless the effect of phylogenetic relatedness was stronger for aquatic species than for terrestrial ones, causing them to be affected disproportionately by our failure to consider transfers between very close species. Currently, we have no reason to believe such a bias could be present. Thus, whereas previous observations of HTT among aquatic species (Galbraith et al. 2020; Zhang et al. 2020; Gozashti et al. 2022, 2025) indicate high transfer rates within certain specific aquatic taxa, they may not necessarily reflect a general positive effect of the aquatic medium on HTT. In this regard, the excess of HTT in teleost fishes observed here and previously (Zhang et al. 2020) would not result solely from their aquatic lifestyle but from other features. One possibility is that external fertilization, used by most fish lineages, may increase the exposure of their germline genome to foreign DNA (Graham and Davies 2021). In addition, teleost fishes harbor endogenous herpesvirus-like elements called Teratorn (Inoue et al. 2017). These elements, which generally constitute a fusion between herpesviruses (Alloherpesviridae) and Piggybac transposons, are able to multiply in fish genomes, and their distribution in the fish phylogeny is suggestive of recurrent cross-species transmission (Inoue et al. 2018; Inoue and Takeda 2023). Much like baculoviruses have been suggested to foster HTT among lepidopterans (Gilbert et al. 2016; Reiss et al. 2019), Teratorn may have acted as vectors of HTT among teleost fishes. The approach developed here could be used to investigate whether patterns of Teratorn presence/absence among teleost fishes, as well as other factors such as fertilization modes, covary with HTT rates.

Regarding the effect of phylogenetic relatedness on HTT, our modeling approach produced results at an unprecedented resolution by explicitly quantifying the strength of the effect and its variability between lineages. The true effect may be even more widespread and stronger than estimated by our models, as HTT detection is negatively affected by taxon relatedness (Peccoud et al. 2018) and was not even attempted between species that were too closely related. In addition, more divergent species have had more time to exchange TEs, producing an opposite phylogenetic effect, which is mitigated by the fact that we mostly recovered recent transfers. Finally, the regularizing prior used biases towards an underestimation of the effect size. The fact that closely related lineages incur higher rates of HTT is consistent with studies of other types of horizontal transfer, such as gene transfers among viruses and prokaryotes, as well as cross-species transmission of viruses within ecosystems (Groussin et al. 2021; Sheinman et al. 2021; French et al. 2023; Wu et al. 2024). The effect of phylogenetic relatedness can be explained by two factors: (i) a higher exposure of recipient hosts to TEs coming from closely related species, due to more frequent interspecific interactions (Venner et al. 2017); and (ii) a higher probability that TEs persist in a new host that is closely related to their emitter, because they rely on host components to transpose (Levin and Moran 2011; Peddigari et al. 2013; Protasova et al. 2021). Our results support a moderate contribution of the first factor, given that habitat similarity, hence, potentially, engagement in more frequent interspecific interactions, induces higher HTT rates, as found in bacteria exchanging genes (Popa and Dagan 2011; Dmitrijeva et al. 2024). Regarding the second factor, compatibility between TEs and recipient hosts may indeed play a role in conditioning horizontal transfers, as previously

proposed (Palazzo et al. 2017, 2019; Peccoud et al. 2017). Under this hypothesis, our results suggest that molecular or physiological compatibilities modulate horizontal TE transfer in most animal lineages. Our findings also hint at a stronger effect of phylogenetic relatedness for Class 1 TEs, in line with the higher reliance on host factors of these TEs, compared to Class 2 TEs (Lampe et al. 1996; Levin and Moran 2011; Peddigari et al. 2013; Palazzo et al. 2017, 2019; Protasova et al. 2021). Thus, although Class 1 TEs are more numerous in genomes, their overall lower HT rate could, in part, be explained by stronger interhost genetic barriers than for Class 2 TEs. To some extent, the lower rate of HTT detected here and in earlier studies for Class 1 TEs might partly be due to a higher mutation rate, as Class 1 TEs undergo error-prone reverse-transcription. Higher mutation rates are expected to cause horizontally transferred Class 1 TEs to exceed divergence thresholds used for HTT detection more rapidly than Class 2 elements. This could lower our power to detect ancient HTT events for these TEs compared to Class 2 TEs. However, a study comparing daughter LINE 1 copies to their progenitor sequence suggests that the error rate is rather low during reverse-transcription (≈ 1 misincorporated nucleotide per LINE 1 copy) (Gilbert et al. 2005). Furthermore, a lower HTT rate for Class 1 compared to Class 2 TEs was measured in insects (Peccoud et al. 2017). In that study, most detected HTT events occurred relatively recently, that is, within the last 10 million years. There is thus no reason to think that our ability to detect such recent horizontal transfers of Class 1 TE would be lower between distantly related lineages than between closely related ones. Beyond these general trends, our fine-scale approach revealed several atypical taxa, such as the polypterid fish *Erpetoichthys calabaricus*, which show a particularly high rate of transfers, even with distantly related species, and the syrphid flies, which did not show any transfers with any other groups of animals. In-depth investigation of these taxa might improve our understanding of the factors underlying HTT.

In conclusion, our study has fundamental implications not only for understanding the determinants that shape HTT but also from a methodological viewpoint. If the taxonomic scale and the structure of the phylogeny are not considered, differences in HTT rates between taxa may simply reflect the varying proportion of close relatives included in the data set. The strength of the phylogenetic relatedness effect might make the analysis of other predictors particularly challenging. This challenge should motivate the development of novel statistical approaches to unravel the mechanisms that underlie, and the factors that shape, the horizontal transfer of transposable elements.

Methods

Selecting genome assemblies

In the NCBI database, as of January 2022, we identified 19 taxonomic groups for which genome assemblies were available both for aquatic (or semiaquatic) and terrestrial animal species. Whereas 17 of the 19 taxonomic groups contain both aquatic and terrestrial species, the two remaining groups (fishes and Palaeoptera) only comprise aquatic (or semiaquatic) species.

To limit computational workload and reduce pseudoreplication, we selected a single species per genus, the one comprising the most contiguous genome assembly, as estimated by the N50 metric. We also selected a maximum of 10 genera per group (except for fishes, for which we selected 15 genera), favoring taxonomical diversity, a consistent assembly size, and a high N50. This strategy yielded a data set of 247 animal species

(Supplemental Dataset S1), on which we built a dated phylogenetic tree (Supplemental Methods; Supplemental Dataset S2). We ran BUSCO v.5.4 (Manni et al. 2021) on the 247 genomes to assess their completeness and to extract core genes for downstream analyses, using the most specific BUSCO data set for each genome (option -l, Supplemental Dataset S1). See Supplemental Figure S12 for results.

Transposable element annotation

We de novo annotated TEs in the 247 genomes (Fig. 1; Supplemental Fig. S1) with RepeatModeler v2 (Flynn et al. 2020) using the option LTRStruct. Annotation failed on two genomes: *Trachelipus rathkii* and *Armadillidium vulgare*. The TE consensus library of *A. vulgare* was taken from Chebbi et al. (2019). We concatenated TE consensus sequences from the 246 species with those obtained on 195 insects and 307 vertebrate genomes (Peccoud et al. 2017; Zhang et al. 2020) and those from Repbase (Bao et al. 2015) (downloaded in February 2022), excluding SINEs, satellites, and tandem repeats. Keeping only TE consensus sequences longer than 300 bp and classified in a superfamily resulted in a single database of 284,643 TE consensus sequences (223,414 of which are available in Supplemental Dataset S7). Clustering this database with MMseqs2 (Steinegger and Söding 2017), using options `-cluster-reassign, -c 0.8, and -min-seq-id 0.8`, yielded 217,391 TE consensus sequences. We used this large library to annotate TE copies in all 247 genomes, using RepeatMasker with the options `-nolow, -no_is, -norma, and -engine ncbi` (Smit et al. 2013). Using a unique library helps annotating TEs in genomes where de novo TE detection was less effective, either because of a lower genome quality or because the TE burst did not happen yet in the recipient genome (Le Rouzic and Capy 2005). Copies, or fragments of copies, shorter than 300 bp and those included in a higher-scoring match by RepeatMasker were discarded. Doing so, we extracted a total of 111,102,018 TE copies, or fragments of copies, from the 247 assemblies. Although this approach is not an exhaustive detection of TEs, it allowed to automatically recover most TE copies from all genomes, as reflected by TE contents listed in Supplemental Dataset S1, which are similar to TE contents obtained in previous studies. For example, we found that nearly 23% of the macaque genome is made of LINES, a figure typical of primate genomes. Furthermore, using BLASTN (Camacho et al. 2009), we confirmed that our library of TE consensus contains members of TE families that were subjected to recent detailed characterization and analysis such as DD34E, Hiker, and DD37D (Supplemental Dataset S8; Wang et al. 2021, 2024; Shi et al. 2023).

Identifying hits resulting from horizontal transfer

To identify and count HTT events, we used the approach we developed previously (Zhang et al. 2020), to which we made a number of improvements (see Supplemental Fig. S1; Supplemental Methods). Reciprocal sequence similarity searches of extracted TE copies were performed for the 30,313 pairs of species (see Supplemental Methods). Filtering out alignments shorter than 300 bp, with nucleotide sequence identity $<75\%$, and with quality score <200 resulted in 247,248,663 TE-TE hits (corresponding to 47,844,407 TE copies). Of these, 97,187,587 hits involved a protein-coding region long enough (≥ 300 bp) to calculate d_s . To reduce the workload, redundant hits were discarded (see Supplemental Methods). d_s values between copies involved in the retained hits were calculated following Zhang et al. (2020). We then selected hits ($n = 17,956,805$) whose d_s were <0.5 and lower than the 0.5% quantile of the distribution of d_s expected under vertical inheritance in the appropriate clade (see Supplemental Methods).

Clustering of hits between transposable elements

We clustered selected TE-TE hits using the approach developed in Zhang et al. (2020), with some improvements (see [Supplemental Methods](#)). This approach considers two cases where several TE-TE hits can result from the same HTT event. The first case is due to TE duplication within genomes. For such hits, we expect the sequence identity between TE copies of different species to be lower than the identity of TE copies in the recipient taxon. This criterion allows clustering hits between two taxa, each taxon corresponding to a species or a larger clade, depending on the analysis (see below). In practice, any two hits whose sequence identity was equal to or lower than the within-taxon identity of TE copies involved (for at least one taxon) were connected to constitute a graph, on which a clustering algorithm was applied (see [Supplemental Methods](#)). This generates “communities,” ideally representing different transfer events. A second case where a single HTT event can lead to several hits is when annotated copies are fragmented. Our approach thus considers that TE-TE hits involving nonoverlapping parts of the same TE would be erroneously assigned to two transfers. It measures protein sequence identity between TEs of different communities to aggregate them into larger “hit groups” (see [Supplemental Methods](#)).

This clustering scheme was used for two independent analyses to (i) estimate the minimal number of HTT events across animals, and (ii) to count HTT events independently in each species pair (Fig. 1). In the first analysis, clustering was applied in a two-step procedure described in Zhang et al. (2020), where taxa correspond to young clades (of <80 My) in the first step, then to deeper sister clades in the second step ([Supplemental Methods](#)). This yielded 55,156 “hit groups,” of which 17,973 were retained after applying stringent filters to avoid false positives (see [Supplemental Methods](#); [Supplemental Fig. S2](#)). One of these filters is based on the expectation that Class 2 TEs evolve neutrally within genomes (nonautonomous copies can be transposed by proteins encoded by others), whereas only autonomous TEs can successfully transfer horizontally, inducing purifying selection ([Supplemental Fig. S3B](#); Zhang et al. 2020). It turns out that, for species separated by a median d_s of BUSCO genes lower than 0.85, the Class 2 TEs originally selected as resulting from HTT showed a lower signal of purifying selection (higher d_N/d_S) ([Supplemental Fig. S2A](#)). This pattern suggested that the divergence of some TE copies belonging to genetically related species (median d_s of BUSCO genes <0.85) may not have involved horizontal transfer. We therefore decided to ignore all HTT events inferred in these species pairs. In the second analysis, the clustering scheme was applied on each species pair separately and comprised only one step. This yielded 96,908 “hit groups,” of which 25,409 were retained after applying the same stringent filters as in the previous analysis. Retained “hit groups” can be found in [Supplemental Datasets S3 and S4](#) for the first and second analysis, respectively.

Counting independent horizontal transfer events

Counting HTT events among more than two lineages must take into consideration that a horizontal transfer between taxa A and B and another between taxa A and C might lead to counting a third transfer between taxa B and C, which never took place. To avoid this, we applied the method detailed in Zhang et al. (2020) on the 17,973 “hit groups,” which is illustrated in their [Supplemental Figure 8](#) (Zhang et al. 2020). Briefly, this method relies on sequence identities between TE copies hosted by species composing the different taxa to conservatively identify apparent transfers that may be explained by others. These nonindependent

transfers were removed from counts and are not shown on Figure 1.

Testing the effect of phylogenetic relatedness

To test the effect of phylogenetic relatedness on HTT rates, we fitted a negative binomial regression model with a logarithmic link to each species involved in at least one HTT (219 out of 247 species). For clarity, we will refer to the species being modeled as the focal species, and to the species with which there has been an HTT event as the partner species. As the method is agnostic as to the direction of the transfer, it counts both the events where the focal species is the recipient as well as the ones where it is the source of the transfer. The same HTT event could also be counted several times if some of the partner species are closely related to each other ([Supplemental Fig. S4](#)). To avoid this issue, we grouped partner species that were too closely related into “units.” We chose a relatedness threshold corresponding to an average d_s of BUSCO genes of <0.85. We considered such a unit as a single partner species in the model, taking the median of the number of HTT events between the focal species and such a group.

The likelihood was of the form:

$$n_{sp} \sim \text{NegativeBinomial}(\mu_{sp}, \phi_s),$$

$$\ln \mu_{sp} = \beta_{0s} + \beta_{1s} * d_{sp} + \beta_{2s} * h_{sp},$$

where n_{sp} is the number of HTT events for focal species s with partner species (or species unit) p , μ_{sp} is the mean number of HTT events expected for focal species s with partner species p , and ϕ_s is the negative binomial shape parameter for this focal species. β_{0s} corresponds to the regression intercept in focal species s . β_{1s} is the regression coefficient for the divergence time d_{sp} , and β_{2s} is the regression coefficient for shared habitat h_{sp} (encoded as 1 when habitat is shared and 0 when it is not). Before modeling, divergence times were centered and scaled by subtracting the mean and dividing by the standard deviation, both computed over the whole data set. β_{0s} thus corresponds to the log number of transfers expected in a given species when the divergence time with the partner species corresponds to the mean of the data set and habitat is not shared. Fairly loose priors were used for the intercept ($\beta_{0s} \sim \text{Normal}(0, 10)$) and the shape ($\phi_s \sim \text{Gamma}(0.001, 0.001)$), as well as the regression coefficient of shared habitat ($\beta_{2s} \sim \text{Normal}[0, 3]$). A stricter prior was used for the regression coefficient of divergence time, ($\beta_{1s} \sim \text{Normal}(0, 1)$), to help avoid overfitting and make conclusions more conservative. We used a weaker prior for the effect of habitat because this predictor was included primarily as a potential confound for phylogenetic relatedness and it was therefore more conservative to risk overestimating the effect than to risk underestimating it. We also fitted the models separately for Class1 and Class2 TEs ([Supplemental Methods](#)). Each model was fitted in R 4.4.1 (R Core Team 2024) with `brms` 2.21.0 (Bürkner 2017) using four Markov chain Monte Carlo chains with 10,000 iterations per chain, 1000 of which for warm-up. This yielded 36,000 post-warmup draws per model. `adapt_delta` was set to 0.99 in the `brm()` function. Rhat values did not exceed 1.01, indicating that the chains converged successfully. Bulk and tail effective sample size (ESS) values were at least 2000, suggesting that posterior estimates are reliable. The rethinking 2.40 R package (McElreath 2020) was used to calculate highest posterior density intervals. All ranges of parameter estimates and posterior predictions that are presented correspond to 95% HPDIs; that is, the model gives 95% confidence that the population parameter value is within the range.

To determine the median effect of the divergence time (see Results), we used the `brms` `posterior_epred()` function to estimate

the average number of HTT events expected at 500 My a.d.t., as well as 1300 My a.d.t., for each species separately. This calculation was performed for each post-warmup Markov chain Monte Carlo draw, after which the median across the draws was taken. Finally, the species medians at 500 My a.d.t. were divided by those at 1300 My a.d.t. to obtain a fold difference, the median of which is reported in the Discussion.

Data access

All the scripts used in this study were written in R (R Core Team 2024) and can be found in Supplemental Code S1 and at GitHub (<https://github.com/HeloiseMuller/HTvertebrates>). All Supplemental Datasets are available at Zenodo (<https://doi.org/10.5281/zenodo.15297793>). Supplemental Datasets S1, S2, S5–S8 are also available as Supplemental Data. Legends for all Supplemental Datasets can be found in the Supplemental Information file, which also contains Supplemental Methods and Supplemental Figures S1–S12.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

The following infrastructures provided computational support: the genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul [<https://doi.org/10.15454/1.5572369328961167E12>]), the GenOuest bioinformatics core facility (<https://www.genouest.org>), the Core Cluster of the Institut Français de Bioinformatique (IFB) financed under the Programme d'Investissements d'Avenir funded by the Agence Nationale de la Recherche (RENABI-IFB ANR-11-INBS-0013 and MUDIS4LS ANR-21-ESRE-0048), and the Roscoff Bioinformatics platform ABiMS (<http://abims.sb-roscoff.fr>). Antoine Fouquet provided his expertise on amphibians to identify their lifestyle. This work was supported by Agence Nationale de la Recherche, projects ANR-17-CE02-0021-01 HORIZON, ANR-18-CE02-0021-01 Transp Horizon, and VIRHOZFER ANR-24-CE02-1004. Part of this work benefitted from intramural funds from the CNRS and the University of Poitiers.

Author contributions: H.M., R.S., J.P., S.C., and C.G. designed the study. H.M. and R.S. carried out all the analyses. All authors have read, edited, and approved the final manuscript.

References

- Abe K, Nomura N, Suzuki S. 2020. Biofilms: hot spots of horizontal gene transfer (HGT) in aquatic environments, with a focus on a new HGT mechanism. *FEMS Microbiol Ecol* **96**: fiae031. doi:10.1093/femsec/fiae031
- Baidouri ME, Carpentier M-C, Cooke R, Gao D, Lasserre E, Llauro C, Mirouze M, Picault N, Jackson SA, Panaud O. 2014. Widespread and frequent horizontal transfers of transposable elements in plants. *Genome Res* **24**: 831–838. doi:10.1101/gr.164400.113
- Bao W, Kojima KK, Kohany O. 2015. Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob DNA* **6**: 11. doi:10.1186/s13100-015-0041-9
- Bartolomé C, Bello X, Maside X. 2009. Widespread evidence for horizontal transfer of transposable elements across *Drosophila* genomes. *Genome Biol* **10**: R22. doi:10.1186/gb-2009-10-2-r22
- Brann T, de Oliveira FS, Protasio AV. 2024. Horizontal transfer of a LINE-RTE retrotransposon among parasite, host, prey and environment. *bioRxiv* doi:10.1101/2024.11.24.625053v2
- Bürkner P-C. 2017. brms: an R package for Bayesian multilevel models using Stan. *J Stat Softw* **80**: 1–28. doi:10.18637/jss.v080.i01
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421. doi:10.1186/1471-2105-10-421
- Carvalho TL, Cordeiro J, Vizentin-Bugoni J, Fonseca PM, Loreto ELS, Robe LJ. 2023. Horizontal transposon transfer and their ecological drivers: the case of flower-breeding *Drosophila*. *Genome Biol Evol* **15**: evad068. doi:10.1093/gbe/evad068
- Chebbi MA, Becking T, Moumen B, Giraud I, Gilbert C, Peccoud J, Cordaux R. 2019. The genome of *Armadillidium vulgare* (Crustacea, Isopoda) provides insights into sex chromosome evolution in the context of cytoplasmic sex determination. *Mol Biol Evol* **36**: 727–741. doi:10.1093/molbev/msz010
- Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A. 1990. Evidence for horizontal transmission of the P transposable element between *Drosophila* species. *Genetics* **124**: 339–355. doi:10.1093/genetics/124.2.339
- de Melo ES, Wallau GL. 2020. Mosquito genomes are frequently invaded by transposable elements through horizontal transfer. *PLoS Genet* **16**: e1008946. doi:10.1371/journal.pgen.1008946
- Dmitrijeva M, Tackmann J, Matias Rodrigues JF, Huerta-Cepas J, Coelho LP, von Mering C. 2024. A global survey of prokaryotic genomes reveals the eco-evolutionary pressures driving horizontal gene transfer. *Nat Ecol Evol* **8**: 986–998. doi:10.1038/s41559-024-02357-0
- Dupuy C, Periquet G, Serbielle C, Bézier A, Louis F, Drezén J-M. 2011. Transfer of a chromosomal Maverick to endogenous bracovirus in a parasitoid wasp. *Genetica* **139**: 489–496. doi:10.1007/s10709-011-9569-x
- Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci* **117**: 9451–9457. doi:10.1073/pnas.1921046117
- French RK, Anderson SH, Cain KE, Greene TC, Minor M, Miskelly CM, Montoya JM, Wille M, Muller CG, Taylor MW, et al. 2023. Host phylogeny shapes viral transmission networks in an island ecosystem. *Nat Ecol Evol* **7**: 1834–1843. doi:10.1038/s41559-023-02192-9
- Gacesa R, Hung JY, Bourne DG, Long PF. 2020. Horizontal transfer of a natterin-like toxin encoding gene within the holobiont of the reef building coral *Acropora digitifera* (Cnidaria: Anthozoa: Scleractinia) and across multiple animal lineages. *J Venom Res* **10**: 7–12.
- Galbraith JD, Ludington AJ, Suh A, Sanders KL, Adelson DL. 2020. New environment, new invaders—repeated horizontal transfer of LINES to sea snakes. *Genome Biol Evol* **12**: 2370–2383. doi:10.1093/gbe/evaa208
- Gilbert C, Cordaux R. 2017. Viruses as vectors of horizontal transfer of genetic material in eukaryotes. *Curr Opin Virol* **25**: 16–22. doi:10.1016/j.coviro.2017.06.005
- Gilbert C, Feschotte C. 2018. Horizontal acquisition of transposable elements and viral sequences: patterns and consequences. *Curr Opin Genet Dev* **49**: 15–24. doi:10.1016/j.gde.2018.02.007
- Gilbert N, Lutz S, Morrish TA, Moran JV. 2005. Multiple fates of L1 retrotransposition intermediates in cultured human cells. *Mol Cell Biol* **25**: 7780–7795. doi:10.1128/MCB.25.17.7780-7795.2005
- Gilbert C, Schaack S, Pace JK, Brindley PJ, Feschotte C. 2010. A role for host-parasite interactions in the horizontal transfer of DNA transposons across animal phyla. *Nature* **464**: 1347–1350. doi:10.1038/nature08939
- Gilbert C, Peccoud J, Chateigner A, Moumen B, Cordaux R, Herniou EA. 2016. Continuous influx of genetic material from host to virus populations. *PLoS Genet* **12**: e1005838. doi:10.1371/journal.pgen.1005838
- Gozashti L, Roy SW, Thornlow B, Kramer A, Ares M, Corbett-Detig R. 2022. Transposable elements drive intron gain in diverse eukaryotes. *Proc Natl Acad Sci* **119**: e2209766119. doi:10.1073/pnas.2209766119
- Gozashti L, Nakamoto A, Russell S, Corbett-Detig R. 2025. Horizontal transmission of functionally diverse transposons is a major source of new introns. *Proc Natl Acad Sci* **122**: e2414761122. doi:10.1073/pnas.2414761122
- Graham LA, Davies PL. 2021. Horizontal gene transfer in vertebrates: a fishy tale. *Trends Genet* **37**: 501–503. doi:10.1016/j.tig.2021.02.006
- Groussin M, Poyet M, Sistiaga A, Kearney SM, Moniz K, Noel M, Hooker J, Gibbons SM, Segurel L, Froment A, et al. 2021. Elevated rates of horizontal gene transfer in the industrialized human microbiome. *Cell* **184**: 2053–2067.e18. doi:10.1016/j.cell.2021.02.052
- Han Z, Xu S, Gao T. 2023. Unexpected complex horizontal gene transfer in teleost fish. *Curr Zool* **69**: 222–223. doi:10.1093/cz/zoac032
- Hassan NT, Galbraith JD, Adelson DL. 2024. Multiple horizontal transfer events of a DNA transposon into turtles, fishes, and a frog. *Mob DNA* **15**: 7. doi:10.1186/s13100-024-00318-9
- Heisserer C, Muller H, Jouan V, Musset K, Periquet G, Drezén J-M, Volkoff A-N, Gilbert C. 2023. Massive somatic and germline chromosomal integrations of polydnnaviruses in lepidopterans. *Mol Biol Evol* **40**: msad050. doi:10.1093/molbev/msad050
- Inoue Y, Takeda H. 2023. Teratom and its relatives – a cross-point of distinct mobile elements, transposons and viruses. *Front Vet Sci* **10**: 1158023. doi:10.3389/fvets.2023.1158023

- Inoue Y, Saga T, Aikawa T, Kumagai M, Shimada A, Kawaguchi Y, Naruse K, Morishita S, Koga A, Takeda H. 2017. Complete fusion of a transposon and herpesvirus created the Teratorn mobile element in medaka fish. *Nat Commun* **8**: 551. doi:10.1038/s41467-017-00527-2
- Inoue Y, Kumagai M, Zhang X, Saga T, Wang D, Koga A, Takeda H. 2018. Fusion of piggyBac-like transposons and herpesviruses occurs frequently in teleosts. *Zoological Lett* **4**: 6. doi:10.1186/s40851-018-0089-8
- Ivancevic AM, Kortschak RD, Bertozzi T, Adelson DL. 2018. Horizontal transfer of BovB and L1 retrotransposons in eukaryotes. *Genome Biol* **19**: 85. doi:10.1186/s13059-018-1456-7
- Kambayashi C, Kakehashi R, Sato Y, Mizuno H, Tanabe H, Rakotoarison A, Künzel S, Furuno N, Ohshima K, Kumazawa Y, et al. 2022. Geography-dependent horizontal gene transfer from vertebrate predators to their prey. *Mol Biol Evol* **39**: msac052. doi:10.1093/molbev/msac052
- Lampe DJ, Churchill ME, Robertson HM. 1996. A purified mariner transposase is sufficient to mediate transposition in vitro. *EMBO J* **15**: 5470–5479. doi:10.1002/j.1460-2075.1996.tb00930.x
- Le Rouzic A, Capy P. 2005. The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. *Genetics* **169**: 1033–1043. doi:10.1534/genetics.104.031211
- Levin HL, Moran JV. 2011. Dynamic interactions between transposable elements and their hosts. *Nat Rev Genet* **12**: 615–627. doi:10.1038/nrg3030
- Manni M, Berkeley MR, Seppely M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* **38**: 4647–4654. doi:10.1093/molbev/msab199
- McDaniel LD, Young E, Delaney J, Ruhnau F, Ritchie KB, Paul JH. 2010. High frequency of horizontal gene transfer in the oceans. *Science* **330**: 50. doi:10.1126/science.1192243
- McElreath R. 2020. *Statistical rethinking: a Bayesian course with examples in R and Stan*, 2nd ed. CRC Press, Boca Raton, FL.
- Metzger MJ, Paynter AN, Siddall ME, Goff SP. 2018. Horizontal transfer of retrotransposons between bivalves and other aquatic species of multiple phyla. *Proc Natl Acad Sci* **115**: E4227–E4235. doi:10.1073/pnas.1717227115
- Ono R, Yasuhiko Y, Aisaki K, Kitajima S, Kanno J, Hirabayashi Y. 2019. Exosome-mediated horizontal gene transfer occurs in double-strand break repair during genome editing. *Commun Biol* **2**: 57. doi:10.1038/s42003-019-0300-2
- Palazzo A, Caizzi R, Viggiano L, Marsano RM. 2017. Does the promoter constitute a barrier in the horizontal transposon transfer process? Insight from Bari transposons. *Genome Biol Evol* **9**: 1637–1645. doi:10.1093/gbe/evx122
- Palazzo A, Lorusso P, Miskey C, Walisko O, Gerbino A, Marobbio CMT, Ivics Z, Marsano RM. 2019. Transcriptionally promiscuous “blurry” promoters in Tc1/mariner transposons allow transcription in distantly related genomes. *Mob DNA* **10**: 13. doi:10.1186/s13100-019-0155-6
- Paulat NS, Storer JM, Moreno-Santillán DD, Osmanski AB, Sullivan KAM, Grimshaw JR, Korstian J, Halsey M, Garcia CJ, Crookshanks C, et al. 2023. Chiropterans are a hotspot for horizontal transfer of DNA transposons in Mammalia. *Mol Biol Evol* **40**: msad092. doi:10.1093/molbev/msad092
- Peccoud J, Loiseau V, Cordaux R, Gilbert C. 2017. Massive horizontal transfer of transposable elements in insects. *Proc Natl Acad Sci* **114**: 4721–4726. doi:10.1073/pnas.1621178114
- Peccoud J, Cordaux R, Gilbert C. 2018. Analyzing horizontal transfer of transposable elements on a large scale: challenges and prospects. *Bioessays* **40**: 1700177. doi:10.1002/bies.201700177
- Peddigari S, Li PW-L, Rabe JL, Martin SL. 2013. hnRNPL and nucleolin bind LINE-1 RNA and function as host factors to modulate retrotransposition. *Nucleic Acids Res* **41**: 575–585. doi:10.1093/nar/gks1075
- Pérez-Etayo L, González D, Vitas AI. 2020. The aquatic ecosystem, a good environment for the horizontal transfer of antimicrobial resistance and virulence-associated factors among extended spectrum β -lactamases producing *E. coli*. *Microorganisms* **8**: 568. doi:10.3390/microorganisms8040568
- Popa O, Dagan T. 2011. Trends and barriers to lateral gene transfer in prokaryotes. *Curr Opin Microbiol* **14**: 615–623. doi:10.1016/j.mib.2011.07.027
- Protasova MS, Andreeva TV, Rogaev EI. 2021. Factors regulating the activity of LINE1 retrotransposons. *Genes (Basel)* **12**: 1562. doi:10.3390/genes12101562
- R Core Team. 2024. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna <https://www.R-project.org/>.
- Reiss D, Mialdea G, Miele V, de Vienne DM, Peccoud J, Gilbert C, Duret L, Charlat S. 2019. Global survey of mobile DNA horizontal transfer in arthropods reveals Lepidoptera as a prime hotspot. *PLoS Genet* **15**: e1007965. doi:10.1371/journal.pgen.1007965
- Scarpa A, Pianezza R, Wierzbicki F, Kofler R. 2024. Genomes of historical specimens reveal multiple invasions of LTR retrotransposons in *Drosophila melanogaster* during the 19th century. *Proc Natl Acad Sci* **121**: e2313866121. doi:10.1073/pnas.2313866121
- Schaack S, Gilbert C, Feschotte C. 2010. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol Evol* **25**: 537–546. doi:10.1016/j.tree.2010.06.001
- Sheinman M, Arkhipova K, Arndt PF, Dutilh BE, Hermsen R, Massip F. 2021. Identical sequences found in distant genomes reveal frequent horizontal transfer across the bacterial domain. *eLife* **10**: e62719. doi:10.7554/eLife.62719
- Shi S, Puzakov MV, Puzakova LV, Ulupova YN, Xiang K, Wang B, Gao B, Song C. 2023. *Hiker*, a new family of DNA transposons encoding transposases with DD35E motifs, displays a distinct phylogenetic relationship with most known DNA transposon families of *IS630-Tc1-mariner (Ttm)*. *Mol Phylogenet Evol* **188**: 107906. doi:10.1016/j.ympev.2023.107906
- Silva JC, Loreto EL, Clark JB. 2004. Factors that affect the horizontal transfer of transposable elements. *Curr Issues Mol Biol* **6**: 57–71. doi:10.21775/cimb.006.057
- Smit A, Hubley R, Green P. 2013. RepeatMasker Open-4.0. <http://www.repeatmasker.org>.
- Steinegger M, Söding J. 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* **35**: 1026–1028. doi:10.1038/nbt.3988
- Van Etten J, Bhattacharya D. 2020. Horizontal gene transfer in eukaryotes: not if, but how much? *Trends Genet* **36**: 915–925. doi:10.1016/j.tig.2020.08.006
- Venner S, Miele V, Terzian C, Biémont C, Daubin V, Feschotte C, Pontier D. 2017. Ecological networks to unravel the routes to horizontal transposon transfers. *PLoS Biol* **15**: e2001536. doi:10.1371/journal.pbio.2001536
- Wallau GL, Ortiz MF, Loreto ELS. 2012. Horizontal transposon transfer in Eukarya: detection, bias, and perspectives. *Genome Biol Evol* **4**: 801–811. doi:10.1093/gbe/evs055
- Wallau GL, Capy P, Loreto E, Le Rouzic A, Hua-Van A. 2016. VHICA, a new method to discriminate between vertical and horizontal transposon transfer: application to the *Mariner* family within *Drosophila*. *Mol Biol Evol* **33**: 1094–1109. doi:10.1093/molbev/msv341
- Wang X, Liu X. 2016. Close ecological relationship among species facilitated horizontal transfer of retrotransposons. *BMC Evol Biol* **16**: 201. doi:10.1186/s12862-016-0767-0
- Wang S, Diaby M, Puzakov M, Ullah N, Wang Y, Danley P, Chen C, Wang X, Gao B, Song C. 2021. Divergent evolution profiles of DD37D and DD39D families of *Tc1/mariner* transposons in eukaryotes. *Mol Phylogenet Evol* **161**: 107143. doi:10.1016/j.ympev.2021.107143
- Wang S, Guan Z, Diaby M, Asare E, Ullah N, Jia W, Gao B, Yu D, Song C. 2024. Evolution of *Skipper (SK)*, a family of DD34E/*Tc1* transposons, in animals. *Biol J Linn Soc* **142**: 220–232. doi:10.1093/biolinnean/blad141
- Widen SA, Bes IC, Koreshova A, Pliota P, Krogull D, Burga A. 2023. Virus-like transposons cross the species barrier and drive the evolution of genetic incompatibilities. *Science* **380**: eade0705. doi:10.1126/science.ade0705
- Wu J, Meng L, Gaia M, Hikida H, Okazaki Y, Endo H, Ogata H. 2024. Gene transfer among viruses substantially contributes to gene gain of giant viruses. *Mol Biol Evol* **41**: msae161. doi:10.1093/molbev/msae161
- Zhang H-H, Peccoud J, Xu M-R-X, Zhang X-G, Gilbert C. 2020. Horizontal transfer and evolution of transposable elements in vertebrates. *Nat Commun* **11**: 1362. doi:10.1038/s41467-020-15149-4

Received January 15, 2025; accepted in revised form July 2, 2025.