

## Article

# Closing the Knowledge Gap: Horizontal Transfer of *Mariner* Transposons between *Rhus* Gall Aphids and Other Insects

Aftab Ahmad<sup>1</sup>, Xu Su<sup>2</sup>, AJ Harris<sup>3</sup> and Zhumei Ren<sup>1\*</sup>

<sup>1</sup> School of life Science, Shanxi University, 92 Wucheng Rd, Taiyuan Shanxi 030006, China; aftabahamad-khan003@gmail.com

<sup>2</sup> School of Life Sciences, Qinghai Normal University, Xining, 810008, China; xusu8527972@126.com

<sup>3</sup> South China Botanical Garden, Chinese Academy of Sciences, 723 Xingke Rd, Tianhe District, Guangzhou 510650; aj.harris@inbox.com

\* Correspondence: zmren@sxu.edu.cn

**Simple Summary:** Transposable elements (TEs) are mobile genetic elements which invades and multiply in host genomes. Besides vertical inheritance, TEs can be transferred from one species to another by phenomena called horizontal transfer of TEs (HTT). HT is crucial for TEs survival in genomes, but also a great disadvantage for host genomes, and recurrent HTT event between different host could shape and affect their genome architecture. HTT could be harmful for host genomes, but sometimes it can be useful too, and may play role in adaptive evolution of host genome. HTT is well reported in many eukaryotes but still there is a huge gap of knowledge about HTT in some organism. In this study close the knowledge gap about HTT regarding *Rhus* gall aphids, and reported multiple events of HTT involving *Rhus* gall aphids and other insects.

**Abstract:** Horizontal transfer of transposons (HTT) is an important source of genomic evolution in eukaryotes. The HTT dynamics is well characterized in eukaryotes, including insects, but there is a huge gap of knowledge about HTT regarding many eukaryotes' species. In this study we analyze the events of the HTT between *Rhus* gall aphids (Hemiptera) and other insects. We analyze the *Mariner*-like transposable elements (MLEs) that belong to *Rhus* gall aphids for possible HT events. The MLEs have patchy distribution and have high similarity over the entire length of element among insects MLEs from different orders. We selected representative sequences from the *Rhus* gall MLEs and identified five events of HT between MLEs of *Rhus* gall aphids and other insects from five different orders. We also found multiple HTT events among the MLEs of insects from the five orders which demonstrate that these *Mariner* elements have been involved in recurrent HT between these six orders of insects. Our current study closes the knowledge gap of HTT and reports the events between *Rhus* gall aphids and other insects for the first time. We believe that this study about HTT events will help to understand the evolution and spread of transposable elements in the genomes of *Rhus* gall aphids.

**Keywords:** horizontal transfer; *Rhus* gall aphids; *Mariner* transposable elements

## 1. Introduction

Transposable elements (TEs) are mobile DNA sequences that can translocate in the host genome and replicate their number. This ability of TEs allows them to invade virtually all kinds of organisms, from prokaryotes to higher vertebrates, including those of humans [1] and plants [2,3]. However, the mobility of TEs can lead to significant adaptive changes by promoting chromosomal rearrangements such as segmental duplications, deletions, and inversions through phenomena such as non-allelic homologous recombination [4,5]. On the other hand, TEs expansion in the host genomes can also be harmful, leading eukaryotes to evolve various defence and regulatory mechanisms [6].

Like all other nuclear genes, TEs can also be inherited vertically, from parents to offspring; however, TEs can be transmitted among different organisms through a

phenomenon known as horizontal transfer (HT) [7-9] which can have an immediate or delayed effect on host organism [10]. The exact mechanisms and pathway of HT of TEs (HTT) is not well understood, but most certainly, it is related to the mobile nature of TEs, as the normal genes, in comparison, are much more rarely found to be transferred horizontally in eukaryotes [11-15]. HTT is typically inferred when the nucleotide divergence between TE copies from two distantly related hosts is much lower than expected due to vertical inheritance since the last common ancestor of the two hosts [16].

TEs are classified into two classes based on their transposition mechanism: class I elements or retrotransposons move by a copy-paste mechanism and class II elements, or DNA transposons, move through a DNA intermediate [17]. HT appears to be more persistent in DNA transposons of which the *Tc1/Mariner* superfamily is the most common type to be transferred horizontally, and many studies showed the prevalence of *Tc1/Mariner* TEs in HTT among diverse animals taxon [18-21]. The underlying mechanism and vectors involved in HTTs is unclear; however, recent studies speculated hypotheses and indirect evidence of HTT species [22]. With the advancement and invention of new technologies, large scale analysis of different organisms genomes, several host and parasitic features can be considered to facilitate the occurrence of HTT, including the occurrence of some parasites with multiple host species, while a symbiotic association between different species could also leverage the phenomena of HTT among different species [22]. TE research community has evaluated many vectors with little or no success while some studies have hypothesized that parasites can mediate transfer of TEs from one species to another while some recent studies proposed that viruses can be the possible vectors that mediate HTT [23-26].

Class Insecta has one of the most extensive species diversity on earth and represents one of the main eukaryotic evolutionary branches. Insects genomes have been studied extensively for the detection of TEs, and several HTTs events have been reported in insects; [27-31], including the first case of HTTs of P elements in *Drosophila* [16]. A recent study reported more than two thousand HTT events among 195 insects species and closed significant gaps related to insects' HTT occurrence [20]; however, there were no HTT events reported related to gall-forming aphids due to the unavailability of their genomes and TEs in the public database. Though HTT is well studied in most insects and some aphids species [9,32,33], there is no information present about the existence of HTT events in the genome of galling aphids. We recently uncovered the existence of *Mariner* transposons in seven species of *Rhus* gall aphids [34], while phylogenetic analysis of detected TEs showed patchy distribution, which predicted the occurrence of HT among *Rhus* gall aphids and organisms belonging to different insect orders. In the present study, we performed a detailed analysis of the of *Mariner*-like transposable elements (MLEs) of *Rhus* gall aphids and unveiled several events of HTT for the first time in *Rhus* gall aphids and any other galling aphids.

*Rhus* gall aphids (Aphidoidea: Eriosomatinae: Fordini), are sap-feeding aphids and parasitized plant hosts of the *Rhus* genus. In contrast to other aphids, *Rhus* gall aphids are not very harmful to the host plant and do not damage their host plant, while recent studies also reported a symbiotic association and complex and nutrient exchange between these gall-forming aphids and plants [35]. *Rhus* gall aphids' life cycle alternates between two hosts, i.e., *Rhus* genus of plants and a few moss species. They induce gall in the leaves of their primary host plant, *Rhus* species (Anacardiaceae) and live inside the gall for several generations. The galls formed by these aphids are rich in tannin, which can be used in medicines, tanning, and military industry; hence they have practical economic importance [35-37].

In our previous study, we reported the diversity of MLEs in seven species of *Rhus* gall aphids. The seven species of aphids belonged to six major genera of *Rhus* gall aphids [36]. The *Mariner* family of DNA transposons is very well known to perform recurrent and successful horizontal transfers, supported by previous studies' conclusions [18,38,39]. The present study aimed to analyze all MLEs detected in *Rhus* gall aphids for possible HTTs between the *Rhus* gall aphids and other insects. This study will present the first report of

HTTs between *Rhus* gall aphids and other insects belonging to different orders of Class Insecta.

## 2. Materials and Methods

### 2.1. Data availability

We used 121 *Mariner* transposons sequences as queries in databases, along with accession numbers in Supplementary file S1. The qualified MLEs for the HT analysis in this study constitute the set detected in seven species of *Rhus* gall aphids and is available on GenBank as of August 2021. We built species phylogenetic tree based on 20 genes downloaded from GenBank (Supplementary file S2), while *Mariner*-like elements for the target insect's species were extracted from genome assemblies at NCBI after BLASTn search (Supplementary file S3).

### 2.2. Identification and annotation of MLEs in targeted species

To infer any possible transfer of *Rhus* gall MLEs with other insects, we followed the widely used two-step approach in reference to the previous studies [40]. Homology-based approaches were followed to find and extract similar nucleotide sequences from the genome of the target species. For this purpose, we used the transposable elements database i.e., RepBase [41], and the whole genomes database i.e., NCBI GenBank. RepBase searches for homologous sequences were done using the "CENSOR" tool [42], implementing Repeatmasker in RepBase [41] using the default parameters against the whole database. We retrieved few sequences from the Repbase as a result blast in the whole database, and extract sequences that follow our designed criteria (Sequence similarity  $\geq 85\%$ , and Query coverage  $\geq 80\%$ ), while blastn searches at NCBI were done with default parameters MLEs of *Rhus* gall aphids in queries. We extracted the resulting blastn hit produced by the qualified MLEs from the genome of the insect's species and were manually analyzed for their terminal inverted repeats (TIRs). To confirm the placement of the extracted sequences in the *Mariner* family, we searched the conserved motifs and domain of the sequences i.e., Helix-turn-Helix HTH DNA binding motif and DDE catalytic domain using CD-search [43,44] and motif search online ([www.genome.jp/tools/motif/](http://www.genome.jp/tools/motif/)). The open reading frame (ORF) was also predicted using ORF finder (implemented in Geneious) and the MLEs were annotated using Geneious v11.1.

Overall, MLEs nucleotide sequence's identity between *Rhus* gall aphids and other insects included in the present study was calculated using BLASTN. We used alignments with query coverage  $\geq 80\%$  and nucleotide sequences shorter than this was filtered out from the analysis to avoid false-positive results.

### 2.3. Phylogenetic analysis

We achieved the phylogenetic analysis in four steps: First, we aligned all the 121 MLEs of *Rhus* gall aphids with MAFFT v.7.1.1 [45] with the default parameter. The alignment was manually curated, followed by the construction of ML phylogenetic tree. We used jModelTest v2.1.10 [46,47] to select the best evolutionary model that fitted adequately and resulted in a good tree. Second, we aligned the qualified sequences from *Rhus* gall aphids and the extracted sequences from targeted species i.e., other insects orders using MAFFT implemented in Geneious. Alignment was trimmed manually, and MLEs phylogenetic tree was constructed with IQTREE using model GTR+I+G, suggested by jModelTest.

In the third step, we constructed the species phylogenetic tree of seven species of *Rhus* gall aphid and the targeted insects belonging to four different orders, using fifteen mitochondrial genes and five nuclear genes. We aligned all the 20 genes with MAFFT implemented in Geneious with default parameters. The aligned sequences were manually curated, and ML phylogenetic tree was constructed with IQTREE using model GTR+I+G, suggested by jModelTest. The tree was visualized and modified using Figtree v1.4.4 software (<http://treebioedacuk/software/figtree/>), and posterior probability was used as

statistical support for each branch. Last, we estimated the divergence time between the species to infer the HT between the species, especially the five different clades representing each order. Divergence time was estimated using TimeTree online at (www.time-tree.org).

#### 2.4. Estimating the minimal number of horizontal transfer events

To infer multiple events of HTs, we analyzed and estimated the minimum number of HTT events in our present data, taking into consideration the possibility that a single HT event may be sufficient to explain several cases of shared MLEs through horizontal transfer if they happened in the common ancestor of recently diverged species. We evaluate and compare the species tree and MLEs tree, considering all the nodes in the tree and predicted one HTT event if the descending clades sharing the same MLE were connected by a common ancestor.

Strictly speaking, we inferred and concluded that most of the HT events took place in ancestor branches, which passed to the descendant through a vertical transfer with slight divergence under natural selection, but for simplicity, we approximated that species of our sample (*Rhus* gall aphids) whom MLEs has the closest similarity with MLEs of the other species were potentially involved in HTT. We estimated the confirmed minimal HTT events between *Rhus* gall aphids and other insects following this procedure.

### 3. Results

In this study, we analyzed the MLEs we detected in our previous study for HTT events between *Rhus* gall aphids and other insects from belonging to different orders. To further understand the origin of evolution and inheritance of MLEs in *Rhus* gall aphids, we did a detailed comparative phylogenetic analysis of the *Rhus* gall aphids and other insects MLEs. A significant number of *Rhus* gall aphids MLEs from different lineages showed high pairwise identity with the TEs of phylogenetically distantly related insects species. The unique identities between TEs of *Rhus* gall aphids and other insects prompted us to have a systematic search for the HTT events involving *Rhus* gall aphids *Mariner* transposons. Many studies had successfully documented thousands of HTT events among the species of class Insecta [20,38], but due to the lack of TEs data about *Rhus* gall aphids, there was no evidence of HTT between *Rhus* gall aphids and other insects. We followed a detailed two-step approach to discriminate the HTT event in the present studies, based on homology or nucleotide sequence identity and species phylogenetic tree comparisons with MLEs phylogenetic tree. We have found several events of HTT between *Rhus* gall aphids and species from five different orders of class Insecta.

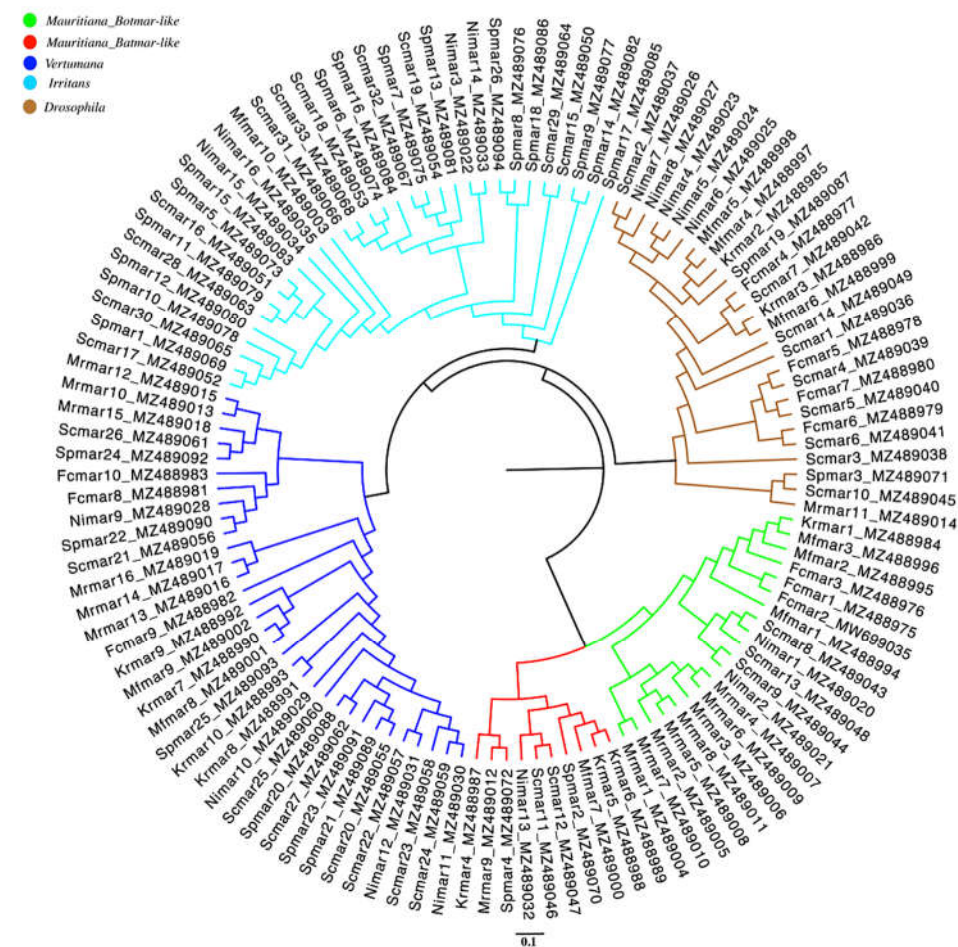
#### 3.1. Phylogenetic relationship of the *Rhus* gall aphids MLEs

We determined the phylogenetic relationship of all the MLEs of *Rhus* gall aphids analyzed in the present study with already known MLEs from the Tc1/Mariner Superfamily in our previous study [34]. All the MLEs in this study belonged to four subfamilies of the *Mariner* family, i.e., *Mauritiana*, *Irritans*, *Vertumana*, and *Drosophila*. Phylogenetically we further classified MLEs of the *Mauritiana* subfamily into two sub-lineages based on sequence similarities, i.e., *Botmar-like* elements reported in *Bombyx mori* for the first time and *Batmar-like* elements found in *Bactrocera tryoni*. Phylogenetic distribution of all MLEs used in this study can be seen in (Fig. 1) across the seven species of *Rhus* gall aphids among which *Mauritiana*, *Vertumana*, and *Drosophila* subfamily are distributed in all the seven species while MLEs from the *Irritans* subfamily is present in four of the seven species.

In conclusion, the phylogenetic analysis shows the distribution of DNA transposons of the *Mariner* family in all the studied *Rhus* gall aphids species used for HTT analysis in the present study. All the MLEs reported in seven *Rhus* gall aphids showed patchy distribution and are not congruent with species phylogenetic tree, i.e., MLEs from *Irritans* are present in only four species, which might also reflect HTT event within *Rhus* gall aphids species. To infer the horizontal transfer events, we selected representative sequences from



each lineage of *Rhus* gall aphids and searched identical sequences in other insects species following homology based approaches.



**Figure 1.** Phylogenetic tree and lineage information of all MLEs of *Rhus* gall aphids for HTT analysis. Tips of branches indicate MLE name followed by GenBank accession number.

3.2. Selection of the MLEs Representative Sequence for HTT

Initially, to infer the HTT event between transposons of *Rhus* gall aphids and other species, specifically insects, we include all the 121 *Rhus* gall MLEs as queries (Supplementary file S1) in NCBI Blastn against the genomic databases and extensively search the highly similar DNA sequences in distantly related genomes. Among 121 sequences, forty produced an excellent hit against genomes of other species with query coverage >90% and similarity ≥80%. Although it is possible to include all copies of MLEs detected within one species, but most sequences from the same lineages result in similar hits, so it is much simpler and quicker to use only a few representatives. As one sequence from the same lineage (subfamily) is enough to infer the desired result, we chose one representative from each lineage or more than one where required. We selected the complete sequence from each lineage, resulting in good Blastn hits, and truncated sequences were discarded from the MLEs. To further simplify our search and choose the best representative among the complete copies from the same lineages, we chose the sequence which results in the best Blastn hit (high similarity and query coverage). For instance, Fcmar1, Fcmar2 and Fcmar3 belong to the same lineage of MLEs in *Floraphis choui* and resulted in similar hits, so we discarded Fcmar3 as it was incomplete, and Fcmar2 was selected as it results best hit against the genome of other species at NCBI Blastn search. Similar rules were applied to MLEs of all seven species, and fifteen MLEs were selected.

Furthermore, to avoid repetitions and false-positive HTT events among *Rhus* gall aphids and other non-related species, only one orthologous sequence (MLE) was selected as representative from all the MLEs in seven different species of *Rhus* gall aphids following the rules explained above. For instance, Fcmar2, Krmar1, Mfmar1, Nimar1, and Scmar2 belong to the same lineage (Fig. 1), i.e., the *Botmar*-like elements of *Mauritiana* sub-family result in similar hits in BLASTn, so only one Fcmar2 with the best hit was selected as representative for this lineage. Finally, we selected eight MLEs from four different sub-families involved in the HTT event between *Rhus* gall aphids and other insects. We followed the above-discussed criteria again to search similar MLEs sequences in RepBase, which could be involved in HTT with *Rhus* gall aphids.

### 3.3. Inference of horizontal transfer between *Rhus* gall TEs with other insects

Horizontal transfer event is well documented in many species of insects, but there are no reports of HTT of transposons in *Rhus* gall aphids. HT of transposons can be conferred either based on DNA sequences similarities or phylogenetic incongruences of TEs compared to neutrally evolving vertically transmitted genes or by combining both methods. Sequences of distantly related species having query coverage and similarity >90% are considered to be horizontally transferred [38], while sequences sharing terminal inverted repeats (TIRs) similarity >90% could also be a result of HTT event [17]. We also followed the commonly used two-step approach to unveil the phenomena of HT in *Rhus* gall aphids and other insects.

### 3.4. Inference of HTT based on nucleotides sequence similarities

The first step to uncover the possible events of HT in *Rhus* gall aphids was based on DNA sequences similarities with other non-related species. We search for homologous sequences of the *Rhus* gall MLEs in Repbase using the "censor" tool implemented in the Repbase database with default parameters, and BLASTn searches at NCBI database using default parameters. Deleted and truncated small copies can lead to false-positive results; we extracted copies only with query >90% and similarity >80%. We found very few good hits in Repbase as per the designed criteria for the study, but retrieved many homologous sequences from Blastn searched at NCBI. Many sequences that belong to the same lineages within the same species result in similar hits, but we selected sequences having the highest DNA similarity throughout the length as a representative sequence from each lineage as explained above. The top-ranking results from NCBI Blastn, i.e. query coverage >90%, %identity >85% with lowest E-value and higher bit score, are shown in Table 1, while top-ranking result from RepBase, i.e., query coverage >90%, similarity with consensus sequence >90%, Pos-value and higher bit score, are shown in Table 2. The detected sequences in other insects display higher nucleotide identity, which exceeds the expected identity values when comparing transposable elements in distantly related species. We found many sequences in NCBI blastn result, and RepBase wuBlast result, very similar to *Rhus* gall transposons as per criteria (identity  $\geq 85\%$ , and query  $\geq 95\%$ ) in a total of seventeen insects species belonging to five different orders, i.e. (Hymenoptera, Diptera, Coleoptera, Lepidoptera, and Neuroptera). *Rhus* gall aphid's species belong to the order Hemiptera of class Insecta, so the current study represent HTT events among six different order of Class Insecta (Table.1 & 2).

**Table 1.** The top-ranking results from NCBI Blastn, i.e., query coverage >90%, %identity >85% with lowest E-value and higher bit score along with target species and their respective order.

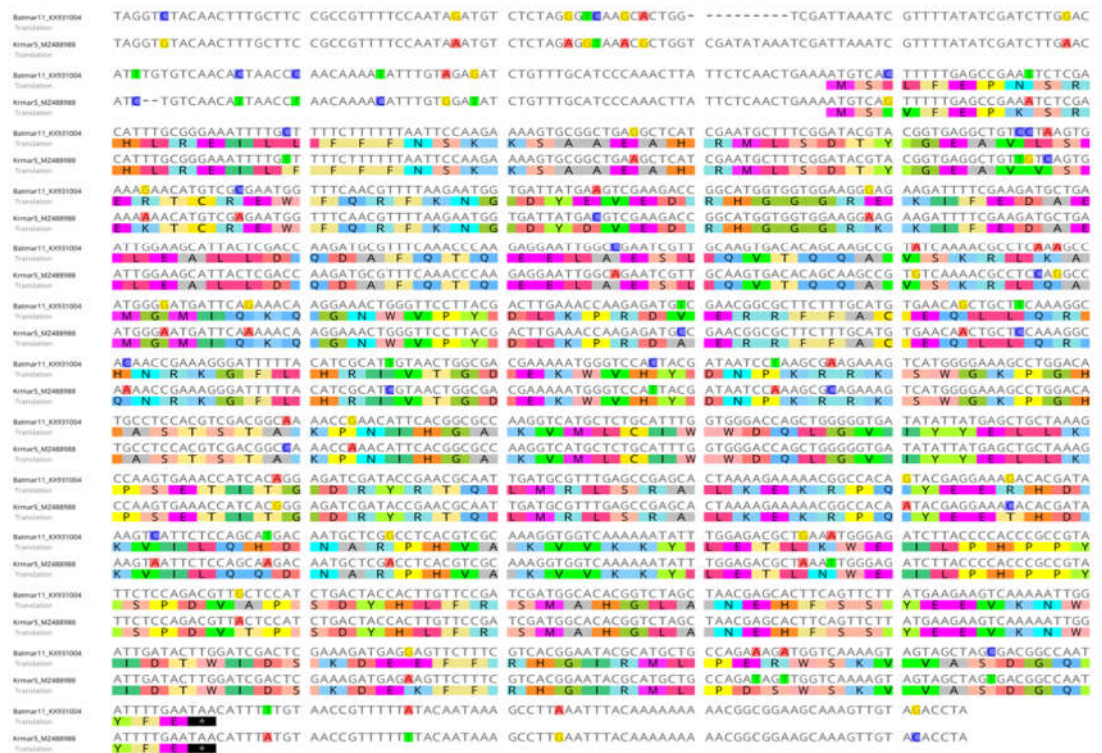
Rhus gall aphid Species	Rhus gall Aphids MLE	Target species element	Species Order	Target accession	NCBI Blast Alignment parameters			
					% Query coverage	%Identity	E-value	Bit score
Floraphis choui	Fcmar2	Myrmica ruginodis	Hymenoptera	AY652426	99	90.10	0	1657
		Bombus campestris	Hymenoptera	HG995146	100	87.66	0	1489
		Nomada fabriciana	Hymenoptera	OU015690	99	87.48	0	1476
		Ocypus olens	Coleoptera	OU343056	99	87.55	0	1469
		Bombus terrestris	Hymenoptera	OU342929	99	87.39	0	1469
		Bombus pascuorum	Hymenoptera	HG995272	99	87.25	0	1454
		Osmia bicornis	Hymenoptera	OU015504	99	85.60	0	1339
Kaburagia rhusicola	Krmar4	Tinea trinitella	Lepidoptera	HG992316	99	85.39	0	1315
	Krmar5	Bactrocera tryoni	Diptera	KX931004	100	94.6	0	1750
Meitanaphis flavo-gallis	Mfmar7	Bactrocera tryoni	Diptera	KX931004	100	93.57	0	1720
Melpahis rhois	Mrmar11	Melicta athalia	Lepidoptera	HG992203	99	96.80	0	2185
		Bactrocera tryoni	Diptera	KX930994	100	95.67	0	2108
		Chrysoperla carnea	Neuroptera	FR997756	99	91.56	0	1783
		Tinea trinitella	Lepidoptera	HG992328	99	90.98	0	1307
Schlechtendalia chinensis	Scmar11	Bactrocera tryoni	Diptera	KX931004	100	93.64	0	1910
Schlechtendalia peitan	Spmar2	Bactrocera tryoni	Diptera	KX931004	100	94.49	0	1910

**Table 2.** The top-ranking results from RepBase, i.e., query coverage >90%, %identity >85% with consensus sequence, Pos value and higher bit score along with target species and their respective order.

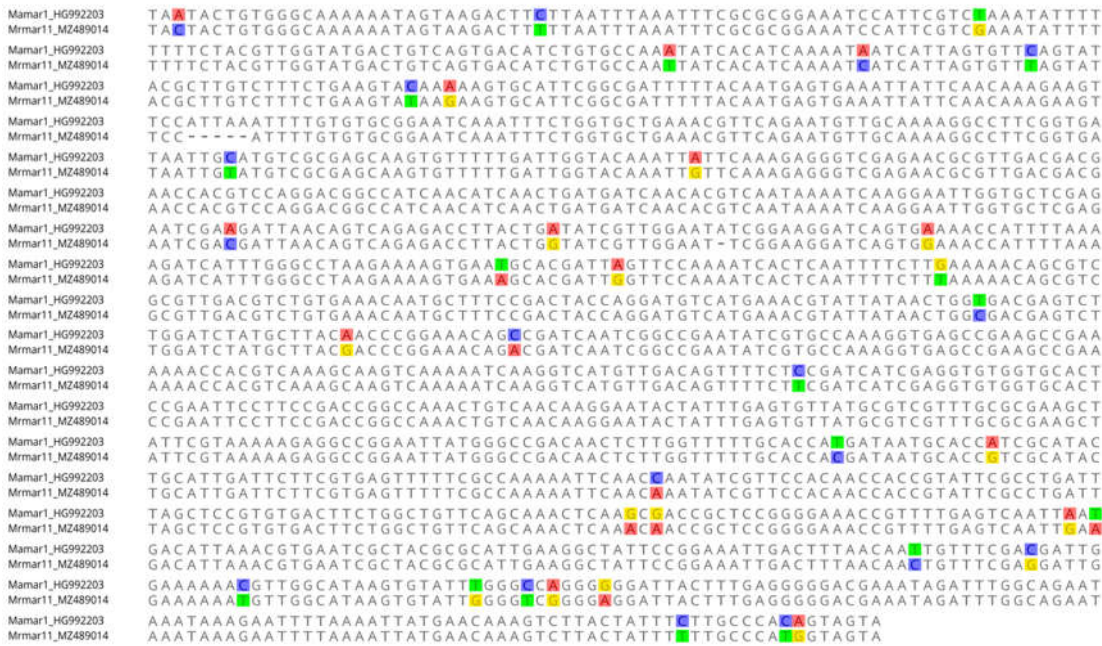
Rhus gall MLEs	Rhus gall aphids	Target Species name	Order	Target MLE name	Query	Similarity to (Consensus)	Pos	Bit Score
Mrmar16	Melaphis rhois	Herpegnathos saltator	Hymenoptera	Mariner-33_HSal	92%	0.8307 (>97%)	1.8131	7128
Fcmar2	Floraphis choui	Acromyrmex echinator	Hymenoptera	Mariner-2_AEc	99%	0.8925 (89.2%)	1.5227	8806
Krmar4	Kaburagia rhusicola	Drosophila elegans	Diptera	Mariner-7_DEL	95%	0.8003 (~96%)	1.5844	6914
Krmar5	Kaburagia rhusicola	Acromyrmex echinator	Hymenoptera	Mariner-18_AEc	100%	0.8213 (~96%)	1.6214	7732
Mrmar11	Melaphis rhois	Solenopsis invicta	Hymenoptera	Mariner-5_SIn	100%	0.8088 (>97%)	1.8926	6926

As expected for TEs, in some cases, the extent of sequence identity was observed throughout the length of elements of TEs, including TIRs, which strongly support the hypothesis of the HTT event. Some of the examples which support this can be seen in (Fig. 2 & 3 and Table 1 & 2) in which the DNA sequence similarity between TEs of *Rhus* gall aphid species with distantly related species of different orders are more than 96%, while the amino acid sequences similarities are also more than 96%.





**Figure 2.** Pairwise alignment of nucleotide and amino acids sequences of *Rhus* gall MLE (Krmars5) from *Kaburagia rhusicola* and *Bactrocera tryoni* MLE (Batmar11) with intact ORF for transposase, showing high pairwise similarity (94.6%) throughout nucleotide length and pairwise similarity (96.1%) for throughout amino acid sequence, and >96% similarity between the TIRs.



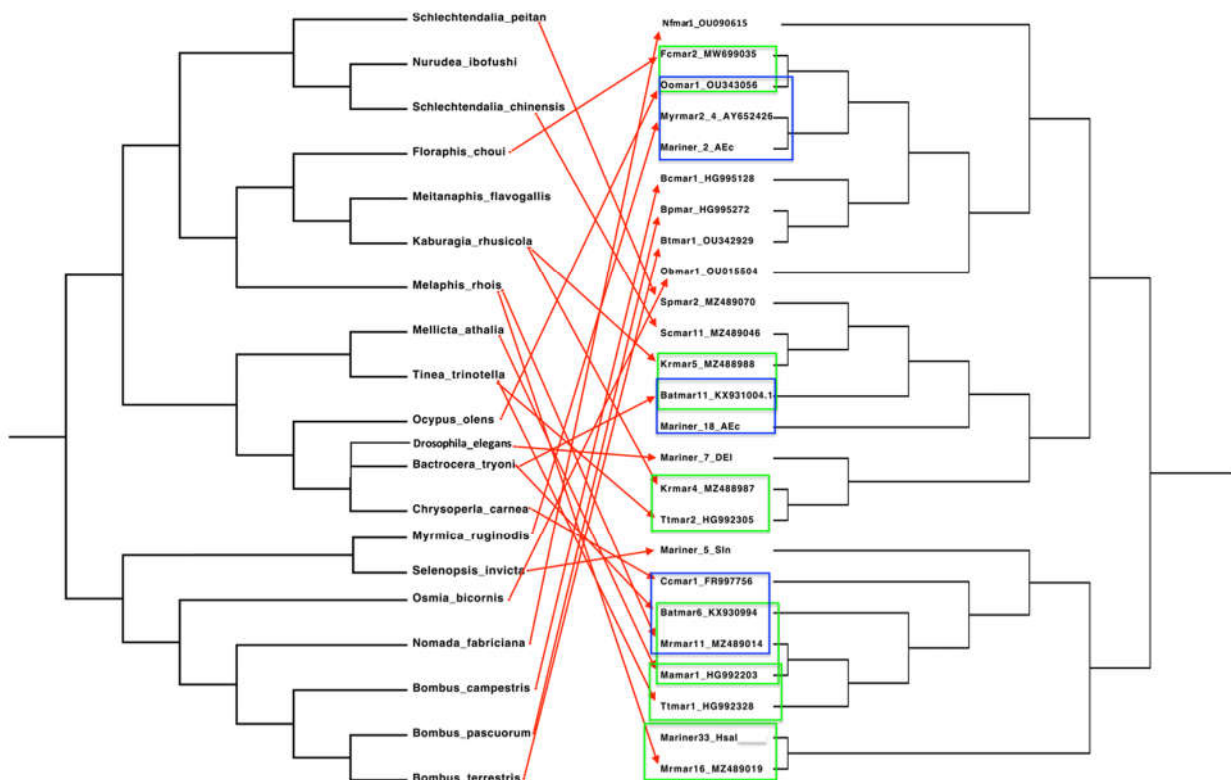
**Figure 3.** Pairwise alignment of nucleotide sequences of MLEs from *Rhus* gall aphid *Melpahis rhois* (Mmar11) and butterfly *Mellicta athalia* (Mamar1), showing high similarity (96.80%) throughout the length of sequences, with no intact ORF, and more than > 93% similarity for TIRs.



### 3.5. Phylogenetic analysis of HTT among *Rhus* gall aphids and other insects

To further confirm the phenomenon of HTT between *Rhus* gall aphids and five distantly related insects' orders, we construct the transposons tree (Fig. 4a) of representative sequences of *Rhus* gall aphids, and the other insects MLEs recovered from NCBI GenBank and Repbase. We also constructed the species tree (See Fig. 4b) of all included species in this study by selecting 20 highly conserved orthologous genes, resulting in a good quality species phylogenetic tree. We couldn't retrieve the mitochondrial and nuclear gene for *Drosophila elegans* and *Herpegnathos saltator* due to unavailability in public databases. *Drosophila elegans* was branched manually in species tree using its divergence information from Timetree online, while *Herpegnathos saltator* position can be represented by other ant species i.e., *Acromyrmex echinator*. We concatenate the selected 20 genes of the species; 15 were mitochondrial genes, i.e., 12S rRNA, 16S rRNA, ATP6, APT8, COX1, COX2, COX3, Cyt-b, ND1, ND2, ND3, ND4, ND4L, ND5, and ND6, while five are nuclear genes, i.e., Long-wavelength rhodopsin resistance gene (*lwrh*), wingless (*wnt-1*), Elongation factor 1-alpha (*EF1-alpha*), Histone (*H3*) and 18S rRNA gene (See Supplementary file 2).

We compared the MLEs tree to the species tree of the *Rhus* gall aphids and other insects included in this study. These comparisons clearly showed at least five events of HT among *Rhus* gall aphids and other insects belonging to five different orders and few events of HTT within the insects of the other five orders. Seven MLEs from three subfamilies and four lineages of *Rhus* gall aphids clustered with the MLEs of insects from distantly related orders provide strong evidence of HTT. For instance, a *Rhus* gall MLE from *Botmar*-like lineage of *Mauritiana* subfamily (Fcmar2\_MW699035) clustered with MLE of *Ocypus olens* (Oomar1\_OU343056) from Coleoptera order, *Myrmica ruginodis* (Myrmar2\_AY652426), and other insects elements of *Bombus* genus of order Hymenoptera. However, *Rhus* gall aphids (order: Hemiptera) are very distantly related to these species and diverged between 323-392 MYA from Coleoptera and Hymenoptera.



**Figure 4.** Comparison of species tree with MLEs tree to infer HTT events between *Rhus* gall aphids and other insects. A. Left side species tree of seven *Rhus* gall aphids and thirteen insects belongs to five different orders, constructed based on 20 genes. B. Phylogenetic tree of *Rhus* gall MLEs and extracted MLEs from other insects' species. The green rectangles indicate HTT events between *Rhus* gall MLE and other insects of a different order. In contrast, the blue rectangle shows HTT events

involving *Rhus* gall aphids and between insects of different orders in this study. Red arrows indicate the position of each MLE of each species and describe the highly patchy distribution of the MLEs.

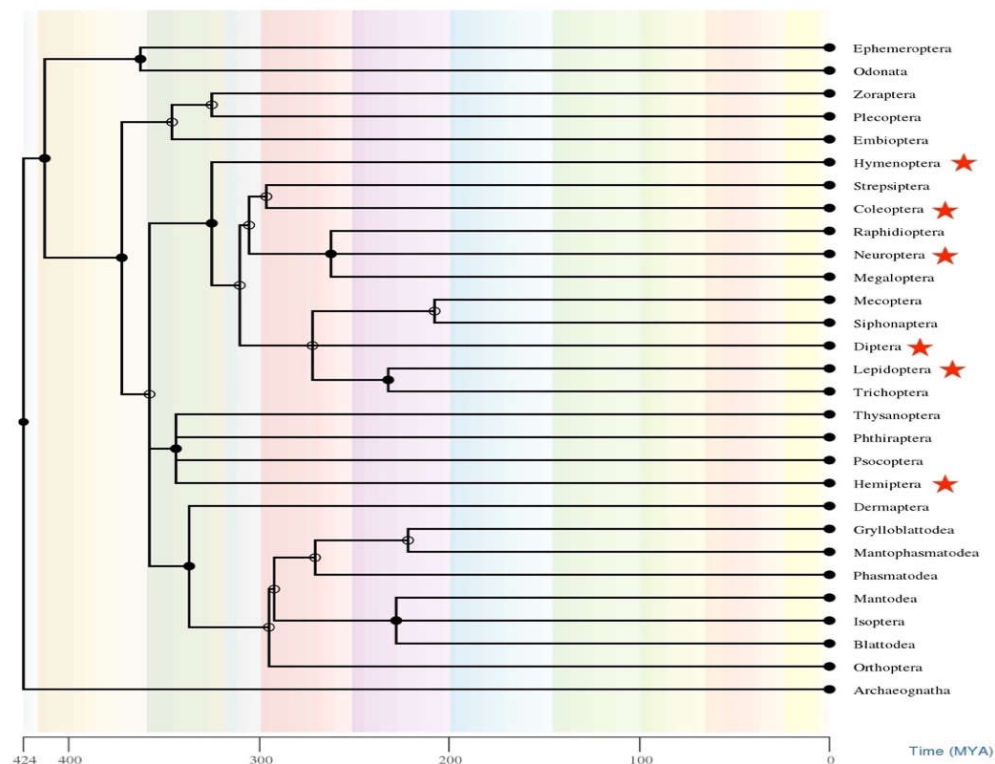
Several HTTs events have also been observed within the other insect's orders included in this study. For instance, a Botmar-like MLE from *Mauritiana* subfamily (Oommar1\_OU343056) extracted from *Ocypus olens* of Coleoptera order, nested with the MLEs of *Myrmica ruginodis* and *Acromyrmex echinator* of order Hymenoptera, which reflects multiple events of HTT of this MLE between *Myrmica ruginodis*, *Ocypus olens* and *Floraphis choui* (Fig. 4b). Another such event of HTT can be observed in a Batmar-like MLE of *Mauritiana* subfamily in *Acromyrmex echinator* (Mariner-18\_Ace) of Hymenoptera order, which nested with *Bactrocera tryoni* MLE (Batmar-11\_KX931004) of order Diptera, which also reflect multiple events of HTTs of these elements between *Acromyrmex echinator*, *Bactrocera tryoni*, and *Kaburagia rhusicola*. Furthermore, one MLE of *Drosophila* subfamily (Ccmarr\_FR997756) extracted from *Chrysoperla carnea* (Order: Neuroptera) have been nested with Batmar6\_KX930994 of *Bactrocera tryoni* (Order: Diptera), Mamar1\_HG992203 of *Mellicta athalia* and Ttmar1\_HG992328 of *Tinea trinotella* from order Lepidoptera, which indicates several events of HTT by this MLE between *Chrysoperla carnea*, *Bactrocera tryoni*, *Mellicta athalia*, *Tinea trinotella*, and *Melaphis rhois* (Fig. 4b).

We couldn't discriminate the HTT events within *Rhus* gall aphids species, as all the aphids in this study are closely related and belong to same family Eriosominae, with closed divergence time. As recently diverged species share a high nucleotide similarities in their sequences, which also make it difficult to detect HT in them based on nucleotide identities. Though the *Rhus* gall MLEs shows patchy distribution among the *Rhus* gall aphids but did not produce positive HT signals.

In conclusion of the above results, *Rhus* gall aphids transposons have striking identities with transposons of distantly related species and clustered together with them in the MLEs tree. Meanwhile in the species tree, based on nuclear genes and mitochondrial genes sequences all the *Rhus* gall aphids clustered apart. These unexpected sequences similarities of galling aphids MLEs with other insects from different orders, and the uneven distribution of MLEs in the phylogenetic tree is assumed as solid evidence of HTT between *Rhus* gall aphids and other insects.

### 3.6. Estimation of Divergence time

All the insects in the tree shared a common ancestor and belonged to the same class, we estimated the divergence time of each species and the order of insects included in this study (Fig. 5). Divergence times were estimated to know the distances between the species which will help to discriminate the HT events between the sequences. Species having common ancestors, and diverged long time ago tend to accumulate more mutations and changes in their nucleotide sequences due to evolution. We estimated divergence time of all species using Timetree online at (<https://www.timetree.org>) and the divergence time between *Rhus* gall aphids with all other insects in the tree were more than 300-350 Myr, with an average 325 Myr. Comparatively the MLEs sequences of *Rhus* gall aphids and other insects showing contrasting nucleotide similarities between them, which is not possible for the neutrally evolving genes diverged so long time ago. For example, *Schlechtendalia chinensis* diverged 350-375 Myr from *Bactrocera tryoni* nested very apart in species tree, but their MLEs showed 96% nucleotide sequence similarities between them and nested as sister sequences in MLEs tree. The very high divergence time between studied species and closed MLEs sequences similarities further supports the claim HTT events between these species.



**Figure 5.** Phylogenetic tree constructed with time tree online, showing the relation and divergence time (MYA) between the insect's orders. Stars\* at the tip of the branch indicate the orders included in the study, which are involved in HTT.

#### 4. Discussion

HTT is a well-known and reported phenomenon documented in many metazoans, including insects. Though the exact mechanism of HTT is poorly understood, the geographic proximity and host-parasite interactions might help in the exchange of genetic material in distantly related species [8,48]. To date (visited 1<sup>st</sup> January 2022), there are 5689 cases of HTT that have been reported in the Horizontal transfer of transposons database (HTT-DB) [49]. Among all the HTT events reported in HTT-DB, *Tc1/Mariner* Superfamily of DNA transposons contribute to most cases, i.e., 2523 out of 4271 DNA transposon HT events. HTT is well reported in most insects. Recent studies uncovered thousands of HTT events in class Insecta [20], while other studies reported thousands of HTT events in vertebrates [50]. However, there is no report of HTT in *Rhus* gall aphids to date due to the unavailability of TEs data in these aphids. We recently reported *Mariner* transposons in *Rhus* gall aphids for the first time. We followed the previous findings to uncover the events of HTT among *Rhus* gall aphids and other insects and closed the knowledge gap about HTT in galling aphids.

Methodologically, it is not easy to infer HTT events. Several tools and methods can be used; for example, the VHICA tool infers HTT events based on codon usage analyses and comparison synonymous non-synonymous substitutions rate [19]. Unfortunately, there are limitations to these automatic tools. These perform poorly if the divergence between the species in which HTT is inferred increases, leading to substitution saturation which causes loss of the phylogenetic signal [51]. In this study, the divergence time between *Rhus* gall aphids and other species for which HTT was inferred is huge (>300Myr), and the methods based on phylogeny and genetic distance seem to be more suitable, which is also suggested by previous studies [52,53]. We followed a two-step approach explained above and uncovered a few events of HTTs between *Rhus* gall aphids and eleven other species of insects belonging to five orders of class Insecta. HTT events detected from more than one species of the same order for the same MLEs were considered



single HTT events. Among the seven species of *Rhus* gall aphids, TEs from three species seem to have undergone HTT events.

We have inferred HTT events involving six DNA transposons of the Mariner family in *Rhus* gall aphids. Our findings suggest that HTT events between *Rhus* gall aphids and distantly related insects might have occurred several times. Our finding also suggests the MLEs involved in *Rhus* gall aphids' HTT events have also undergone many HT between other orders of insects. Besides, the common ancestor of all the insects, *Rhus* gall aphids (Order: Hemiptera), is quite distantly related from other insects in this study, and the HTT scenario is very clear from the results; however, it is challenging to infer both the direction and the vector of the HTT events described in this study.

Interestingly all the putative cases of HTT detected in this study involved 21 insect species from five different orders. In this respect, Order Hymenoptera seems to be the preferred order in exchanging MLEs since seven different bee species belonging to 3 different genera and five ants species from different genera are putatively involved in HTT events of four *Rhus* gall aphids *Mariner* elements (namely Fcmar2, Krmar4, Krmar5, Mrmar16). Insects from order Hymenoptera have been involved in many HTT events in previous studies [40], which is supported by the present study. While, four of the 21 insects involved in HTT events in the study belongs to order Lepidoptera and four from order Diptera, respectively with five *Rhus* gall elements (Krmar4, Krmar5, Spmar2, Scmar11, Mrmar11). A recent study in insects found that Lepidoptera order has been the hotspot of HTT in insects [21], while fruitflies from order Diptera were also suggested as good horizontal transfer candidates [38,49,54,55]. At the same time, one species, each from beetles (Order: Coleoptera) and Laecwings (order: Neuroptera) were, also seemed to be involved in the HTT event with *Rhus* gall aphid *Mariner* elements (Fcmar2, Mrmar11). All the *Mariner* elements involved in HTT events in this study belong to two subfamilies in which elements of the *Mauritiana* subfamily seem to be dominantly involved in HTT events.

Though there is a patchy distribution of MLEs among the *Rhus* gall aphids and do not follow the species tree, it is difficult to infer the HTT events among closely related species [20]. All the *Rhus* gall aphids belong to the same subfamily Eriosomatinae and are closely related phylogenetically. To avoid false-positive results, we could not infer HTT events within *Rhus* gall aphids, but have drawn observations from the study, that *Rhus* gall genomes are equally targeted to transposition and HT events of transposons. The presence of potentially active TEs in genome could also be involved in HT events within *Rhus* gall aphids yet undetectable due to closed evolutionary relationship.

## 5. Conclusions

Our study reveals that the evolutionary history of Mariner transposons in *Rhus* gall aphids has been subjected to many events of HT, involving total of five other orders of insects at the same time. Moreover, our results show that Mariner elements from *Mauritiana* subfamily is involved in more HT events compared to other MLEs. These results contribute to the description of transposons as genomic symbionts, that mobilize and move between different host lineages, evolving and shaping their host genomes. Overall, our study represents the HT events, involving the *Rhus* gall aphids for the first time, and closed the gap of information about the occurrence of HTs events in galling aphids.

**Patents:** Not applicable.

**Supplementary Materials:** Supplementary file S1: Fasta. file containing all MLEs of *Rhus* gall aphids, Supplementary file S2: CSV spread sheet containing nucleotide sequences of the 20 genes of all insects used for the construction of species tree, Supplementary file S3: CSV file contain the *Mariner* elements extracted from the insects species and used in the construction of tree, and a table showing the genetic distances between them.

**Author Contributions:** Conceptualization, A.A. and Z.R.; methodology, A.A.; software, A.A. and A.H.; validation, Z.R., X.S. and A.H.; formal analysis, Z.R.; investigation, Z.R.; writing—original draft preparation, A.A.; writing—review and editing, Z.R., X.S. and A.A.; supervision, Z.R.; funding acquisition, Z.R. All authors have read and agreed to the published version of the manuscript.”

**Funding:** This study was partially supported by The National Natural Science Foundation of China (31870366), Shanxi International Science and Technology Cooperation Project (201803D421051), Research Project Supported by Shanxi Scholarship Council of China (2020-018), the National High Technology Research and Development “863” Program (2014AA021802).

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All the data used in this study is present in the manuscript and supplementary materials.

**Acknowledgments:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. de Koning, A.J.; Gu, W.; Castoe, T.A.; Batzer, M.A.; Pollock, D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS genetics* 2011, 7, e1002384.
2. Wicker, T.; Gundlach, H.; Spannagl, M.; Uauy, C.; Borrill, P.; Ramírez-González, R.H.; De Oliveira, R.; Mayer, K.F.; Paux, E.; Choulet, F. Impact of transposable elements on genome structure and evolution in bread wheat. *Genome biology* 2018, 19, 1-18.
3. Feschotte, C.; Jiang, N.; Wessler, S.R. Plant transposable elements: where genetics meets genomics. *Nature Reviews Genetics* 2002, 3, 329-341.
4. Bourque, G. Transposable elements in gene regulation and in the evolution of vertebrate genomes. *Current opinion in genetics & development* 2009, 19, 607-612.
5. Warren, I.A.; Naville, M.; Chalopin, D.; Levin, P.; Berger, C.S.; Galiana, D.; Volff, J.-N. Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. *Chromosome Research* 2015, 23, 505-531.
6. Dion-Côté, A.-M.; Renaut, S.; Normandeau, E.; Bernatchez, L. RNA-seq reveals transcriptomic shock involving transposable elements reactivation in hybrids of young lake whitefish species. *Molecular biology and evolution* 2014, 31, 1188-1199.
7. Panaud, O. Horizontal transfers of transposable elements in eukaryotes: the flying genes. *Comptes rendus biologies* 2016, 339, 296-299.
8. Schaack, S.; Gilbert, C.; Feschotte, C. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends in ecology & evolution* 2010, 25, 537-546.
9. Wallau, G.L.; Vieira, C.; Loreto, É.L.S. Genetic exchange in eukaryotes through horizontal transfer: connected by the mobilome. *Mobile DNA* 2018, 9, 1-16.
10. Husnik, F.; McCutcheon, J.P. Functional horizontal gene transfer from bacteria to eukaryotes. *Nature Reviews Microbiology* 2018, 16, 67-79.
11. Graham, L.A.; Li, J.; Davidson, W.S.; Davies, P.L. Smelt was the likely beneficiary of an antifreeze gene laterally transferred between fishes. *BMC evolutionary biology* 2012, 12, 1-13.
12. Gasmi, L.; Boulain, H.; Gauthier, J.; Hua-Van, A.; Musset, K.; Jakubowska, A.K.; Aury, J.-M.; Volkoff, A.-N.; Huguet, E.; Herrero, S. Recurrent domestication by Lepidoptera of genes from their parasites mediated by bracoviruses. *PLoS genetics* 2015, 11, e1005470.
13. Martin, W.F. Too much eukaryote LGT. *BioEssays* 2017, 39, 1700115.
14. Leger, M.M.; Eme, L.; Stairs, C.W.; Roger, A.J. Demystifying eukaryote lateral gene transfer (Response to Martin 2017 DOI: 10.1002/bies. 201700115). *BioEssays* 2018, 40, 1700242.
15. Dunning, L.T.; Olofsson, J.K.; Parisod, C.; Choudhury, R.R.; Moreno-Villena, J.J.; Yang, Y.; Dionora, J.; Quick, W.P.; Park, M.; Bennetzen, J.L. Lateral transfers of large DNA fragments spread functional genes among grasses. *Proceedings of the National Academy of Sciences* 2019, 116, 4416-4425.
16. Daniels, S.B.; Peterson, K.R.; Strausbaugh, L.D.; Kidwell, M.G.; Chovnick, A. Evidence for horizontal transmission of the P transposable element between *Drosophila* species. *Genetics* 1990, 124, 339-355.
17. Wicker, T.; Sabot, F.; Hua-Van, A.; Bennetzen, J.L.; Capy, P.; Chalhoub, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* 2007, 8, 973-982.
18. Lampe, D.J.; Witherspoon, D.J.; Soto-Adames, F.N.; Robertson, H.M. Recent horizontal transfer of mellifera subfamily mariner transposons into insect lineages representing four different orders shows that selection acts only during horizontal transfer. *Molecular biology and evolution* 2003, 20, 554-562.
19. Wallau, G.L.; Capy, P.; Loreto, E.; Le Rouzic, A.; Hua-Van, A. VHICA, a new method to discriminate between vertical and horizontal transposon transfer: Application to the mariner family within *Drosophila*. *Molecular biology and evolution* 2016, 33, 1094-1109.

20. Peccoud, J.; Loiseau, V.; Cordaux, R.; Gilbert, C. Massive horizontal transfer of transposable elements in insects. *Proceedings of the National Academy of Sciences* 2017, 114, 4721-4726.
21. Reiss, D.; Mialdea, G.; Miele, V.; de Vienne, D.M.; Peccoud, J.; Gilbert, C.; Duret, L.; Charlat, S. Global survey of mobile DNA horizontal transfer in arthropods reveals Lepidoptera as a prime hotspot. *PLoS genetics* 2019, 15, e1007965.
22. Silva, J.C.; Loreto, E.L.; Clark, J.B. Factors that affect the horizontal transfer of transposable elements. *Current issues in molecular biology* 2004, 6, 57-72.
23. Houck, M.A.; Clark, J.B.; Peterson, K.R.; Kidwell, M.G. Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis*. *Science* 1991, 253, 1125-1128.
24. Gilbert, C.; Schaack, S.; Pace II, J.K.; Brindley, P.J.; Feschotte, C. A role for host-parasite interactions in the horizontal transfer of transposons across phyla. *Nature* 2010, 464, 1347-1350.
25. Gilbert, C.; Cordaux, R. Viruses as vectors of horizontal transfer of genetic material in eukaryotes. *Current opinion in virology* 2017, 25, 16-22.
26. Loiseau, V.; Peccoud, J.; Bouzar, C.; Guillier, S.; Fan, J.; Gueli Alletti, G.; Meignin, C.; Herniou, E.A.; Federici, B.A.; Wennmann, J.T. Monitoring insect transposable elements in large double-stranded DNA viruses reveals host-to-virus and virus-to-virus transposition. *Molecular biology and evolution* 2021, 38, 3512-3530.
27. Herédia, F.; Loreto, E.L.S.; Valente, V.L.S. Complex evolution of gypsy in *Drosophilid* species. *Molecular biology and evolution* 2004, 21, 1831-1842.
28. Ludwig, A.; Valente, V.d.S.; Loreto, E. Multiple invasions of Errantivirus in the genus *Drosophila*. *Insect molecular biology* 2008, 17, 113-124.
29. de Setta, N.; Van Sluys, M.-A.; Capy, P.; Carareto, C. Multiple invasions of Gypsy and Micropia retroelements in genus *Zapionus* and melanogaster subgroup of the genus *Drosophila*. *BMC evolutionary biology* 2009, 9, 1-18.
30. Sormacheva, I.; Smyshlyaev, G.; Mayorov, V.; Blinov, A.; Novikov, A.; Novikova, O. Vertical evolution and horizontal transfer of CR1 non-LTR retrotransposons and Tc1/mariner DNA transposons in Lepidoptera species. *Molecular biology and evolution* 2012, 29, 3685-3702.
31. Zhang, H.-H.; Li, G.-Y.; Xiong, X.-M.; Han, M.-J.; Zhang, X.-G.; Dai, F.-Y. TRT, a vertebrate and protozoan Tc1-like transposon: current activity and horizontal transfer. *Genome biology and evolution* 2016, 8, 2994-3005.
32. Kharrat, I.; Mezghani, M.; Casse, N.; Denis, F.; Caruso, A.; Makni, H.; Capy, P.; Rouault, J.-D.; Chénais, B.; Makni, M. Characterization of mariner-like transposons of the mauritiana subfamily in seven tree aphid species. *Genetica* 2015, 143, 63-72.
33. Bouallègue, M.; Filée, J.; Kharrat, I.; Mezghani-Khemakhem, M.; Rouault, J.-D.; Makni, M.; Capy, P. Diversity and evolution of mariner-like elements in aphid genomes. *BMC genomics* 2017, 18, 1-12.
34. Ahmad, A.; Wallau, G.L.; Ren, Z. Characterization of Mariner transposons in seven species of *Rhus* gall aphids. *Scientific reports* 2021, 11, 1-11.
35. Chen, X.; Yang, Z.; Chen, H.; Qi, Q.; Liu, J.; Wang, C.; Shao, S.; Lu, Q.; Li, Y.; Wu, H. A Complex nutrient exchange between a gall-forming aphid and its plant host. *Frontiers in Plant Science* 2020, 11, 811.
36. Ren, Z.; Harris, A.; Dikow, R.B.; Ma, E.; Zhong, Y.; Wen, J. Another look at the phylogenetic relationships and intercontinental biogeography of eastern Asian-North American *Rhus* gall aphids (Hemiptera: Aphididae: Eriosomatinae): Evidence from mitogenome sequences via genome skimming. *Molecular phylogenetics and evolution* 2017, 117, 102-110.
37. Ren, Z.; Von Dohlen, C.D.; Harris, A.; Dikow, R.B.; Su, X.; Wen, J. Congruent phylogenetic relationships of Melaphidina aphids (Aphididae: Eriosomatinae: Fordini) according to nuclear and mitochondrial DNA data with taxonomic implications on generic limits. *PloS one* 2019, 14, e0213181.
38. Wallau, G.L.; Capy, P.; Loreto, E.; Hua-Van, A. Genomic landscape and evolutionary dynamics of mariner transposable elements within the *Drosophila* genus. *BMC genomics* 2014, 15, 1-19.
39. Dupeyron, M.; Leclercq, S.; Cerveau, N.; Bouchon, D.; Gilbert, C. Horizontal transfer of transposons between and within crustaceans and insects. *Mobile DNA* 2014, 5, 1-10.
40. Palazzo, A.; Escuder, E.; D'Addabbo, P.; Lovero, D.; Marsano, R.M. A genomic survey of Tc1-mariner transposons in nematodes suggests extensive horizontal transposon transfer events. *Molecular phylogenetics and evolution* 2021, 158, 107090.
41. Jurka, J.; Kapitonov, V.V.; Pavlicek, A.; Klonowski, P.; Kohany, O.; Walichiewicz, J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and genome research* 2005, 110, 462-467.
42. Kohany, O.; Gentles, A.J.; Hankus, L.; Jurka, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC bioinformatics* 2006, 7, 1-7.
43. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic acids research* 2020, 48, D265-D268.
44. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I. CDD: NCBI's conserved domain database. *Nucleic acids research* 2015, 43, D222-D226.
45. Yamada, K.D.; Tomii, K.; Katoh, K. Application of the MAFFT sequence alignment program to large data—reexamination of the usefulness of chained guide trees. *Bioinformatics* 2016, 32, 3246-3251.
46. Posada, D.; Buckley, T.R. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic biology* 2004, 53, 793-808.
47. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nature methods* 2012, 9, 772-772.



- 
48. Gilbert, C.; Feschotte, C. Horizontal acquisition of transposable elements and viral sequences: patterns and consequences. *Current opinion in genetics & development* 2018, 49, 15-24.
  49. Dotto, B.R.; Carvalho, E.L.; Silva, A.F.; Duarte Silva, L.F.; Pinto, P.M.; Ortiz, M.F.; Wallau, G.L. HTT-DB: horizontally transferred transposable elements database. *Bioinformatics* 2015, 31, 2915-2917.
  50. Zhang, H.-H.; Peccoud, J.; Xu, M.-R.-X.; Zhang, X.-G.; Gilbert, C. Horizontal transfer and evolution of transposable elements in vertebrates. *Nature communications* 2020, 11, 1-10.
  51. Strimmer, K.; von Haeseler, A.; Salemi, M. Genetic distances and nucleotide substitution models. *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing* 2009, 111-141.
  52. Dunemann, S.M.; Wasmuth, J.D. Horizontal transfer of a retrotransposon between parasitic nematodes and the common shrew. *Mobile DNA* 2019, 10, 1-13.
  53. Suh, A.; Witt, C.C.; Menger, J.; Sadanandan, K.R.; Podsiadlowski, L.; Gerth, M.; Weigert, A.; McGuire, J.A.; Mudge, J.; Edwards, S.V. Ancient horizontal transfers of retrotransposons between birds and ancestors of human pathogenic nematodes. *Nature communications* 2016, 7, 1-9.
  54. Robertson, H.; MacLeod, E. Five major subfamilies of mariner transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. *Insect molecular biology* 1993, 2, 125-139.
  55. Brunet, F.; Godin, F.; Bazin, C.; Capy, P. Phylogenetic analysis of Mos1-like transposable elements in the Drosophilidae. *Journal of molecular evolution* 1999, 49, 760-768.