Article

Diversity and Evolution of *pogo* and *Tc1/mariner* Transposons in the Apoidea Genome

Yibing Liu†, Wencheng Zong†, Mohamed Diaby, Zheguang Lin, Saisai Wang, Bo Gao, Ting Ji and Chengyi Song*

College of Animal Science & Technology, Yangzhou University, Yangzhou, Jiangsu, 225009, China; yibing-liu@foxmail.com (Y.L.); zongzone@outlook.com (W.Z.); agrical6@yahoo.fr (M.D.); z.lin@yzu.edu.cn (Z.L.); wang850980245@hotmail.com (S.W.); bgao@yzu.edu.cn (B.G.); tji@yzu.edu.cn (T.J.)
*Correspondence: cysong@yzu.edu.cn (C. S.); Tel: 86 514 87979034; Fax: 86 514 87350440
† Yibing Liu and Wencheng Zong contributed equally to this work.

Abstract: Bees (Apoidea), the largest and most crucial radiation of pollinators, play a vital role in ecosystem balance. Transposons are widely distributed in nature and are important drivers of species diversity. However, transposons are rarely reported in important pollinators such as bees. Here, we surveyed 37 genomes in Apoidea, annotated the *pogo* and *Tc1/mariner* transposons in the genome of each species and performed a phylogenetic analysis to determine their overall distribution. The *pogo* and *Tc1/mariner* families showed high diversity and low abundance in the 37 species, and their proportion was significantly higher in solitary bees than in social bees. DD34D/*mariner* was found to be distributed in almost all species and was found in *Apis mellifera*, *Apis mellifera carnica*, *Apis mellifera caucasica* and *Apis mellifera mellifera* and *Euglossa dilemma* may still be active. Using horizontal transfer analysis, we found that DD30D/*Tigger*, DD33D/*Tigger* and DD34D/*mariner* may have experienced horizontal transfer events. The current study displayed the evolution profiles (including diversity, activity, and abundance) of the *pogo* and *Tc1/mariner* transposons across 37 species of Apoidea. Our data revealed their contributions to the genomic variations across these species and facilitated in understanding of the genome evolution of this lineage.

Keywords: Apoidea; *pogo*; *Tc1/mariner*; transposons; evolution

1. Introduction

Bees (Hymenoptera: Apoidea), which originated in the early to mid-cretaceous [1], are the largest and most crucial radiation of pollinators with more than 20,000 described species [2]. Bees play a critical role in ecosystem balance, sustainable agriculture and food security around the world [3,4]. Most bee species lead a solitary life; only ~10% of bee species are eusocial[3]. In the superfamily of Apoidea, seven extant families are recognized: Megachilidae and Apidae of the long-tongued families Andrenidae, Colletidae, Halictidae and Melittidae, and Stenotritidae of the short-tongued families [1,5]. Apidae has the largest family size, with >5900 described species, followed by Halictidae and Megachilidae with >4000 species [2]. By contrast, Stenotritidae is the smallest family and includes just 21 recorded species [2]. Melittidae, with 204 known species, is viewed as the most basal bee family [1,6], although some earlier studies have argued that Colletidae is the sister group to the rest of the bees [7,8]. The emergence of different views relates to advances in molecular and genome research, which have substantially changed and continue to change the understanding of classifications and relationships in bees [3].

Transposons, also called jumping genes, were once considered junk sequences but were later confirmed to play an important role in genome evolution and size [9–12]. Transposons account for a significant sequence component of eukaryote genomes, in insects ranging from as little as 1% in the Antarctic midge [13] to as much as 65% in the migratory locust [14]. Transposons can move in the genome and make copies during this movement, and these processes facilitate the ability to invade the genome of almost all organisms and reshape the structure and phenotype of different lineages [15–17].

Transposons are usually divided into two types according to the structural organization and mechanism: Class I represents RNA transposons and Class II DNA transposons. DNA transposons have been widely reported and can be divided into three main types: cut-and-paste, peel-and-paste and self-synthesizing transposons. Widely reported DNA transposons include *Tc1/mariner*, *pogo*, *hAT*, *PiggyBac*, *CACTA*, *Helitron* and *PIF-Harbinger*. The *Tc1/mariner* superfamily is a member of the cut-and-paste group that was first discovered in *Drosophila mauritiana* (transposon *mariner*) and *Caenorhabditis elegans* (transposon *C. elegans* number 1, *Tc1*) and is the most widely distributed transposon superfamily currently known in nature [18–20]. *Tc1/mariner* transposon usually has a single open reading frame (ORF) of about 340 amino acids (aa), flanked by two terminal inverted repeats (TIRs) and dinucleotide target site duplications (TSDs) of TA.

Based on the phylogeny of the DD,E conserved catalytic motif, Tc1/mariner is classified into DD34E/Tc1 [21–25], DD34D/mariner [26–29], DD35E/TR [30], DD36E/IC [31], DD37D/maT [32,33], DD37E/TRT [34], DD39D/GT [35] and DD41D/VS [36]. The pogo transposon was first found in flies [37],and Tigger [38] was then found in the human genome, Fot, Tan1, Pot1, Pot2, Flipper and Aft1-transposons were found in the genome of fungi [39–44], Lemi1 was found in the genome of plants [45] and pogo-like elements were found in the teleost genome [46]. A previous study confirmed that pogo and Tc1/mariner are two distinct superfamilies [47].

The *mariner* element has been reported in studies of the annotation of honeybees and bumblebees [48,49]. However, comparative studies of transposons in the genome of Apoidea species are lacking. Here, we surveyed the genome of 37 bee species in Apoidea, annotated *pogo* and *Tc1/mariner* transposons in the genome of each species and determined their phylogenetic positions, classification, overall distribution and structural characteristics. We also investigated the evolutionary patterns of DD29-36D/*Tigger* and DD34D/*mariner* transposons. Our data reveal the evolutionary landscape of *pogo* and *Tc1/mariner* transposons in Apoidea and will add to the understanding of their contributions to the evolution of the Apoidea genome.

2. Materials and Methods

2.1. Distribution of pogo and Tc1/mariner within Apoidea

To survey the distribution and diversity of *pogo* and *Tc1/mariner* DNA transposons in 37 sequenced genomes of Apoidea (taxid:34735), we selected several representative transposase sequences from the well-defined families of *pogo* (DD35D/*Passer*, DD35D/*Fot*, DD29-42D/*Lemi*, DD29-59D/*pogoR*, DD36D/*Mover*, DD29-36D/*Tigger*) and *Tc1/mariner* (DD34E/*Tc1*, DD35E/*Traveler*, DD36E/*Incomer*, DD37E/*TRT*, DD38E/*Intruder*, DD34D/*mariner*, DD37D/*maT*, DD39D/*Guest* and DD41D/*Visitor*) superfamilies and subjected them to a Tblastn search against the Apoidea genomes deposited in a whole-genome shotgun database (https://www.ncbi.nlm.nih.gov/) with default parameters. Significant hits were extracted with 2000-base pair (bp) flanking sequences, and the transposon boundaries (TIRs) were then determined manually by alignment using the BioEdit program. The obtained transposon sequence for each species was used to search against its host genome to estimate the copy number using Blast. All Blast hits with >40% coverage and 80% identity to the query were retained and the copy number calculated. In addition, transposons with deficient copy numbers in the genome, which may be false-positive hits resulting from sequence contamination, were verified manually by checking the flanking sequences of the transposon insertion.

2.2. Phylogenetic Analysis and Protein Domain Prediction

For phylogenetic analysis, the conserved DDE domains of the identified transposases were aligned to the representative families of the *pogo* and *Tc1/mariner* transposases separately using MAFFT (v.7.310). The phylogenetic trees were inferred based on the conserved DDE domain using the maximum-likelihood method in the IQ-TREE program. According to the Bayesian information criterion, the best-suited aa substitution model for these data was the LG+G4 model, which was selected by ModelFinder. The reliability of

the maximum-likelihood trees was estimated using the ultrafast bootstrap approach with 1000 replicates. Information for all representative sequences involved in constructing the phylogenetic tree is listed in Table S1.

Protein secondary structure predictions were performed using the PSIPRED program (http://bioinf.cs.ucl.ac.uk/psipred/). Putative nuclear localization signal (NLS) motifs were predicted using PSORT (https://www.genscript.com/psort.html?src=leftbar). The protein domains were identified using the profile hidden Markov models on the hmmscan online web server (https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan).

2.3. Horizontal Transfer and Evolutionary Dynamics Analysis

The coding sequences of three host genes and elongation factor1 alpha (*EF1-a*), heat shock cognate 70 (*Hsc70-4*) and tubulin beta-3 (*tub3*) were used for the comparison with transposon distance to test the horizontal transfer (HT) hypothesis. These three genes have been used as internal controls by other studies that included HT analysis [50], and their accession numbers are listed in Table S2. Species that do not have a complete CDS region of the *EF1-a*, *Hsc70-4* and *tub3* genes in the US National Center for Biotechnology Information (NCBI) database were not included in these calculations. Multiple alignments of *EF1-a*, *Hsc70-4* and *tub3* and all transposons in Apoidea were created using MAFFT, and the pairwise distances were then calculated using MEGA software (v.7.2.06; pairwise deletion, maximum composite likelihood) based on two aligned files. The species divergence times were estimated using the online TimeTree program (http://www.timetree.org/).

To detect the evolutionary dynamics of transposon invasion in each genome, the Kimura two-parameter distance was calculated using the calcDivergenceFromAlign.pl package from RepeatMasker.

3. Results

3.1. Diversity and Distribution of pogo and Tc1/mariner Elements in the Apoidea Genome

To determine the diversity and distribution of *pogo* and *Tc1/mariner* transposons in Apoidea (taxid:34735), we executed a Tblastn search using the NCBI whole-genome shotgun database (https://www.ncbi.nlm.nih.gov/) using 15 representative transposase sequences from different families of the *pogo* (DD35D/*Passer*, DD35D/*Fot*, DD29-42D/*Lemi*, DD29-59D/*pogoR*, DD36D/*Mover*, DD29-36D/*Tigger*) and *Tc1/mariner* (DD34E/*Tc1*, DD35E/*Traver*, DD36E/*Incomer*, DD37E/*TRT*, DD38E/*Intruder*, DD34D/*mariner*, DD37D/*maT*, DD39D/*Guest* and DD41D/*Visitor*) superfamilies. We identified 164 elements in 37 species of Apoidea (Table S3). The phylogenetic tree was then used to define the evolutionary relationships of the *pogo* and *Tc1/mariner* elements in Apoidea that we identified. *TP36_RB* and *Zator* transposases, a clade of the *ITm* group [51], were used as the outgroup, and the incomplete DDE domains were excluded from this analysis. Based on the phylogenetic tree and distribution analysis, two families (DD29-36D/*Tigger* and DD35-36D/*Fot*) classified as members of the *pogo* superfamily and seven families (DD34E/*Tc1*, DD36E/*Incomer*, DD38E/*Intruder*, DD34D/*mariner*, DD37D/*maT*, DD39D and DD41D/*Visitor*) classified as members of the *Tc1/mariner* superfamily were identified in 37 species of Apoidea (Figure 1 and Figure S1).

Differential evolutionary profiles of these families were observed across the Apoidea superfamily in four bee families: Apidae, Megachilidae, Colletidae and Halictidae (Figure 2 and Table S3). The Apidae family displayed high diversity of species (29 species) but represented a lower diversity of *pogo* and *Tc1/mariner* transposons than the other three families (Megachilidae, Colletidae and Halictidae) of Apoidea, in which most species contained seven to eight transposon families. Most species of Apidae contained only two to four transposon families, particularly in Apis genus, and most species exhibited only DD34D/*mariner* and DD41D/*VS* families. The exception was *Habropoda laboriosa*, which had one *pogo* family (*Tigger*), but with variations of DDE domains (DD30D, DD33E and DD36D), and six *Tc1/mariner* families (DD34D/*mariner*, DD37D/*MaT*, DD41D/*VS*, DD34E/*Tc1*, DD36E/*IC* and DD38E/*IT*). This species exhibited the highest diversity of *pogo* and *Tc1/mariner* transposons in Apoidea. Four transposon families (DD29-36D/*Tigger*,

DD34D/*mariner*, DD41D/*VS* and DD34E/*Tc1*) were detected in all bumblebees (five species), *Melipona quadrifasciata* and *Frieseomelitta varia*, whereas the five species of stingless bee (*Tetragonula davenporti*, *Tetragonula hockingsi*, *Tetragonula carbonaria*, *Tetragonula clypearis* and *Tetragonula mellipes*) exhibited only the DD29-36D/*Tigger* and DD34D/*mariner* transposon families. DD34E/*Tc1* and DD36E/*IC* were detected in *Eufriesea mexicana*, DD34E/*Tc1* and DD41D/*VS* in *Euglossa dilemma* and *Ceratina calcarata*, and DD34E/*Tc1* in *Ceratina australensism*, in addition to the DD29-36D/*Tigger* and DD34D/*mariner* families (Figure 1 and Table S3).

Most species of Megachilidae, Colletidae and Halictidae displayed high diversity of *pogo* and *Tc1*/*mariner* transposons, where seven to eight transposon families appeared within each species, except for *Nomia melanderi* and *Megalopta genalis*, where only three and four transposon families of *pogo* and *Tc1*/*mariner* were detected, respectively. Eight families of *pogo* and *Tc1*/*mariner* were detected in *Megachile rotundata* (DD29-36D/*Tigger*, DD35-36D/*Fot*, DD34D/*mariner*, DD34E/*Tc1*, DD36E/*IC*, DD37D/*maT*, DD38E/*IT* and DD41D/*VS*) and *Dufourea novaeangliae* (DD29-36D/*Tigger*, DD35D-36D/*Fot*, DD34D/*mariner*, DD34E/*Tc1*, DD36E/*IC*, DD37D/*MaT*, DD38E/*IT* and DD41D/*VS*), which represented the highest diversity across Apoidea species (Figure 1). Interestingly, almost all solitary bees exhibited a very high transposon abundance and diversity compared with social bees.

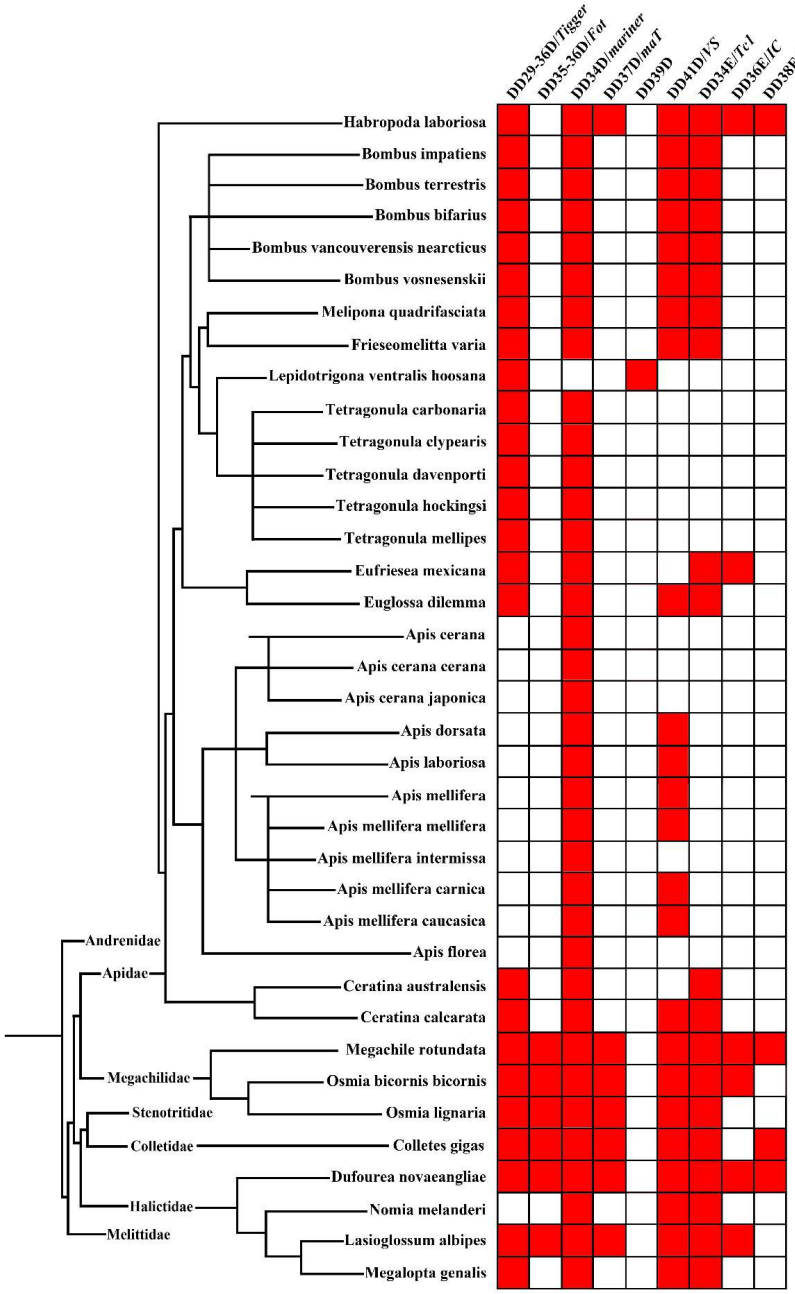

Figure 1. Nine transposon families in various species of Apoidea. Each red square represents a transposon hit.

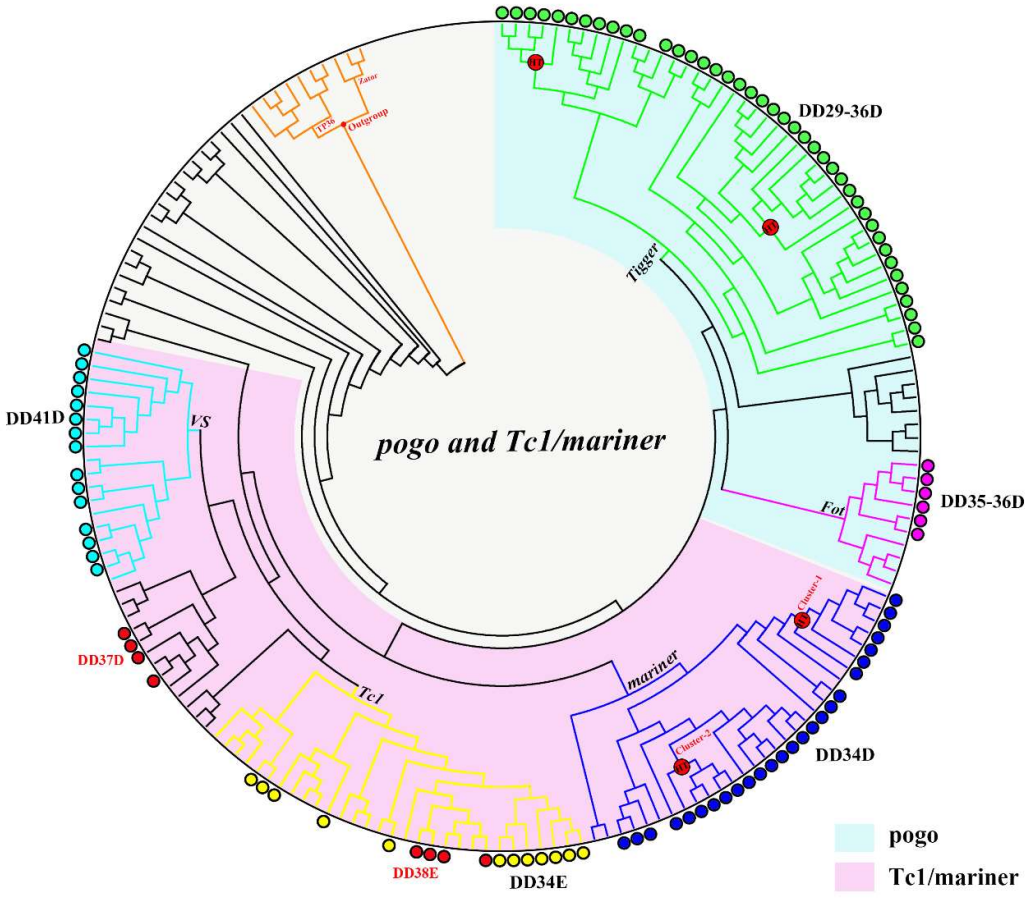

Figure 2. Phylogenetic tree of all transposons. Each colour represents a different family, and branches with horizontal transfer (HT) signals are represented by red circles.

3.2. Activities of pogo and Tc1/mariner in the Apoidea Genome

For further analysis, we investigated the copy numbers of all transposons (Table 1 and Table S3). A copy was defined as Blast result coverage >40% and identity >80%. An intact copy was defined as a Blast result with the full length of the transposon and the complete ORF encoded. Overall, a total of 164 transposon elements were found in 37 bee species. However, we found that most transposons had low copy numbers in the genome, ranging from one to 203 in 37 species, and 114 of 164 transposable elements had <10 copies (>80% identity and >40% coverage) in the host genomes. We found a total of 20 transposons in the bumblebee genus, of which 15 transposons had >15 copies and 14 had complete copies. Thirteen species had complete copies, and 11 species had only one copy (Table 1 and Table S3).

With the exception of DD34D/*mariner*, all other transposon families had relatively low intact copy numbers. For example, neither DD39D/*GT* nor DD36E/*IC* contained an intact copy, whereas in the DD35D-36D/*Fot*, DD37D/*maT* and DD38E/*IT* families, most transposons contained an intact copy. In addition, *Colletes gigas* had two intact copies in the DD35D-36D/*Fot* family, *Colletes gigas* and *Dufourea novaeangliae* had two intact copies in DD37D/*maT*, and *Habropoda laboriosa* had two intact copies in DD38E/*IT*. Most of the intact copies of DD29-36D/*Tigger*, DD41D/*VS* and DD34E/*Tc1* were were ≤5 in hosts. In the DD34D/*mariner* family, we found five species with >10 intact copy numbers: *Apis mellifera*, *Apis mellifera carnica*, *Apis mellifera caucasica*, *Apis mellifera mellifera* and *Euglossa dilemma*. The intact copy number for *Euglossa dilemma* was the largest found in this study (Table 1 and Table S3).

To examine more comprehensively the evolutionary dynamics of transposons in the superfamily, we investigated the DD34D/*mariner* family of transposons, which are distributed in all bees except *Lepidotrigona ventralis hoosana*. Species that cannot be annotated with RepeatMasker were excluded from this analysis. The most transposons of several of the gregarious bee genera, including Bumble, Tetragonula and Apis, seem to have been very young at the time of invasion Some of these species displayed very recent activities with insertion ages <5 million years ago (Figure 3), which suggests that these elements are highly active and may still be functional. Interestingly, almost all solitary bees have a relatively old insertion age, which confirmed our previous prediction.

Table 1. Copy number information for each transposon family.

	DD29D–36D	DD35D–36D	DD34D	DD37D	DD39D	DD41D	DD34E	DD36E	DD38E
	Tigger	*Fot*	*mariner*	*maT*	*GT*	*VS*	*Tc1*	*IC*	*IT*
Copy number	**1–67**	**1–5**	**1–320**	**1–6**	**3**	**1–401**	**1–61**	**3–24**	**3–50**
Intact copy number	**1–4**	**1–2**	**1–28**	**1–2**	**0**	**1–3**	**1–5**	**0**	**1–2**

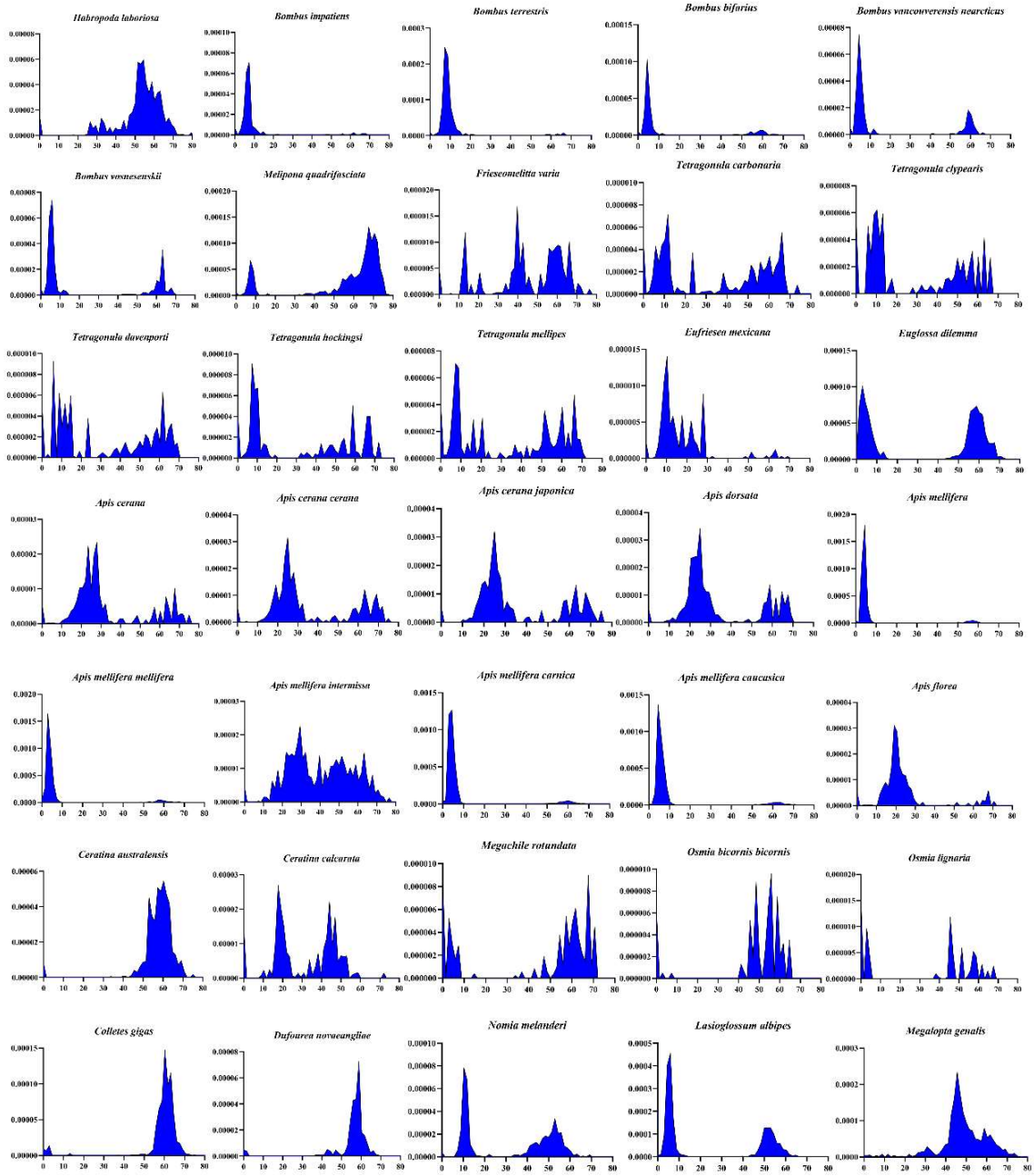

Figure 3. Evolutionary dynamics analysis of DD34D/*mariner* transposons. Species that cannot be executed with RepeatMasker were excluded from this analysis.

3.3. Evolutionary Patterns of pogo and Tc1/mariner in Apoidea

HT is an important form of asexual transmission in nature, and there is increasing evidence that transposons are important participants in HT. In our analysis of four families of Apoidea, namely Apidae, Megachilidae, Colletidae and Halictidae, the average divergence time was 110 million years ago. For example, *Dufourea novaeangliae*, *Nomia melanderi*, *Lasioglossum albipes* and *Megalopta genalis* belong to the Halictidae family, but they had obvious irregular distributions of transposons, indicating putative HT events of these transposons across species within this lineage. By contrast, two very close species, *Ceratina*

calcarata and *Ceratina australensis*, also had different distributions of transposons. Moreover, some relatively distantly related species on the phylogenetic tree were closer than closely related species, for example, the DD30D/*Tigger* family in *Dufourea novaeangliae* and *Megalopta genalis*, which belong to Halictidae, but *Dufourea novaeangliae* was closer to the *Osmia lignaria* in the Megachilidae family on the clade than to *Megalopta genalis*, which is a more closely related species of *Nomia melanderi*. Considering these findings, we speculate that some transposons might have been exposed to several episodes of HT.

To illustrate further the evolutionary patterns of *pogo* and *Tc1*/*mariner* in the bee genome, the pairwise distances between the *EF1-a*, *Hsc70-4* and *tub3*, and four widely distributed transposon families including DD29D-36D/*Tigger*, DD34D/*mariner*, DD34E/*Tc1* and DD41D/*VS* were calculated and compared. All incomplete ORF sequences were excluded from this analysis. *EF1*-a, *Hsc70-4* and *tub3* are conserved host genes and have been used to examine HT in insects [50]. We detected the presence of HT signs in the four clusters (Figure 4). The average distance was smaller for DD30D/*Tigger* (0.081 ± 0.05) than for *EF1-a* (0.145 ± 0.01), *Hsc70-4* (0.164 ± 0.01) and *tub3* (0.205 ± 0.05). The average distance was smaller for DD33D/*Tigger* (0.012 ± 0.03) than for *EF1-a* (0.026 ± 0.02), *Hsc70-4* (0.021 ± 0.02) and *tub3* (0.018 ± 0.02). The average distance was smaller for the DD34D/*VS*-cluster 1 (0.035 ± 0.02) than for *Hsc70-4* (0.059 ± 0.05) and *tub3* (0.095 ± 0.09), and close to *EF1-a* (0.032 ± 0.03). The average distance was also smaller for the DD34D/*VS*-cluster 2 (0.073 ± 0.08) than for *Hsc70-4* (0.106 ± 0.13), and close to *EF1-a* (0.072 ± 0.08) and *tub3* (0.072 ± 0.09) (Table S4) (Table S4).

The four transposon ORFs of DD30D/*Tigger* showed very high overall average sequence identity (93.25% ± 6.49%) across these species (Figure 4). However, their average divergence time was >100 million years ago, which strongly suggests exposure of *Tigger* transposons to HT. Interestingly, the remaining three HTs may reflect the same pattern; that is, these three branches exhibited a common feature that included almost all species sequenced for that genus. For example, DD33D/*Tigger* included the *Tetragonula* genus, DD34D/*mariner* cluster-1 included the *Bombus* genus and DD34D/*mariner* cluster-2 included all western honeybees. Considering the genetic relationship of each species within these genera, this finding suggests that transposons invaded the common ancestor of these genera before species differentiation and then spread vertically to each subspecies. In addition, the divergence time of *Bombus*, *Lasioglossum* and *Osmia* is >110 million years ago, and the identity of DD34D/*mariner* transposons in the Bumble bees with *Lasioglossum albipes* and *Osmia lignaria* is close to 90%. Interestingly, the same *mariner* element was not found in species closely related to *Lasioglossum albipes* or *Osmia lignaria*. This irregular distribution suggests that the five bumblebees and the other two bees experienced independent HT events.

In addition, we did not find HT signals in most other species, which suggests that most of these elements that differ between bee species were not obtained by HT. Overall, these data suggest that the *pogo* and *Tc1*/*mariner* transposons in bees were obtained by both HT and vertical transfer.

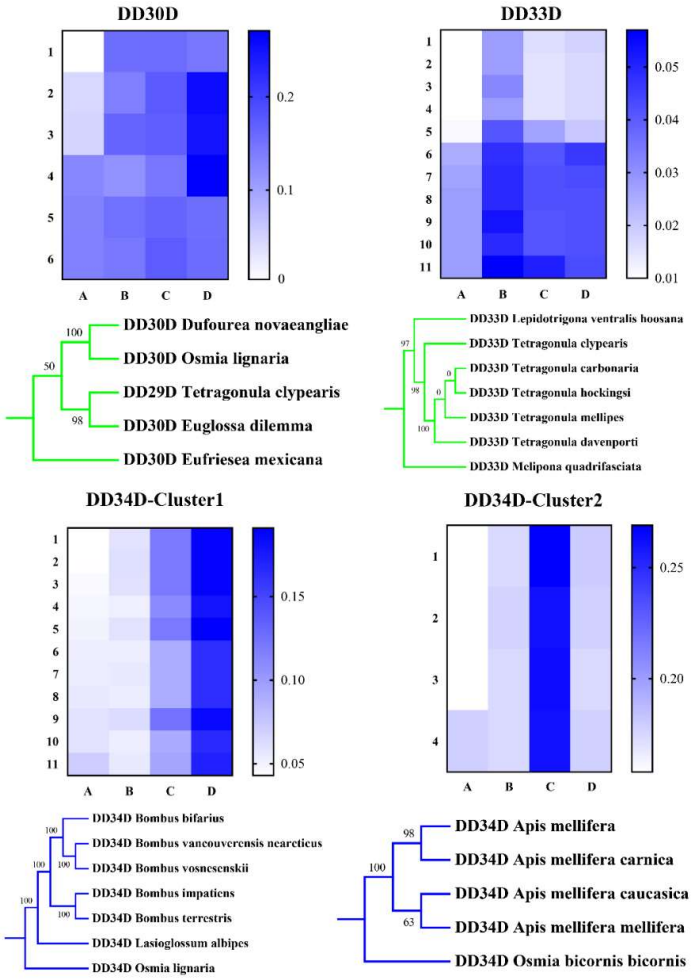

Figure 4. HT analysis in Apoidea. A lighter colour represents a smaller genetic distance, and a darker colour represents a greater genetic distance.

3.4. Structural Organization of the detected pogo and Tc1/mariner Transposons

Previous studies have shown that the *pogo* and *Tc1/mariner* protein structure is highly conserved along with the transposons, as summarized in Figure 5. The transposase of *Tigger* comprises a CENP-B protein and helix-turn-helix (HTH) domain at the N-terminus and a DD29D-36D/*Tigger* catalytic domain at the C-terminus. The N-terminal of *Fot-like* transposase comprises HTH and DNA-binding domain regions, and the C-terminal is a DD35D-36D catalytic domain. The *Tc1/mariner* transposase comprises an HTH domain and a catalytic domain related to transposon transposition.

The DD29-36D/*Tigger* transposons include almost all species except Apis genus, among which 26 species contain 56 transposons, which shows a rich abundance of transposons. The length of all transposons ranged from 907 bp to 4213 bp, and the length of the TIR was 10–33 bp. The overall copy number content was very low, but most transposons included a complete copy. Generally speaking, the last aa of the DD/D catalytic domain of the *pogo* transposon was 'D', but we found an exception in this study. DD33D and DD33E are similar in structure and have a very close phylogenetic relationship. Considering their phylogenetic location and all species that appear in stingless bees, DD33E may have been caused by the DD33D mutation at some stage. DD34D/*mariner* was found to be

the most widely distributed transposon in this study. Almost all bees except *Lepidotrigona ventralis hoosana* and that it appears to represent a very young transposon family in Apoidea.

We found that all transposons were around 1300 bp in length and that 24 transposons contained complete transposase. However, among the six *Apis* species studied (*Apis cerana*, *Apis cerana cerana*, *Apis cerana japonica*, *Apis dorsata*, *Apis florea* and *Apis mellifera intermissa*), we could identify only complete ORFs but could not determine the boundary. DD34E/*Tc1* is considered to be the origin of many *Tc1/mariner* transposons and was found to be a very widely distributed transposon in this analysis. All transposons were 952–1744 bp in length, and the TIRs were 24–221 bp in length. We also found five *Fot*-like transposons, namely DD35D-36/*Fot* in *Colletes gigas*, *Lasioglossum albipes*, *Megachile rotundata*, *Osmia bicornis bicornis*, *Osmia lignaria* and *Dufourea novaeangliae*. Six transposons were found in DD36E/*IC*, but they all lost the function of encoding complete transposase. All DD37D/*maT* started with CAGGG, and DD38E/*IT* started with CACTA and CAGTG. The first five base changes seem to be typical characteristics of some subfamilies. *Lepidotrigona ventralis hoosana* had an incomplete DD39D/*GT* transposase with a total length of 2381 bp, which retains the typical characteristics of DD39D/*GT* transposon beginning with CTCCC. In addition, the recently discovered DD41D/*VS* family was also found in this study. Eleven of the 23 transposons had complete ORFs and were 1024–2258 bp in length. All of the bumblebee catalytic domains had changed from DD41D to DD40D.

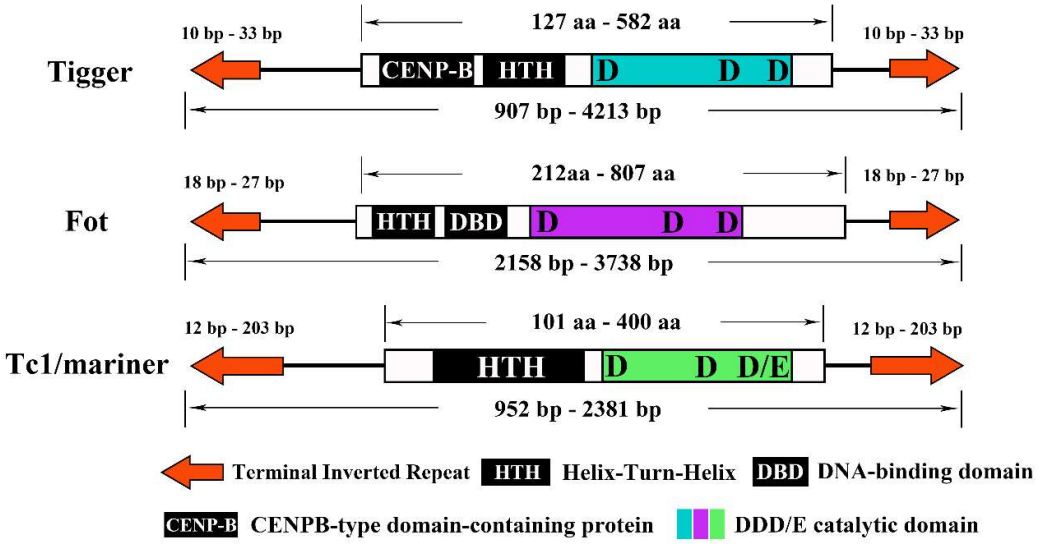

Figure 5. The structure of the two subfamilies of the *pogo* and *Tc1/mariner* transposons. 1) terminal inverted repeats (TIRs) are represented by orange arrows; 2) different protein elements in the DNA-binding domain are represented by black rectangles, each with its name; 3) catalytic domains of the three types of transposons are shown in blue, purple and green. Amino acid length range, full-length nucleotide length range and TIR length range are indicated.

4. Discussion

4.1. Distribution, Diversity and Copy Number in Apoidea

Transposons can generate new genes through genome rearrangement and overexpression, which have played a role in the behavioural diversification of *Apis dorsata* and other honeybees [52]. Given their transposable properties, TEs can accumulate a large number of copies in the host genome, and evidence suggests that transposons are the main cause of differences in genome size between some species [11,12]. The genome of the North African honeybee *Apis mellifera intermissa* contains 5.16% TEs with low diversity, most of which comprise simple repeat sequences [53]. In the European honeybee *Apis mellifera* genome, the percentage of TEs is close to 8% [54]. DNA transposons are the major

repeat sequence, and *mariner* transposons are the most common element within this class [55]. A recent study reported that TEs in the genome of the Asian honeybee *Apis cerana* accounted for 9.15% and simple repeats accounted for most TEs [56]. The DNA transposons constitute 0.11% (247 kb) of the *Apis cerana* genome and 0.57% (1.34 Mb) of the *Apis mellifera* genome. The TEs are significantly smaller in *Apis cerana* than in *Apis mellifera* [57]. In a study of bumblebees, TEs accounted for 14.8% and 17.9% in *Bombus terrestris* and *Bombus impatiens*, respectively, and most of them were retrotransposons, of which *Gypsy* was the most common element. The main types of DNA transposons are TIRs of *mariner*, which account for 1.6% and 2.7%, and their abundance is higher than that of *Apis cerana* and *Apis mellifera* [58]. In addition, compared with other Hymenoptera, the abundance of TEs is low in honeybees, and the lack of TEs is one of the main genomic characteristics of these species [59–62]. Overall, these data suggest that TEs occupy <10% of genomic sequences in honeybees but >10% in bumblebees.

As expected, in our analysis, the transposon diversity was also higher for bumblebees than honeybees. We detected at least seven families of *Tc1/mariner* and two families of *pogo*, which suggests that the genome of some bees represents high diversity of *Tc1/mariner* and *pogo* transposons. In our study, the genomes of 37 bee genomes were rich in *pogo* and *Tc1/mariner* transposons, and included nine different transposon families: DD29D-36D/*Tigger*, DD35D-DD36D/*Tigger*, DD34D/*mariner*, DD34E/*Tc1*, DD36E/*IC*, DD37D/*maT*, DD38E/*IT*, DD39D/*GT* and DD41D/*VS*. Among all transposon families discovered to date, DD34D/*mariner* is the most widely distributed and appears in almost all bee genomes. In addition, DD34D/*mariner* is also widely distributed in nature and has been reported in all species [26,63–67]. The distribution of transposons differs between species, but the distribution of transposons is similar in sister species, which suggests that transposons invade the host's genome before species divergence.

It is generally believed that the copy number is related to transposon activity and that a genome with more intact transposon copies, which include the full length transposase (no frameshift mutation) and the flanking TIRs with very high identity, usually reflects a recent invasion activity, suggesting that some copies may be still active [68]. We examined the copy numbers for all transposons included in this study and found that most families of transposons had only one or few copies, which suggests that most *Tc1/mariner* and *pogo* transposons do not display significant amplification in these lineages. This is consistent with previous findings that most *Tc1/mariner* is detected with low genomic coverage [53,56]. In addition, we found >10 intact copies of DD34D/*mariner* in the genomes of *Apis mellifera*, *Apis mellifera carnica*, *Apis mellifera caucasica*, *Apis mellifera mellifera* and *Euglossa dilemma*. The evolutionary dynamics analysis showed that the DD34D/*mariner* elements of these five genomes are very young, and we speculate that these elements may remain active in the genomes of some species. European honeybee is an important economic animal introduced into China in the last century, where there is a considerable amount of breeding. Active transposon elements may play an important role in the genome evolution and species differentiation of European honeybees, and their existence raises the possibility of molecular operations on honeybees. In addition, the number of newly discovered families of the *pogo* and *Tc1/mariner* superfamily continues to increase as more genomic sequencing data are obtained, and this trend indicates that the diversity of the *pogo* and *Tc1/mariner* superfamilies may be much larger than the currently known families.

4.2. Differential Susceptibility of Apoidea Species to DNA Transposons

In this analysis, we detected a total of nine subfamilies of *Tc1/mariner* transposons in 37 species, for a total of 164 transposons. Interestingly, we found that the genomes of some solitary bees were rich in DNA transposons. For example, *Habropoda laboriosa* exhibited DD29-36D/*Tigger*, DD34D/*mariner*, DD37D/*MaT*, DD41D/*VS*, DD34E/*Tc1*, DD36E/*IC* and DD38E/*IT* transposons; *Dufourea novaeangliae* exhibited DD29-36D/*Tigger*, DD34D/*mariner*, DD35-36D/*Fot*, DD34E/*Tc1*, DD36E/*IC*, DD37D/*MaT*, DD38E/*IT* and DD41D/*VS* transposons; and *Colletes gigas* exhibited DD29-36D/*Tigger*, DD35-36D/*Fot*, DD34D/*mariner*,

DD34E/*Tc1*, DD36E/*IC*, DD37D/*maT*, DD38E/*IT* and DD41D/*VS* transposons. Social bees exhibited significantly fewer transposons than solitary bees. For example, the *Tetragonula* genus exhibited only two transposon subfamilies, DD29-36D/*Tigger* and DD34D/*mariner*, whereas the *Bombus* genus exhibited four families, DD29-36D/*Tigger*, DD34D/*mariner*, DD41D/*VS* and DD34E/*Tc1*. Interestingly, Bombus bees are group that once transformed from solitary bees to social bees. In *Apis*, only two family transposons (DD34D/*mariner* and DD41D/*VS*) appear in the genome of honeybees. This seems to be related to the living environment. Solitary bees act alone and their long-term migration and living environment are more complicated and unstable, whereas social bees live in a group involving social activities and their living environment is more stable [69].

We also found another interesting phenomenon. *Apis mellifera*, *Apis mellifera carnica*, *Apis mellifera caucasica* and *Apis mellifera mellifera* exhibited a complete DD34D/*mariner* element (full-length and complete ORF). However, only the complete ORF of DD34D/*mariner* was exhibited by *Apis cerana*, *Apis cerana cerana*, *Apis cerana japonica*, *Apis dorsata*, *Apis florea* and *Apis mellifera intermissa*, and the boundary could not be determined. Compared with its close relative the European honeybee, the Asian honeybee seems to have a stronger selection of transposons in the genome. It is possible that the drones that develop from unfertilized eggs and carry haploid chromosomes experience strong selective pressure on some active elements [70]. In our further analysis of DD41D/*VS*, we only found that *Apis mellifera*, *Apis mellifera carnica*, *Apis mellifera caucasica*, *Apis mellifera mellifera* and *Apis dorsata* have the complete full-length transposon sequence, but no complete ORF was found. Moreover, *Apis cerana*, *Apis cerana cerana* and *Apis cerana japonica* did not exhibit any DD41D/*VS* transposon elements. It seems that DD41D/*VS* may be an ancient invasion and that, during long-term genome evolution, European honeybees gradually lost its complete copy and Asian honeybees completely lost the entire transposon. Finally, some studies have noted that a host contains an endogenous enzyme system that can silence the transposon or some virus expression, such as RNA interference [71]. However, the mechanism responsible for the interaction between the disappearance of transposon copies and the host in the honeybee genome requires further research.

4.3. HT Events in Apoidea

Transposons are parasitic DNAs whose only role is to replicate and propagate. When a transposon invades a host, the germline genome must be colonized to ensure that it remains in the population. This then increases the copy number [72], and the transposon remains in the genome until, through vertical inhibition, all copies of the transposon become inactive and remain as fossils only [73]. Through genetic drift, these inactive elements can even vanish [73]. To evade this cycle, a transposon must invade a new species or spread to multiple species. That is, to guarantee its longevity, the transposon must transfer to a new genome through HT to start its life cycle again. Although the molecular mechanism responsible for HT is uncertain, studies have provided evidence that HT is the main reason why transposons are widely distributed in nature. Instances include the exchange across marine crustaceans [74], among insects of various orders [75,76] and even between species of different phylae as diverse as the human and parasitic nematode [77].

In this study, although most transposons did not show obvious HT signs, we detected HT traces in DD30D/*Tigger*, DD33D/*Tigger* and DD34D/*mariner*. DD34D/*mariner* is reported to have experienced frequent HT [78]. To find the HT origin of the DD30D/*Tigger* transposon, we performed Blast searching of this transposase in *Dufourea novaeangliae* in the NCBI nucleotide collection (nr/nt) database. The complete DD,D spacing (DD30D) catalytic domains were detected in *Nosema bombycis*, *Strongyloides stercoralis* and *Ctenocephalides felis*. Interestingly, these three creatures have parasitic ability: *Nosema bombycis* is a parasite of silkworm, *Strongyloides stercoralis* is parasitic in humans, and the host of *Ctenocephalides felis* is the cat. These species may serve as potential vectors for *Tigger* transposons that may have invaded the bee genome at a certain time; if so, this strongly suggests the existence of a host–parasitic relationship in bees. Moreover, we have found some

DD34D/*mariner* fragments in ants, which have a high identity with bees. It has been reported that the *mariner* transposon reflects an HT event from ant to mammal [79]. *Mariner* transposons also have been reported in flatworms and nematodes with parasitic ability. These findings suggest that the host–parasitic mode may also be the primary mode of HT of DD34D/*mariner* transposon.

5. Conclusion

Our research provides, for the first time, information about the distribution of *pogo* and *Tc1*/*mariner* family transposons in 37 sequenced honeybee genomes. In general, *pogo* and *Tc1*/*mariner* show high diversity in most species, except for a few honeybees, but have a low abundance. In addition, only DD34D/*mariner* has a high copy number in several species, appears as a complete structure and may have potential activity. Finally, our results also show that DD30D/*Tigger*, DD33D/*Tigger* and DD34D/*mariner* experienced an HT event, and we speculate that they may have invaded their common ancestor before some species formed.

Author Contributions: Conceptualization, Yibing Liu, Wencheng Zong, Zheguang Lin, Bo Gao, Ting Ji, Chengyi Song; methodology, Yibing Liu, Wencheng Zong, Saisai Wang, Chengyi Song; writing—original draft preparation, Yibing Liu, Wencheng Zong, Mohamed Diaby, Zheguang Lin, Chengyi Song; writing—review and editing, Yibing Liu, Wencheng Zong, Bo Gao, Chengyi Song. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by grants from the National Genetically Modified Organisms Breeding Major Projects program (2018ZX08010-08B), the National Natural Science Foundation of China (31671313), the Priority Academic Program Development of Jiangsu Higher Education Institutions, and the High-end Talent Support Program of Yangzhou University to Chengyi Song.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: Thanks to Dr. Yang and Cai for help in bioinformatics analysis.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Danforth, B.N.; Sipes, S.; Fang, J.; Brady, S.G. The history of early bee diversification based on five genes plus morphology. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 15118–15123, doi:10.1073/pnas.0604033103.
2. Ascher, J.S.; Pickering, J. Discover Life bee species guide and world checklist (Hymenoptera:Apoidea:Anthophila). https://bugguide.net/node/view/468583.
3. Danforth, B.N.; Cardinal, S.; Praz, C.; Almeida, E.A.B.; Michez, D. The impact of molecular data on our understanding of bee phylogeny and evolution. *Annu. Rev. Entomol.* **2013**, *58*, 57–78.
4. Potts, S.G.; Imperatriz-Fonseca, V.; Ngo, H.T.; Aizen, M.A.; Biesmeijer, J.C.; Breeze, T.D.; Dicks, L. V.; Garibaldi, L.A.; Hill, R.; Settele, J.; et al. Safeguarding pollinators and their values to human well-being. *Nature* **2016**, *540*, 220–229.
5. Hepburn, H.R. The Bees of the World. *African Zool.* **2001**, doi:10.1080/15627020.2001.11657126.
6. Danforth, B.N.; Fang, J.; Sipes, S. Analysis of family-level relationships in bees (Hymenoptera: Apiformes) using 28S and two previously unexplored nuclear genes: CAD and RNA polymerase II. *Mol. Phylogenet. Evo*l. **2006**, *39*, 358–372, doi:10.1016/j.ympev.2005.09.022.
7. Roberts, R.B. Genesis of the Hymenoptera and the Phases of Their Evolution. *Bull. Entomol. Soc. Am.* **1969**, *15*, doi:10.1093/besa/15.3.258a.
8. Engel, M.S. A monograph of the baltic amber bees and evolution of the apoidea (Hymenoptera). *Bull. Am. Museum Nat. Hist.* **2001**, *259*, 1–192, doi:10.1206/0003-0090(2001)259<0001:amotba>2.0.co;2.
9. Orgel, L.E.; Crick, F.H.C. Selfish DNA: The ultimate parasite. *Nature* **1980**, *284*, 604–607, doi:10.1038/284604a0.
10. Oliver, K.R.; Greene, W.K. Transposable elements and viruses as factors in adaptation and evolution: An expansion and strengthening of the TE-Thrust hypothesis. *Ecol. Evol.* **2012**, *2*, 2912–2933, doi:10.1002/ece3.400.
11. Tenaillon, M.I.; Hollister, J.D.; Gaut, B.S. A triptych of the evolution of plant transposable elements. *Trends Plant Sci.* **2010**, *15*, 471–478, doi:10.1016/j.tplants.2010.05.003.

12. Gao, B.; Shen, D.; Xue, S.; Chen, C.; Cui, H.; Song, C. The contribution of transposable elements to size variations between four teleost genomes. *Mob. Dna* **2016**, *7*, 4.
13. Kelley, J.L.; Peyton, J.T.; Fiston-Lavier, A.S.; Teets, N.M.; Yee, M.C.; Johnston, J.S.; Bustamante, C.D.; Lee, R.E.; Denlinger, D.L. Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. *Nat. Commun.* **2014**, *5*, 4611, doi:10.1038/ncomms5611.
14. Wang, X.; Fang, X.; Yang, P.; Jiang, X.; Jiang, F.; Zhao, D.; Li, B.; Cui, F.; Wei, J.; Ma, C.; et al. The locust genome provides insight into swarm formation and long-distance flight. *Nat. Commun.* **2014**, *5*, 2957, doi:10.1038/ncomms3957.
15. Chénais, B.; Caruso, A.; Hiard, S.; Casse, N. The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments. *Gene* **2012**, *509*, 7–15, doi:10.1016/j.gene.2012.07.042.
16. Rebollo, R.; Zhang, Y.; Mager, D.L. Transposable elements: not as quiet as a mouse. *Genome Biol.* **2012**, *13*.
17. Feschotte, C.; Pritham, E.J. DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.* **2007**, *41*, 331.
18. Wicker, T.; Sabot, F.; Hua-Van, A.; Bennetzen, J.L.; Capy, P.; Chalhoub, B.; Flavell, A.; Leroy, P.; Morgante, M.; Panaud, O.; et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **2007**, *8*, 973–982.
19. Jacobson, J.W.; Medhora, M.M.; Hartl, D.L. Molecular structure of a somatically unstable transposable element in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 8684–8688, doi:10.1073/pnas.83.22.8684.
20. Emmons, S.W.; Yesner, L.; Ruan, K. san; Katzenberg, D. Evidence for a transposon in caenorhabditis elegans. *Cell* **1983**, *32*, 55–65, doi:10.1016/0092-8674(83)90496-8.
21. Vos, J.C.; van Luenen, H.G.; Plasterk, R.H. Characterization of the Caenorhabditis elegans *Tc1* transposase in vivo and in vitro. *Genes Dev* **1993**, *7*, 1244–1253.
22. Radice, A.D.; Bugaj, B.; Fitch, D.H.A.; Emmons, S.W. Widespread occurrence of the *Tc1* transposon family: *Tc1*-like transposons from teleost fish. *Mol. Genet. Genom.* **1994**, *244*, 606–612.
23. Lam, W.L.; Seo, P.; Robison, K.; Virk, S.; Gilbert, W. Discovery of amphibian *Tc1*-like transposon families. *J. Mol. Biol.* **1996**, *257*, 359–366, doi:10.1006/jmbi.1996.0168.
24. Ivics, Z.; Hackett, P.B.; Plasterk, R.H.; Izsvák, Z. Molecular Reconstruction of Sleeping Beauty, a *Tc1*-like Transposon from Fish, and Its Transposition in Human Cells. *Cell* **1997**, *91*, 501–510.
25. Sinzelle, L.; Pollet, N.; Bigot, Y.; Mazabraud, A. Characterization of multiple lineages of *Tc1*-like elements within the genome of the amphibian *Xenopus tropicalis*. *Gene* **2005**, *349*, 187–196, doi:10.1016/j.gene.2004.12.023.
26. Robertson, H.M. The *mariner* transposable element is widespread in insects. *Nature* **1993**, *362*, 241–245, doi:10.1038/362241a0.
27. Plasterk, R.H.A.; Izsvák, Z.; Ivics, Z. Resident aliens the *Tc1/mariner* superfamily of transposable elements. *Trends Genet.* **1999**, *15*, 326–332, doi:10.1016/S0168-9525(99)01777-1.
28. Arkhipova, I.R.; Meselson, M. Diverse DNA transposons in rotifers of the class Bdelloidea. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 11781–11786, doi:10.1073/pnas.0505333102.
29. Nguyen, D.H.; Hermann, D.; Caruso, A.; Tastard, E.; Marchand, J.; Rouault, J.D.; Denis, F.; Thiriet-Rupert, S.; Casse, N.; Morant-Manceau, A. First Evidence of *Mariner*-like Transposons in the Genome of the Marine Microalga *Amphora acutiuscula* (Bacillariophyta). *Protist* **2014**, *165*, 730–744, doi:10.1016/j.protis.2014.08.003.
30. Zong, W.; Gao, B.; Diaby, M.; Shen, D.; Wang, S.; Wang, Y.; Sang, Y.; Chen, C.; Wang, X.; Song, C. Traveler, a new DD35E family of *Tc1/mariner* transposons, invaded vertebrates very recently. *Genome Biol. Evol.* **2020**, *12*, 66–76, doi:10.1093/gbe/evaa034.
31. Sang, Y.; Gao, B.; Diaby, M.; Zong, W.; Chen, C.; Shen, D.; Wang, S.; Wang, Y.; Ivics, Z.; Song, C. Incomer, a DD36E family of *Tc1/mariner* transposons newly discovered in animals. *Mob. DNA* **2019**, *10*, 1–12.
32. Robertson, H.M.; Asplund, M.L. *Bmmar1*: A basal lineage of the *Mariner* family of transposable elements in the silkworm moth, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **1996**, *26*, 945–954, doi:10.1016/S0965-1748(96)00061-6.
33. Gilchrist, A.S.; Shearman, D.C.A.; Frommer, M.; Raphael, K.A.; Deshpande, N.P.; Wilkins, M.R.; Sherwin, W.B.; Sved, J.A. The draft genome of the pest tephritid fruit fly *Bactrocera tryoni*: resources for the genomic analysis of hybridising species. *BMC Genomics* **2014**, *15*, 1153.
34. Zhang, H.H.; Li, G.Y.; Xiong, X.M.; Han, M.J.; Dai, F.Y. *TRT*, a Vertebrate and Protozoan *Tc1*-Like Transposon: Current Activity and Horizontal Transfer. *Genome Biol. Evol.* **2016**, *8*, evw213.
35. Wang, S.; Diaby, M.; Puzakov, M.; Ullah, N.; Wang, Y.; Danley, P.; Chen, C.; Wang, X.; Gao, B.; Song, C. Divergent evolution profiles of DD37D and DD39D families of *Tc1/mariner* transposons in eukaryotes. *Mol. Phylogenet. Evol.* **2021**, *161*, 107143, doi:10.1016/j.ympev.2021.107143.

36. Shen, D.; Gao, B.; Miskey, C.; Chen, C.; Sang, Y.; Zong, W.; Wang, S.; Wang, Y.; Wang, X.; Ivics, Z.; et al. Multiple invasions of visitor, a DD41D family of *tc1/mariner* transposons, throughout the evolution of vertebrates. *Genome Biol. Evol.* **2020**, *12*, 1060–1073, doi:10.1093/GBE/EVAA135.
37. Tudor, M.; Lobocka, M.; Goodell, M.; Pettitt, J.; O'Hare, K. The pogo transposable element family of *Drosophila melanogaster*. *Mol. Gen. Genet. MGG* **1992**, *232*, 126–134.
38. Smit, A.F.A.; Riggs, A.D. *Tiggers* and other DNA transposon fossils in the human genome. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 1443–c1448, doi:10.1073/pnas.93.4.1443.
39. Dufresne, M.; Lespinet, O.; Daboussi, M.J.; Hua-Van, A. Genome-wide comparative analysis of *pogo*-like transposable elements in different *Fusarium* species. *J. Mol. Evol.* **2011**, *73*, 230–243, doi:10.1007/s00239-011-9472-1.
40. Hey, P.; Robson, G.; Birch, M.; Bromley, M. Characterisation of *Aft1* a *Fot1/Pogo* type transposon of *Aspergillus fumigatus*. *Fungal Genet. Biol.* **2008**, *45*, 117 – 126, doi:10.1016/j.fgb.2007.10.009.
41. Nyyssönen, E.; Amutan, M.; Enfield, L.; Stubbs, J.; Dunn-Coleman, N.S. The transposable element *Tan1* of *Aspergillus niger* var. *awamori*, a new member of the *Fot1* family. *Mol. Gen. Genet.* **1996**, *253*, 50 – 56, doi:10.1007/s004380050295.
42. Levis, C.; Fortini, D.; Brygoo, Y. Flipper, a mobile Fot1-like transposable element in *Botrytis cinerea*. *Mol. Gen. Genet.* **1997**, *254*, 674 – 680, doi:10.1007/s004380050465.
43. Kachroo, P.; Leong, S.A.; Chattoo, B.B. Pot2, an inverted repeat transposon from the rice blast fungus *Magnaporthe grisea*. *Mol. Genet. Genom.* **1994**, *245*, 339 – 348, doi:10.1007/BF00290114.
44. Daboussi, M.J.; Langin, T.; Brygoo, Y. Fot1, a new family of fungal transposable elements. *MGG Mol. Gen. Genet.* **1992**, *232*, 12–16, doi:10.1007/BF00299131.
45. Feschotte, C.; Mouchès, C. Evidence that a family of miniature inverted-repeat transposable elements (MITEs) from the *Arabidopsis thaliana* genome has arisen from a *pogo*- like DNA transposon. *Mol. Biol. Evol.* **2000**, *17*, 730 – 737, doi:10.1093/oxfordjournals.molbev.a026351.
46. Gao, B.; Chen, W.; Shen, D.; Wang, S.; Chen, C.; Zhang, L.; Wang, W.; Wang, X.; Song, C. Characterization of autonomous families of Tc1*/mariner* transposons in neoteleost genomes. *Mar. Genomics* **2017**, *34*, 67–77.
47. Gao, B.; Wang, Y.; Diaby, M.; Zong, W.; Shen, D.; Wang, S.; Chen, C.; Wang, X. Evolution of *pogo*, a separate superfamily of *IS630-Tc1-mariner* transposons, revealing recurrent domestication events in vertebrates. **2020**, *11*, 1–15.
48. Sumitani, M.; Lee, J.M.; Hatakeyama, M.; Oishi, K. Cloning and characterization of Acmar1, a *mariner*-like element in the Asiatic honey bee, *Apis cerana japonica* (Hymenoptera, Apocrita). *Arch. Insect Biochem. Physiol.* **2002**, *50*, 183 – 190, doi:10.1002/arch.10043.
49. Rouleux-Bonnin, F.; Petit, A.; Demattei, M.V.; Bigot, Y. Evolution of full-length and deleted forms of the *Mariner*-like element, *Botmar1*, in the genome of the bumble bee, *Bombus terrestris* (hymenoptera: apidae). *J. Mol. Evol.* **2005**, *60*, 736 – 747, doi:10.1007/s00239-004-0195-4.
50. Zhang, H.H.; Xu, H.E.; Shen, Y.H.; Han, M.J.; Zhang, Z. The origin and evolution of six miniature inverted-repeat transposable elements in *Bombyx mori* and *Rhodnius prolixus*. *Genome Biol. Evol.* **2013**, *5*, 2020–2031, doi:10.1093/gbe/evt153.
51. Bao, W.; Jurka, M.G.; Kapitonov, V. V.; Jurka, J. New superfamilies of eukaryotic DNA tyransposons and their internal divisions. *Mol. Biol. Evol.* **2009**, *26*, 983–993, doi:10.1093/molbev/msp013.
52. Oppenheim, S.; Cao, X.; Rueppel, O.; Krongdang, S.; Phokasem, P.; Desalle, R.; Goodwin, S.; Xing, J.; Chantawannakul, P.; Rosenfeld, J.A. Whole Genome Sequencing and Assembly of the Asian Honey Bee *Apis dorsata*. *Genome Biol. Evol.* **2019**, *12*, doi:10.1093/gbe/evz277.
53. Haddad, N.J.; Adjlane, N.; Saini, D.; Menon, A.; Krishnamurthy, V.; Jonklaas, D.; Tomkins, J.P.; Loucif-Ayad, W.; Horth, L. Whole-genome sequencing of north African honey bee *Apis mellifera intermissa* to assess its beneficial traits. *Entomol. Res.* **2018**, *48*, 174 – 186, doi:10.1111/1748-5967.12272.
54. Wallberg, A.; Bunikis, I.; Pettersson, O.V.; Mosbech, M.B.; Childers, A.K.; Evans, J.D.; Mikheyev, A.S.; Robertson, H.M.; Robinson, G.E.; Webster, M.T. A hybrid de novo genome assembly of the honeybee, *Apis mellifera*, with chromosome-length scaffolds. *BMC Genomics* **2019**, *20*, doi:10.1186/s12864-019-5642-0.
55. Elsik, C.G.; Worley, K.C.; Bennett, A.K.; Beye, M.; Camara, F.; Childers, C.P.; de Graaf, D.C.; Debyser, G.; Deng, J.; Devreese, B.; et al. Finding the missing honey bee genes: Lessons learned from a genome upgrade. *BMC Genomics* **2014**, *15*, 86, doi:10.1186/1471-2164-15-86.
56. Wang, Z.L.; Zhu, Y.Q.; Yan, Q.; Yan, W.Y.; Zheng, H.J.; Zeng, Z.J. A Chromosome-Scale Assembly of the Asian Honeybee *Apis cerana* Genome. *Front. Genet.* **2020**, *11*, 279, doi:10.3389/fgene.2020.00279.

57. Park, D.; Jung, W.W.; Choi, B.S.; Jayakodi, M.; Lee, J.; Lim, J.; Yu, Y.; Choi, Y.S.; Lee, M.L.; Park, Y.; et al. Uncovering the novel characteristics of Asian honey bee, *Apis cerana*, by whole genome sequencing. *BMC Genomics* **2015**, *16*, doi:10.1186/1471-2164-16-1.
58. Sadd, B.; Barribeau, S.; Bloch, G.; de Graaf, D.; Dearden, P.; Elsik, C.; Gadau, J.; Grimmelikhuijzen, C.; Hasselmann, M.; Lozier, J.; et al. The genomes of two key bumblebee species with primitive eusocial organization. *Genome Biol* **2015**, *16*, 1 – 32, doi:10.5061/dryad.52hj2.
59. Kaminker, J.S.; Bergman, C.M.; Kronmiller, B.; Carlson, J.; Svirskas, R.; Patel, S.; Frise, E.; Wheeler, D.A.; Lewis, S.E.; Rubin, G.M.; et al. The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.* **2002**, *3*, 1 – 20, doi:10.1186/gb-2002-3-12-research0084.
60. Suen, G.; Teiling, C.; Li, L.; Holt, C.; Abouheif, E.; Bornberg-Bauer, E.; Bouffard, P.; Caldera, E.J.; Cash, E.; Cavanaugh, A.; et al. The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. *PLoS Genet.* **2011**, *7*, e1002007, doi:10.1371/journal.pgen.1002007.
61. Werren, J.H.; Richards, S.; Desjardins, C.A.; Niehuis, O.; Gadau, J.; Colbourne, J.K.; Beukeboom, L.W.; Desplan, C.; Elsik, C.G.; Grimmelikhuijzen, C.J.P.; et al. Functional and evolutionary insights from the genomes of three parasitoid *nasonia* species. *Science (80-.).* **2010**, *327*, 343 – 348, doi:10.1126/science.1178028.
62. Richards, S.; Gibbs, R.A.; Gerardo, N.M.; Moran, N.; Nakabachi, A.; Stern, D.; Tagu, D.; Wilson, A.C.C.; Muzny, D.; Kovar, C.; et al. Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol.* **2010**, *8*, e1000313, doi:10.1371/journal.pbio.1000313.
63. Silva, J.C.; Bastida, F.; Bidwell, S.L.; Johnson, P.J.; Carlton, J.M. A potentially functional *mariner* transposable element in the protist *Trichomonas vaginalis*. *Mol. Biol. Evol.* **2005**, *22*, 126 – 134, doi:10.1093/molbev/msh260.
64. Langin, T.; Capy, P.; Daboussi, M.J. The transposable element *impala*, a fungal member of the *Tc1-mariner* superfamily. *Mgg Mol. Gen. Genet.* **1995**, *246*, 19 – 28.
65. Halaimia-Toumi, N.; Casse, N.; Demattei, M. V.; Renault, S.; Pradier, E.; Bigot, Y.; Laulier, M. The GC-rich transposon *Bytmar1* from the deep-sea hydrothermal crab, *Bythograea thermydron*, may encode three transposase isoforms from a single ORF. *J. Mol. Evol.* **2004**, *59*, 747 – 760, doi:10.1007/s00239-004-2665-0.
66. Auge-Gouillou, C.; Bigot, Y.; Pollet, N.; Hamelin, M.H.; Meunier-Rotival, M.; Periquet, G. Human and other mammalian genomes contain transposons of the *mariner* family. *FEBS Lett.* **1995**, *368*, 541 – 546, doi:10.1016/0014-5793(95)00735-R.
67. Jarvik, T.; Lark, K.G. Characterization of *Soymar1*, a *mariner* element in soybean. *Genetics* **1998**, *149*, 1569–1574.
68. Pujolar, J.M.; Astolfi, L.; Boscari, E.; Vidotto, M.; Barbisan, F.; Bruson, A.; Congiu, L. *Tana*1, a new putatively active *Tc1*-like transposable element in the genome of sturgeons. *Mol. Phylogenet. Evol.* **2013**, *66*, 223–232, doi:10.1016/j.ympev.2012.09.025.
69. Plant; John; Paulus; Hannes, F. Evolution and phylogeny of bees - review and cladistic analysis in light of morphological evidence (Hymenoptera, Apoidea) zoologica vol 161 (2016), 364 pages, 232 fi gures, 49 tables, **2016**.
70. Weinstock, G.M.; Robinson, G.E.; Gibbs, R.A.; Worley, K.C.; Evans, J.D.; Maleszka, R.; Robertson, H.M.; Weaver, D.B.; Beye, M.; Bork, P.; et al. Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* **2006**, *443*, 931 – 949, doi:10.1038/nature05260.
71. Waterhouse, P.M.; Wang, M.B.; Lough, T. Gene silencing as an adaptive defence against viruses. *Nature* **2001**, *411*, 834 – 842.
72. Hartl, D.L.; Lohe, A.R.; Lozovskaya, E.R. Modern thoughts on an ancyent marinere: function, evolution, regulation. *Annu. Rev. Genet.* **1997**, *31*, 337–358.
73. Muñoz-López, M.; Garcia-Pérez, J.L. DNA transposons: nature and applications in genomics. *Curr. Genomics* **2010**, *11*, 115–128.
74. Casse, N.; Bui, Q.T.; Nicolas, V.; Renault, S.; Bigot, Y.; Laulier, M. Species sympatry and horizontal transfers of *Mariner* transposons in marine crustacean genomes. *Mol. Phylogenet. Evol.* **2006**, *40*, 609–619.
75. Robertson, H.M.; Lampe, D.J. Recent horizontal transfer of a *mariner* transposable element among and between Diptera and Neuroptera. *Mol. Biol. Evol.* **1995**, *12*, 850–862.
76. Lampe, D.J. Recent Horizontal Transfer of Mellifera Subfamily *Mariner* Transposons into Insect Lineages Representing Four Different Orders Shows that Selection Acts Only During Horizontal Transfer. *Mol. Biol. Evol.* **2003**, *20*, 554–562, doi:10.1093/molbev/msg069.
77. Laha, T.; Loukas, A.; Wattanasatitarpa, S.; Somprakhon, J.; Kewgrai, N.; Sithithaworn, P.; Kaewkes, S.; Mitreva, M.; Brindley, P.J. The bandit, a new DNA transposon from a hookworm possible horizontal genetic transfer between host and parasite. *PLoS Negl. Trop. Dis.* **2007**, *1*, e35.

78. Robertson, H.M.; Soto-adames, F.N.; Walden, K.K.O.; Avancini, R.M.P.; Lampe, D.J. The *mariner* transposons of animals: horizontally jumping genes. In *Horizontal gene transfer*; 2002 ISBN 0126801266.
79. Oliveira, S.G.; Bao, W.; Martins, C.; Jurka, J. Horizontal transfers of *Mariner* transposons between mammals and insects. *Mob. DNA* **2012**, *3*, 14, doi:10.1186/1759-8753-3-14.