Evidence of horizontal transfer of non-autonomous Lep1 Helitrons facilitated by host-parasite interactions

Xuezhu Guo1, Jingkun Gao1, Fei Li2 & Jianjun Wang1

1College of Horticulture and Plant Protection, Yangzhou University, Yangzhou, 225009, China, 2College of Plant Protection, Nanjing Agricultural University, Nanjing, 210095, China.

Horizontal transfer (HT) of transposable elements has been recognized to be a major force driving genomic variation and biological innovation of eukaryotic organisms. However, the mechanisms of HT in eukaryotes remain poorly appreciated. The non-autonomous Helitron family, Lep1, has been found to be widespread in lepidopteran species, and showed little interspecific sequence similarity of acquired sequences at 3′ end, which makes Lep1 a good candidate for the study of HT. In this study, we describe the Lep1-like elements in multiple non-lepidopteran species, including two aphids, Acyrthosiphon pisum and Aphis gossypii, two parasitoid wasps, Cotesia vestalis, and Copidosoma floridanum, one beetle, Anoplophora glabripennis, as well as two bracoviruses in parasitoid wasps, and one intracellular microsporidia parasite, Nosema bombycis. The patchy distribution and high sequence similarity of Lep1-like elements among distantly related lineages as well as incongruence of Lep1-like elements and host phylogeny suggest the occurrence of HT. Remarkably, the acquired sequences of both NbLep1 from N. bombycis and CfLep1 from C. floridanum showed over 90% identity with their lepidopteran host Lep1. Thus, our study provides evidence of HT facilitated by host-parasite interactions. Furthermore, in the context of these data, we discuss the putative directions and vectors of HT of Lep1 Helitrons.

Transposable elements (TEs) are prevalent in the genomes of almost all eukaryotes and are traditionally categorized based on their mode of transposition as class-I elements or retrotransposons and class-II elements or DNA transposons1. Copy and paste retrotransposons replicate via an RNA intermediate, which is reverse transcribed prior to its reintegration into the genome, whereas DNA transposons move through a single or double-stranded DNA intermediate and were divided into three major subclasses including the classic “cut-and-paste” transposons, rolling-circle (RC) transposons called Helitrons, and Mavericks, whose mechanism of transposition is not yet well characterized, but that likely replicate using a self-encoded DNA polymerase2.

The inherent mobility and replication abilities of TEs make them particularly prone to transfer horizontally between organisms to avoid co-evolved host suppression mechanisms leading to vertical inactivation3,4. Horizontal transfer (HT) can be defined as the exchange of genetic material between species by nonvertical inheritance without the aid of any form of sexual mechanism5. Over 200 solid cases of horizontal transfers of TEs (horizontal transposon transfer or HTT) have been described so far in multicellular eukaryotes6,7 with the majority of HTT cases involving drosophilid flies, and it is believed that TEs rely heavily on HT for their propagation and maintenance throughout evolution6–9. However, despite mounting examples of HTT, the unequivocal confirmation of any specific mechanism acting to shuttle DNA among eukaryotes remains poorly appreciated.

Helitrons, a new superfamily of transposons, have recently been uncovered by the computational analysis of genomic sequences of Arabidopsis thaliana, Oryza sativa and Caenorhabditis elegans10. Unlike traditional class DNA TEs, Helitrons are unique in that they do not produce target site duplications on their integration into the host genome and do not contain terminal repeats, and thus are difficult to be identified11–13. However, Helitrons have conserved sequence features including a “TC” motif on the 5′-end and a “CTRR” motif on the 3′-end, and contain a palindromic sequence of 16–20 bp near the 3′-terminus, which can form a hairpin structure10,14. In addition, Helitrons tend to insert preferentially between host nucleotides adenine and thymidine10,13. The non-autonomous Helitrons, Lep1, were originally identified within intron and untranslated regions from eight lepidopteran species14, and subsequently described as lepidopteran-specific common sequence 3(LSCS3)14. Recent
study showed that Lep1 Helitrons were widespread in more than 30 lepidopteran species, and estimated to occupy 1.3 × 10^8 of the Bombyx mori genome sequence.

Although an increasing number of Lep1 elements are being identified in lepidopteran genomes, little is known about Lep1 in non-lepidopteran insect species. In this study, we report the presence of Lep1-like elements in several non-lepidopteran insect species and other distantly related organisms. Our results suggested that the Lep1 Helitrons can undergo horizontal transfer by diverse means.

Results and Discussion

Evolutionary dynamics of Lep1 in Helicoverpa armigera and its related species. While Lep1 Helitrons have been previously described in multiple lepidopteran insects, the evolutionary dynamics of Lep1 had not been further investigated. In this study, a Lep1-like sequence (named HaLep1_1) was identified in H. armigera by genome walking and subsequent sequence analysis. The HaLep1_1 element is 193 bp in length and located at 756 bp upstream of the translation start codon of the CYP6AE12 gene in the reverse orientation. A total of 21 full length sequences with high homology to HaLep1_1 were identified from non-redundant database, and named HaLep1_2-HaLep1_22 (Table S1). Figure S1 shows the alignment of these sequences. As shown in Figure S1, these sequences present the typical structural features of the Lep1 elements: almost all HaLep1 copies have characteristic 5' -TC and 3' -CTRY nucleotide termini as well as CTRR motif at the 3' end of acquired sequence. The integration occurs precisely between the host A and T nucleotides, without duplications or deletions of the target sites, consistent with the RC mechanism. The phylogeny was constructed based on nucleotide sequences of all these HaLep1 elements. Neighbor-joining (NJ) analysis demonstrated the presence of three clear major lineages (Fig. S2), designated Lineage A (HaLep1A), Lineage B (HaLep1B), and Lineage C (HaLep1C), among which, 6 elements form lineage HaLep1A, while HaLep1B and HaLep1C were represented by 6 and 9 elements, respectively. Notably, HaLep1 elements from Lineage A and Lineage B showed relatively high identity with 134 bp Lep1 consensus sequence (83%-89%), while HaLep1 elements from Lineage C showed only 68% to 78% identity with Lep1 consensus sequence (Table S1). These results suggested that HaLep1 lineages might transfer independently into the genome of H. armigera.

The HaLep1_1 sequence was used as a query to search against nucleotide (nr/nt) and EST (est_others) collections to detect sequences with high identity with HaLep1_1 in lepidopteran species other than H. armigera. The result showed that HaLep1_1 sequence shared the highest similarity with two species of Heliothis including Helicoverpa zeas and Heliothis virescens. For example, three sequences from H. zeas (accession number: EF152213, EF152207 and HQ840515) were identified from nucleotide (nr/nt) database to have over 93% identity with HaLep1_1. A total of 103 matches were detected in H. virescens EST database with an E-value less than 1e-50. Representative examples of these sequences are shown in Figure S3. Remarkably, the acquired sequence at 3' end was only found in H. zeas and H. virescens. Further analysis showed that the acquired sequences at 3' end of all other HaLep1 elements were also conserved only in H. zeas and H. virescens (Table S2), suggesting that the acquired sequence was unique to H. armigera and its closely related species. These results consist with previous finding that the acquired sequence at 3' end of Lep1 elements shared little interspecific sequence similarity, while high similarity was only found within species or closely related species.

To understand whether HaLep1 elements mobilized recently, the insertion polymorphism was assessed experimentally or by homology searches. The results of PCR and subsequent sequencing of DNA products showed that in samples of 12 individuals, the percentage of individuals with the band for HaLep1_1 insertion was 25% (Fig. S4A). Paralogous or orthologous empty sites were also analyzed using homology searches. The results showed that no Lep1-like sequence was found in paralogous sites of HaLep1_20 (accession number: FP340435) in H. armigera as well as in orthologous site of HaLep1_8 in H. zeas (accession number: DQ788839) (Fig. S4B, C). The H. armigera is a pest widespread across the Old World from the Western Pacific to the Canary Islands, while H. zeas is found throughout the warm regions of the New World and in Hawaii, and is recently thought to be derived from a founder population of H. armigera approximately 1.5 million years ago. The intra-species insertion polymorphism of HaLep1_1 suggested a very recent transposition. The insertion polymorphism of HaLep1_8 in two different but closely related species suggested that HaLep1_8 might horizontally transfer into a common ancestor of H. armigera and H. zeas, and the absence of orthologous copy in H. zeas was due to the fact that the element had been actively transposing some time after the split of these two species, or to the differential fixation or loss of ancestrally polymorphic insertions in these two species. Further research is necessary to identify the parent TE of the non-autonomous HaLep1 elements.

Identification of Lep1-like sequences in non-lepidopteran species.

To characterize the distribution of Lep1-like elements in non-lepidopteran insect species, Lep1 consensus sequence was used as query in Blastn searches against insect genome assembly. While no significant hits were detected in the genomes of red flour beetle, Tribolium castaneum (Coleoptera: Tenebrionidae), the blood-sucking bug, Rhodius prololis (Hemiptera: reviduinae), the human body louse, Pediculus humanus (Phthiraptera: Pediculidae), the honey bee, Apis mellifera (Hymenoptera: Apidae), the parasitoid wasp Nasonia vitripennis (Hymenoptera: Pteromalidae), and six ants (Hymenoptera: Formicidae) including Camponotus floridanus, Linepithema humile, Pogonomyrmex barbatus, Atta cephalotes, Harpegnathos saltator, and Solenopsis invicta, our Blastn search detected 138 hits with ≥70% identity to the query over >100 bp in the pea aphid, Acrystosiphon pisum (Hemiptera: Aphididae) genome assembly (AphidBase 2.1) (Table S3). However, because of the presence of many chimaeric elements, the acquired sequence regions as well as the proper boundaries of these Lep1-like sequences could not be precisely defined by multiple sequence alignment. Interestingly, one 662 bp EST sequence from the cotton aphid, Aphis gossypii (accession number: GWS06388) also showed high identity with Lep1 consensus sequence (89%) as well as HaLep1_8 (90%).

Lep1 consensus sequence was further used as query in Blastn searches against all the species with sequences deposited in the GenBank databases. A total of 278 significantly similar sequences to Lep1 (≥70% identity to the query over >100 bp) were identified in the genome shotgun sequence of Anoplophora glabripennis (Coleoptera: Cerambycidae). These sequences were subjected to pairwise alignment to reveal the boundaries and evaluated for the presence of structural features typical of Lep1 Helitrons, of these, a total of 175 full length elements were identified and named AglaLep1_1 to AglaLep1_175 (Table S4). The consensus sequence of the AglaLep1 is 209 bp long, shared 86% similarity with Lep1. It also has characteristic 5'-TC and 3'-CTRY nucleotide termini as well as CTRR motif at the 3' end of 65 bp acquired sequence. Comparative analysis showed that the match between the AglaLep1 elements and their consensus sequence ranged from 95% to 100% (excluding indels), with a median similarity of 98%, suggesting a recent transposition activity.

Blastn searches using the Lep1 consensus sequence as a query also yielded several significant hits in two parasitoid wasps, Cotesia vestalis and Cepidosoma floridanum, as well as one microsporidia parasite, Nosema bombycis (Table 1). For example, two elements from C. vestalis (CvLep1_1 and CvLep1_2) showed 90% and 86% identity with
Lep1, which are 190 bp and 201 bp in length including 62 bp and 65 bp acquired sequence, respectively. In C. floridanum, two full length copies of Lep1-like elements, CjLep1_1 and CjLep1_2, were identified, which are 253 bp and 236 bp in length including 122 bp and 100 bp acquired sequence, and showed 75% and 69% identity with Lep1, respectively. Three full length copies of Lep1-like elements were also found in N. bombycis (NbLep1_1-NbLep1_3), which are 445, 208, 218 bp in length including 314 bp, 76 bp and 84 bp acquired sequence, and showed 93%, 83%, 84% identity with Lep1, respectively.

It is also noteworthy that we identified highly similar sequences in two polydnaviruses (PDVs), which are symbiotically associated with hymenopteran wasps, including three copies from C. vestalis bracovirus (CvBVLep1_1-CvBVLep1_3), four copies from C. vestalis (CsKBVLep1_1-CsKBVLep1_4) and Mombasa (CsMBVLep1_1-CsMBVLep1_4) strains of Cotesia sesamiae bracovirus (Table 1). These elements vary in size from 196 bp (CsMBVLep1_1) to 344 bp (CsKBVLep1_4). Pairwise comparisons of individual elements reveal high sequence identity (82%–94%) with Lep1 consensus sequence (Table 1).

Overall, our BLAST searches detected significantly similar sequences to Lep1 element in other non-lepidopteran species. While cross-species contamination is a concern, our Blastx analysis of the flanking sequences of the representative non-lepidopteran Lep1 elements did not find any evidence of contamination (Table S5). The largest number of sequences with significant similarity to Lep1 was identified in A. pism and A. glabripennis. However, this is probably due to the abundant sequence resources for these two species compared with parasitoid wasps. The low copy number of Lep1-like element identified in N. bombycis and polydnaviruses might be explained by the low likelihood of fixation and rapid removal of nonessential DNA in their genomes.

Evidence of horizontal transfer of non-autonomous Lep1 Helitrons.

Traditionally, horizontal transfer has been implied when highly similar TEs have been found in distantly related taxa accompanied by their discontinuous distribution, and such phenomenon could not be explained in terms of vertical inheritance. In this study, a patchy taxonomic distribution of Lep1 was clearly revealed by database searches. While Lep1-like elements were detected in five non-lepidopteran insect species including two aphids (A. pism and A. gossypii, Hemiptera), one beetle (A. glabripennis, Coleoptera), and two parasitoid wasps (C. vestalis and C. floridanum, Hymenoptera), no significant hits were observed in the genomes of R. prolixus (Hemiptera), T. castaneum (Coleoptera), N. bombycis and A. mellifera, as well as six ants (Hymenoptera). Remarkably, Lep1-like elements were also detected in one intracellular microsporidia parasite, N. bombycis, and two bracoviruses which are symbiotically associated with hymenopteran parasitic wasps. In many cases, the sequence identity of the Lep1 Helitrons is exceptionally high compared with the divergence of the hosts. For example, hymenopteran CvLep1_1 showed 90% identity with lepidopteran Lep1 consensus sequence, which diverged 325 million years ago (http://www.timetree.org/)²⁸, and CsBVLepl_1 and NbLep1_1 showed 94% and 93% identity with Lep1, respectively.

In an effort to investigate the relationships within Lep1 more closely, we reconstructed phylogenetic trees that focuses on these elements and representative lepidopteran Lep1 elements. The results obtained with NJ and ML methods were mostly congruent. We chose to present the topologies obtained by NJ method (Fig. 1). The ML tree is provided in Figure S5. The result indicates the existence of two major clades (Fig. 1). The largest clade comprised Lep1-like sequences from bracoviruses, N. bombycis, C. vestalis, A. glabripennis, A. gossypii, and representative Lep1 elements from B. mori (BmLep1_335 and BmLep1_87), Papilio dardanus (PdLep1_1), and H. armigera (HaLep1A and HaLep1B). Inside this clade, two sub-clades formed by CsKBVLep1_4, NbLep1_1, BmLep1_335, and A. gossypii AgosLep1_1, HaLep1A and HaLep1B, respectively, were strongly supported (100% and 99%), and CvLep1_1, GMBVLep1_4, and CvBVLepl_2 were clustered together, with a bootstrap value of 73%. In the second clade, the Lep1-like sequences from C. floridanum (CfLep1_1 and CfLep1_2) were clustered with Trichoplusia ni (TnLep1_1) (EF372817), with a significant bootstrap value of 99%. These results suggested the occurrence of HT and that multiple mechanisms may underlie the horizontal spread of Lep1.

While the inherent abilities of TEs to replicate and integrate into the host genome undoubtedly facilitate HT between organisms, the

| Name          | Genus species   | GenBank accession | location | Size (bp) | % sim. |
|---------------|-----------------|-------------------|----------|-----------|-------|
| CvLep1_1      | Cotesia vestalis| GAKG01025082      | 774-963  | 190       | 90    |
| CvLep1_2      | Cotesia vestalis| GAKG01005573      | 371-171  | 201       | 86    |
| CvBVLepl_1    | Cotesia vestalis BV | HQ009543      | 29755-29864 | 196  | 83    |
| CvBVLepl_2    | Cotesia vestalis BV | HQ009537      | 7652-7557 | 224  | 83    |
| CvBVLepl_3    | Cotesia vestalis BV | DQ075354      | 7743-7648 |       |       |
| CsKBVLep1_1   | Cotesia sesamiae KBV | EFF710635   | 103544-103433 | 213  | 94    |
| CsKBVLep1_2   | Cotesia sesamiae KBV | EFF710634   | 3978-3860 | 196  | 84    |
| CsKBVLep1_3   | Cotesia sesamiae KBV | EFF710633   | 66996-66890 |      |       |
| CsKBVLep1_4   | Cotesia sesamiae KBV | EFF710633   | 63741-63590 | 218  | 83    |
| CsMBVLep1_1   | Cotesia sesamiae MBV | EFF710633   | 3978-3860 | 344  | 90    |
| CsMBVLep1_2   | Cotesia sesamiae MBV | EFF710633   | 11656-11767 | 196  | 84    |
| CsMBVLep1_3   | Cotesia sesamiae MBV | EFF710641   | 98398-98504 | 218  | 83    |
| CsMBVLep1_4   | Cotesia sesamiae MBV | EFF710642   | 36832-36943 | 218  | 83    |
| NbLep1_1      | Nosema bombycis  | ACJZ01002694    | 334-778  | 445  | 93    |
| NbLep1_2      | Nosema bombycis  | ACJZ01001444    | 842-635  | 208  | 83    |
| NbLep1_3      | Nosema bombycis  | ACJZ01001453    | 267-484  | 218  | 84    |
| CvLep1_1      | Cotesiose floridanum | J1831644   | 375-1056 | 253  | 75    |
| CvLep1_2      | Cotesiose floridanum | J1839208   | 363-128  | 236  | 69    |
| AgosLep1_1    | Aphis gossypii   | GWS06388      | 153-362  | 210  | 86    |
precise mechanisms underlying HTT remain largely mysterious. Several hypotheses have been proposed to explain how TEs might be transferred between eukaryotic hosts. For example, TEs can putatively explore events like parasite mediated transfers from one host to another, as in the case of the mariner element transferred between the braconid parasitoid wasp, *Ascogaster reticulatus*, and its moth host, the smaller tea tortrix, *Adoxophyes honmai*. The little interspecific sequence similarity of acquired sequences at 3' end makes *Lep1* a good candidate for the study of HTT mechanisms. In this study, the identification of *Lep1* Helitrons in *C. floridanum* and *N. bombycis* as well as their lepidopteran host insects is of particular interest. *C. floridanum* is a polyembryonic encyrtid that parasitizes the egg stage of *T. ni* and related moth species. The *N. bombycis* is well known as the causal agent of microsporidian disease pebrine of silkworm larvae, *B. mori*. Sequence comparison showed that, across the entire length of the elements, *CfLep1_1* showed 94% identity with...
TnLep₁, NbLep₁ showed 91% identity with BmLep₁, and NbLep₂ and NbLep₃ showed 98% and 94% identity with BmLep₁, respectively. Specifically, the acquired sequences of both NbLep and CfLep₁ showed over 90% identity with their lepidopteran host Lep₁ elements (Fig. 2). Thus, our study provides evidence of the occurrence of HTT facilitated by host-parasite interactions.

Putative directions of horizontal transfer of Lep₁ Helitrons. The Polydnaviruses display an obligatory relationship with endoparasitoid wasps belonging to the Braconidae family and Ichneumonidae family, and have been proposed to be potential vectors for the delivery of TEs among species. During the past few years, there have been several reports of TE-like sequences in the genomes of Polydnaviruses. In this study, Lep₁-like sequences were identified in C. vestalis (CvBV), and C. sesamiae (CsKBV) and CsMBV) strains. These results suggested that Polydnaviruses might be important vectors of HT of Lep₁ Helitrons. Interestingly, Lep₁-like sequences were also identified in the parasitoid wasp, C. vestalis. Considering the widespread distribution of Lep₁-like sequences in lepidopteran species, it is reasonable to propose that Lep₁ Helitrons were transferred from lepidopteran hosts to parasitoid wasps using polydnaviruses to mediate the actual transfer of TE DNA between cells. However, the acquired sequences of CvLep₁ and CvBvLep₁ showed only moderate similarity (72% between CvLep₁ and CvBvLep₁) (Fig. 3). This is possibly because of the current limited availability of C. vestalis sequence. C. vestalis is larval parasitoid of the diamondback moth, Plutella xylostella (Lepidoptera: Plutellidae). However, we also did not find sequences similar to acquired sequences of CvLep₁ and CvBvLep₁ in the genome database of P. xylostella. Because in some cases, parasitoids are likely to oviposit within marginal (or even completely unsuitable) hosts in the laboratory or field, even if suitable hosts are present, and C. vestalis has been reared from several species belonging to different lepidopteran families, we propose that CvLep₁ identified in this study may be transferred from other lepidopteran host to C. vestalis. This hypothesis could be partly supported by the fact observed in this study: the acquired sequence of CsKBVLep₁ showed 90% similarity with BmLep₁, suggesting that C. sesamiae might have oviposited within B. mori (Fig. S6). Alternatively, considering that the Braconidae wasps form a monophyletic assemblage named the microgastroid complex, which evolved 100 million years ago, and BVs evolved from the interaction between the common ancestor of microgastroids and a single ancestral virus, the lepidopteran Lep₁ might repeatedly invade into the common ancestor of BV, and then horizontally transfer to Cotesia parasitoids. This hypothesis could be supported by the facts observed in this study: the acquired sequence of CvBvLep₁ showed 88% similarity with CsKBVLep₁ and CsMBVLep₁, and CsKBVLep₁ and CsMBVLep₁ showed 90% similarity with BmLep₁ (Fig. 3), suggesting
that Lep1-like element might insert into the common ancestor genome of these viruses. Additional experiments and taxon sampling are necessary to further determine the direction and frequency of HT of Lep1 Helitrons.

Other putative mechanisms underlying horizontal transfer of Lep1 Helitrons. While our results indicate the role of host-parasite interactions in HT of Lep1, the presence of Lep1-like elements in A. glabripennis and A. pism as well as A. gossypii is somewhat intriguing. Notably, a recent study also showed the occurrence of horizontal transfer of short interspersed nuclear elements (HaSE2) between Heliothine species and A. gossypii 40. It has been proposed that mechanisms of HT include insect-associated facultative symbionts such as genera Wolbachia, Rickettsia, Cardinium, Arsenophonus, and Sodalis 41–45. In addition to the possibility of HT through facultative symbionts, the Lep1-like elements identified in N. bombycis in this study suggested that the intracellular microsporidia parasite is also a potential vector for HT. It is reported that Wolbachia infect at least 20% of all insect species including aphids 46–48, and apart from the domesticated silkworms, N. bombycis can also infect various lepidopteran insects 49–51, indicative of their broad hosts range. Additionally, a previous study showed that A. glabripennis could be infected by microsporidia parasite, Nosema glabripennis 52. Thus, we proposed that facultative symbionts including Wolbachia and obligate intracellular microsporidia parasites might play a role in the HT of Lep1-like elements in A. glabripennis and A. pism as well as A. gossypii. More widespread sequencing would be required to find exact vectors that would facilitate the HT of Lep1 Helitrons in these species.

Methods
DNA extraction and genome walking. A previous study has shown that TEs were enriched within or in close proximity to xenobiotic-metabolizing cytochrome P450 genes in H. zea 53. To isolate TEs in H. armigera, we performed genome walking to obtain the 5'-flanking sequence of an insecticide resistance-associated cytochrome P450 gene, CYP6AE12, in H. armigera 54. Genomic DNA was isolated from individual third instar larva, using the procedure described by Wang et al 55. Gene-specific primers based on the known sequence of the cDNA (accession number: DQ256407) and four general primers provided by the Genome Walking Kit (TaKaRa, Dalian, China) were used for every genome walking. PCR products were cloned into pGEM-T Easy vector (Promega, Madison, WI, USA) and sequenced.

Database search strategy. Database searches were performed and comprise four steps. Firstly, the Lep1-like element (named as HaLep1_1) identified in the 5'-flanking sequence of the H. armigera P450 gene, CYP6AE12, was compared with NCBI H. armigera nucleotide collection (nr/nt) databases with Blastn (www.ncbi.nlm.gov/cgibin/BLAST), sequences of high homology as well as 500 bp upstream and downstream flanking regions were extracted and analyzed for hallmarks of Lep1 Helitrons such as characteristic 5'-TC and 3'-CTRY nucleotide termini as well as CTRR motif at the 3' end of acquired sequence. Secondly, nucleotide (nr/nt) and EST (est_others) collections were searched using HaLep1_1 as query to detect sequences with high identity with HaLep1_1 in lepidopteran species other than H. armigera. Thirdly, the 134-bp lepidopteran-specific common sequence 3 (LSCS3, Lep1) was searched against non-lepidopteran insect genome sequences, including BeetleBase (http://beetlebase.org/), AphidBase (http://www.aphidbase.com/aphidbase/), NasoniaBase (http://hymenopteragenome.org/nasonia/), BeeBase (http://hymenopteragenome.org/beebase/), vectorbase (https://www.vectorbase.org), and
Ant Genomes Portal (http://hymenopteragenome.org/ant_genomes/). Finally, the 134 bp Lep1 consensus sequence was compared with NCBI non-lepidopteran databases with Blastn, including the whole genome shotgun, nucleotide collection (nr/nt), genome survey sequences, high throughput genomic sequences, and expressed sequence tag databases. Hits that were ≥70% identical to the query over >100 bp were examined and, when possible, full-length Lep1-like elements were manually extracted. These elements were used as queries to find additional related Lep1 Helitrons, the resulting hits were examined, and full-length elements were extracted.

**Assessing polymorphism.** In *H. armigera*, using one pair of primers flanking the insertion site, *HaLep*1 insertion polymorphism was assessed by performing a PCR survey, which yielded products of different sizes in *HaLep*1 insertion individual (about 700 bp) and non-insertion individual (about 500 bp). To further illustrate the mobility of other *HaLep*1 elements, the insertion polymorphisms were also assessed by homology searches. Briefly, paralogous or orthologous sites not containing a *HaLep1* insertion (empty sites) were identified by homology searches utilizing Blastn with a query constructed from the sequences directly flanking the insertion site. The chromic query sequence (about 200 bp in length) was created by extracting both the flanking sequence upstream from the element insertion (about 100 bp) and the flanking sequence downstream from the element insertion (about 100 bp).

**Sequence analysis.** Multiple sequence alignments were performed using ClustalW2 with default settings. Neighbor-joining (NJ) and maximum likelihood (ML, using the DNA sequencing, 10 kb sequences in each direction (upstream and downstream) of analysis with 1000 replicates. To detect putative cross-species contamination during DNA sequencing, 10 kb sequences in each direction (upstream and downstream) of each representative non-lepidopteran Lep1 insertion were extracted from the BAC clone sequences and used to search against the non-redundant databases using the NCBI server with Blastx (www.ncbi.nlm.nih.gov/blast/BLAST).

1. Wicker, T. et al. A unified classification system for euchromatic transposable elements. *Nat. Rev. Genet.* 8, 973–982 (2007).
2. Feschotte, C. & Pritham, E. J. DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.* 41, 331–68 (2007).
3. Hartl, D. L., Lohe, A. R. & Lozovskaya, E. R. Modern thoughts on an ancient mariner: function, evolution, regulation. *Annu. Rev. Genet.* 31, 37–58 (1997).
4. Silva, J. C., Loreto, E. L. & Clark, J. B. Factors that affect the horizontal transfer of transposable elements. *Curr. Issues Mol. Biol.* 6, 57–71 (2004).
5. Gilbert, C., Pace, J. K. & Feschotte, C. Horizontal SPINning of transposons. *Commun. Integr. Biol.* 2, 117–119 (2009).
6. Loreto, E. L., Carareto, C. M. & Capy, P. Revisiting horizontal transfer of transposable elements in *Drosophila*. *Hereditas* 100, 545–554 (2008).
7. Schacka, S., Gilbert, C. & Feschotte, C. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.* 25, 337–546 (2010).
8. Hartl, D. L., Lozovskaya, E. R., Nurminskey, D. I. & Lohe, A. R. What restricts the activity of mariner-like transposable elements? *Trends Genet.* 13, 197–201 (1997).
9. Robertson, H. M. [Evolution of DNA transposons in eukaryotes]. *Annu. Rev. Genet.* 2, 565–660 (2014).
10. Heimpel, G. E., Neuhauser, C. & Hoogendörn, M. Effects of parasitoid fecundity and host resistance on indirect interactions among hosts sharing a parasitoid. *Ecol. Lett.* 6, 556–566 (2003).
11. Cameron, P. J. & Walker, G. P. Host specificity of *Cotesia rubecula* and *Cotesia plutellae*, parasites of white butterfly and diamondback moth. *Proc. 50th NZ Plant Protection Conf.* 236–241 (1997).
12. Whitfield, B. J. Estimating the age of the polydnavirus/braconid wasp symbiosis. *Proc. Natl. Acad. Sci. USA* 99, 7508–7513 (2002).
13. Thomas, J., Schacka, S. & Pritham, E. J. Pervasive horizontal transfer of rolling-circle transposons among animals. *Genome Biol. Evol.* 2, 654–660 (2014).
14. Wang, I. et al. Characterization of three novel SINE families with unusual features in *Helicoverpa armigera*. *PLoS ONE* 7, e31355 (2012).
15. Hotopp, J. C. D. et al. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317, 1753–1756 (2007).
16. Nikoh, N. & Nakabachi, A. Natural acquired symbiotic genes via lateral gene transfer. *BMC Biol.* 7, 12 (2009).
17. Oliver, K. M., Degnan, P. H., Burke, G. R. & Moran, N. A. Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annu Rev. Entomol.* 55, 247–266 (2010).
18. Husnik, F. et al. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* 153, 1567–1578 (2013).
19. Sloan, D. B. et al. Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Mol. Biol. Evol.* 31, 853–871 (2014).
20. Jayaprakash, A. & Hoy, M. A. Long PCR improves *Wolbachia* DNA amplification: WSP sequence found in 76% of thirty-six arthropod species. *Insect Mol. Biol.* 9, 393–405 (1998).
21. Hilgenboecker, K., Hamburger, P., Schallmayer, P., Telschow, A. & Wernery, J. H. How many species are infected with *Wolbachia*? a statistical analysis of current data. *FEBS Microbiol. Lett.* 281, 215–220 (2008).
22. The International Aphid Genomics Consortium. Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol.* 8, e1000313 (2010).
23. Kashkarova, L. & Khalakhanov, A. Range of the hosts of the causal agent of potato *Nosema bombi* in the mulberry silkworm. *Parasitologia* 14, 164 (1980).
24. Hayasaka, S. & Yonemura, N. Infection and development of *Nosema bombycis* (Microsporida: Protozoa) in several lepidopteran insects. *JARQ* 33, 65–68 (1999).
25. Kermani, N. et al. Pathogenicity of *Nosema sp.* (Microsporida) in the diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae). *PLoS One* 8, e62884 (2013).
26. Zhang, Y. A., Wang, Y. Z. & Zhang, L. Preliminary study on a new pathogen (*Nosema glabrifrons* Zhang) parasitizing the longhorned beetle *Anoplophora glabripennis* (Motsch). *Scientia Silvae Sinicae* 39, 217–173 (2003).
Acknowledgments

We thank the editor Chuan-Xi Zhang and anonymous referees for their helpful comments on the manuscript. This work was supported by the National Natural Science Foundation of China under grant no. 30871642.

Author contributions

Conceived and designed the experiments: J.W. Performed the experiments: J.W., X.G., J.G. Analyzed the data: J.W., X.G., J.G., F.L. Wrote the paper: J.W., F.L.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Guo, X.Z., Gao, J.K., Li, F. & Wang, J.J. Evidence of horizontal transfer of non-autonomous Lep1 Helitrons facilitated by host-parasite interactions. Sci. Rep. 4, 5119; DOI:10.1038/srep05119 (2014).

This work is licensed under a Creative Commons Attribution 3.0 Unported License. The images in this article are included in the article’s Creative Commons license, unless indicated otherwise in the image credit; if the image is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the image. To view a copy of this license, visit http://creativecommons.org/licenses/by/3.0/