Published in final edited form as:

Insect Mol Biol. 2011 June ; 20(3): 291–301. doi:10.1111/j.1365-2583.2010.01061.x.

Natural history and intragenomic dynamics of the *Transib* transposon *Hztransib* in the cotton bollworm *Helicoverpa zea*

Erxia Du^{1,2}, Xinzhi Ni³, Huiyan Zhao¹, and Xianchun Li^{2,*}

¹ Department of Plant Protection, Northwest Agriculture and Forestry University, Yangling, Shaanxi, 712100, China

² Department of Entomology, University of Arizona, Tucson, AZ, 85719, USA

³ USDA-ARS Crop Genetics and Breeding Research Unit, Tifton, GA 31793, USA

Abstract

Hztransib recently identified from *Helicoverpa zea* represents the first intact and transcriptionally active *Transib* element. Its open reading frame was detected in *H. armigera*, from which *H. zea* evolved, and *H. assulta*, the common ancestor of *H. zea* and *H. armigera*. But its remaining parts were found only in *H. armigera*. 39 *Hztransib* insertion sites, all of which are polymorphic, were detected from 8 populations of *H. zea*. Out of the 39 insertion sites, 35 were not frequently occupied, with 1 to 33 occurrences in a total of 128 individuals from the 8 populations (16 larvae per population). Its copy number ranged from 5.8 to 14.2 per individual, with putative intact copies always more abundant than internally-deleted ones. Taken together, *Hztransib* had probably transferred to *H. zea* from *H. armigera* and most likely still retains its capacity to maintain structural integrity, increase copy number and remobilize in *H. zea*.

Keywords

Helicoverpa spp.; *Hztransib*; intragenomic dynamics; natural history; transposon display

Introduction

Helicoverpa zea is one of the most destructive pests in the New World. It claims the equivalent of two million acres of corn each year (Metcalf & Metcalf, 1993). In serious infestations, 70 to 98% of ears of field corn can be damaged. It is also a serious pest of cotton (Ring & Benedict, 1993), tomato (Metcalf & Metcalf, 1993; Mcleod *et al.*, 1996), and soybean (Biever *et al.*, 1983). Since the introduction of Bt-transgenic cotton and corn in 1996, the relative importance of *H. zea* has increased considerably, because it is much more tolerant to Bt toxins than other key pests such as the tobacco budworm and European corn borer (Luttrell *et al.*, 1999). It has been ranked as the No. 1 pest on corn and cotton in recent years, especially in Arkansas, Georgia, Louisiana, Mississippi, Missouri, New Mexico, North Carolina, Oklahoma, South Carolina and Tennessee (Robinson, 2001; Brandon, 2004).

Polyphagy and rapid acquiring of insecticide resistance are among the two most important traits contributing to the tremendous success of *H. zea* in a wide range of agro-ecosystems (Fitt, 1989; Mitter *et al.*, 1993). *H. zea* exploits over 100 different plant families (Kogan *et*

*Corresponding Author: Dr. Xianchun Li, Department of Entomology and BIO5 Institute, University of Arizona, Tucson, Arizona 85721, lxc@email.arizona.edu, Phone: (520)626-1749, Fax: (520)621-1150.

al., 1978) and commonly attacks at least 15 cultivated plants including alfalfa, beans, cotton, okra, peanuts, peas, strawberries, sweet peppers, sweet potatoes, tobacco, tomato, corn, sorghum, and soybean. It is capable of rapidly evolving resistance to both conventional insecticides (Abd-Elghafar *et al.*, 1993; Brown *et al.*, 1998; McCaffery, 1998; <http://www.vegedge.umn.edu/ZeaMap/zeamap.htm>) and Bt toxins (Burd *et al.*, 2003). Its capacity to evolve resistance is demonstrated by the fact that it is the first pest with field-evolved resistance to Bt transgenic crops (Tabashnik *et al.*, 2008; Ali *et al.*, 2006).

Cytochrome P450 monooxygenases (P450s) are among the most important detoxification enzymes that facilitate *H. zea* to cope with a tremendous diversity of plant toxins (known as allelochemicals) distributed among its potential host plants and to rapidly evolve resistance to insecticides (Li *et al.*, 2007). Eight P450 genes distributed in the *CYP6B* (*CYP6B8*, *CYP6B9*, *CYP6B27*, *CYP6B28*), *CYP9A* (*CYP9A12*, *CYP9A14*) and *CYP321A* (*CYP321A1*, *CYP321A2*) subfamilies have been characterized from this species (Li *et al.*, 2000, 2002a; Sasabe *et al.*, 2004; Chen & Li 2007). The participations of these *H. zea* P450 genes in allelochemical tolerance and/or insecticide resistance are suggested by their transcriptional activation in response to a range of allelochemicals (xanthotoxin, indole-3-carbinol, chlorogenic acid, and flavone) and the plant defense signaling molecules jasmonate and salicylate naturally encountered in host plants as well as synthetic chemicals not naturally encountered (cypermethrin and phenobarbital) (Li *et al.* 2000, 2002b, 2002c; Li & Ni, 2009; Zhang *et al.*, 2010; Niu *et al.*, 2008; Sasabe *et al.*, 2004; Wen *et al.*, 2009; Zeng *et al.*, 2007; Li and Chen, unpubl. data). More directly, *CYP6B8* and *CYP321A1* have been shown to metabolize a range of allelochemicals (xanthotoxin, chlorogenic acid, quercetin, and flavone) and insecticides (diazinon, cypermethrin, and aldrin) (Li *et al.*, 2004, 2007; Niu *et al.*, 2008; Rupasinghe *et al.*, 2007; Sasabe *et al.*, 2004; Wen *et al.*, 2009).

The aforementioned xenobiotic-metabolizing P450 loci are significantly enriched with transposable elements (TEs) (Chen & Li, 2007). Scanning two alleles (one from a laboratory colony, another from a cell line) per each of six xenobiotic-metabolizing *H. zea* P450 genes (*CYP6B8*, *CYP6B27*, *CYP9A12v3*, *CYP9A14*, *CYP321A1* and *CYP321A2*) has led to identification of twelve novel TEs from their introns, exons or flanking regions (Chen & Li, 2007). These include LINEs (long interspersed nuclear elements), SINEs (short interspersed nuclear elements), MITEs (miniature inverted-repeat transposable elements), two full-length DNA transposons. Upon the transposition and ectopic recombination of these TEs, these xenobiotic-metabolizing P450 loci can gain TE-introduced selectively advantageous variations necessary for *H. zea* to cope with the diversity and unpredictability of plant toxins and insecticides it encounters in a wide range of agro-ecosystems.

One of the twelve TEs, *Hztransib*, is a member of *Transib*, a novel superfamily of DNA transposons originally reconstructed *in silico* from ancient fossil TE sequences in the genomes of *Drosophila melanogaster* and *Anopheles gambiae* (Kapitonov & Jurka, 2003, 2005). The reconstructed *Transib* transposons are 3–4 kb long and have terminal inverted repeats (TIRs) of variable length flanked by 5-bp ‘GC’-rich target site duplications (TSDs) (Kapitonov and Jurka 2003, 2005; Feschotte and Pritham, 2007; Chen & Li, 2008). The encoded transposases share a surprising level of similarity with the core region of the V (variable), D (diversity), and J (joining) (V(D)J) recombination protein RAG1 (Kapitonov and Jurka, 2005). There is also significant sequence similarity between *Transib* TIRs and RAG1/RAG2 recombination signal sequences (RSSs) (Kapitonov and Jurka, 2005). These features support the long-standing hypothesis that RAG1/RAG2, the key recombinase for the V (D) J immune system, may be derived from DNA transposons (Agrawal *et al.*, 1998; Messier *et al.*, 2003; Sakano *et al.*, 1979; Thompson, 1995).

Degenerate *Transib* elements are now known to also be present in the genomes of many other insects, dog hookworm, freshwater flatworm, hydra, the fungus soybean rust, and the plant potato (Chen & Li, 2008). *Hztransib* characterized from the 5'-flanking region of *CYP6B8* in the *H. zea* midgut cell line is by far the only full-length, structurally-intact, and transcriptionally active *Transib* TE (Chen & Li, 2007, 2008). Two copies (one full-length and one internally-deleted) of *Hztransib* have been characterized from the *H. zea* lab colony and midgut cell line (Chen & Li, 2008). The intact *Hztransib* is 3518-bp long and is flanked by 5-bp GC-rich (CGTCG) TSDs. It contains a 5' TIR of 552 bp, a 3' TIR of 502 bp, a single intact ORF (open reading frame) of 1524 bp, and its own promoter sequence (381 bp). It is transcribed as a typical 3'-truncated nonstop mRNA (an mRNA lacking a stop codon) in the larval midgut, fat body, and ovary as well as in the midgut cell line. All these structural and transcriptional features, together with its insertion dimorphism between the laboratory colony and the midgut cell line, suggest that *Hztransib* may represent the first intact and active *Transib* TE found in any organism (Chen & Li, 2007, 2008). In this study, transposon display and genomic PCR-gel analyses were conducted to examine the history and population dynamics of *Hztransib* in four laboratory and four field populations of its natural host organism *H. zea*. Its great insertion polymorphism and variations in copy number among populations and individuals documented suggest that *Hztransib* is probably still active in its host genome.

Results

Presence of *Hztransib* in three *Helicoverpa* species

To determine whether *Hztransib* invaded *H. zea* genome before or after *H. zea* evolved from *H. armigera* via a founder event (Behere *et al.*, 2007), three types of genomic PCR (full-length, ORF-flanking, ORF) were conducted using DNA extracted from a pool of 16 larvae of *H. zea*, *H. armigera* and *H. assulta*, respectively. The full-length PCR with the single primer *Hztransib*IR designed to amplify the full-length or internal deleted copies of *Hztransib* revealed the presence of the putative full-length copy and two internally-deleted copies of *Hztransib* in both *H. zea* and *H. armigera*, but not in *H. assulta* (Fig. 1A), the common ancestor of *H. armigera* and *H. zea* (Behere *et al.*, 2007). Consistent with the results of the full-length PCR, the ORF-flanking PCR with the primers flanking-F and flanking-R annealing to the 5' and 3' UTR of *Hztransib* respectively (see Fig. 2) showed the presence of the putative intact copy and two internally-deleted copies in both *H. zea* and *H. armigera*, but not in *H. assulta* (Fig. 1B). However, the ORF PCR with *transib*-F and *transib*-R annealing the 5'- and 3'-most sequences of the *Hztransib* ORF showed the presence of the intact ORF of *Hztransib* in all of three species (Fig. 1C). These indicate that *H. zea* and *H. armigera* share two forms of *Hztransib* (putative full-length and internally-deleted) that are conserved not only in ORF, but also in 5' UTR, 3' UTR, and TIR. By contrast, the *H. assulta* *Transib* elements are only conserved in ORF with *Hztransib* in *H. zea* and *H. armigera*.

Copy number and insertion site of *Hztransib* in *H. zea*

Transposon display using the adapter-ligated *MspI*-cut genomic DNA fragments from individual larvae and the *Hztransib*-specific primers Gsp1 and Gsp2 designed based on the 54-bp fragment unique to the left TIR (Fig. 2) was conducted to assess *Hztransib*'s insertion sites and copy number in the 8 field or laboratory populations of *H. zea* (Table 1). The 5'-flanking sequences from each *Hztransib* copy to its first 5' *MspI* site in a given individual genome were PCR-amplified with the Cy5-labeled primer Gsp2 and the adapter primer *MspI*-a and visualized on a 6% denaturing polyacrylamide gel using a storm 860 phosphorimager (see a representative gel in Fig. 3). The number of insertion sites per individual is estimated as equal to the number of bands on a gel. This is probably an

underestimate because some degenerate *Transib* elements found in other invertebrates lose one side of their TIR (Chen & Li, 2008). If some of *Hztransib* copies lose their TIR, the transposon display would miss them. For any of the 8 populations or the whole species, its total insertion sites were equal to the sum of different sizes of bands detected in all the individuals from that particular population or the whole species. In other words, a transposon display band that occurs (is occupied) in two or more individuals of a given population or the whole species can only be counted as one insertion site.

Based on this conservative estimation, 39 *Hztransib* insertion sites were detected from a total of 128 larvae (16 larvae per population, $N=16 \times 8 = 128$) randomly chosen from each of the 8 natural or laboratory populations. The total insertion sites per population ranged from 14 in the Benzon, Bt corn and USDA populations to 24 in the Pop2 population (Table 2). The minimum and maximum insertion sites per individual were also quite variable within and between populations (Table 2). The minimum number of insertion sites per individual was 1, found in 1 larva from the Bt corn field population, and 1 larva from the NCSU laboratory strain. By contrast, the maximum number of insertion sites was 11, found in 3 larvae from the DK6971 field population and 1 larva from the Benzon laboratory strain. The insertion site difference within each strain ranged from 5 to 9 sites. The average insertion sites per individual for each population varied from 4.5 in the USDA laboratory strain to 6.56 in the DK6971 field population (Table 2). Multiple comparisons indicated that the average insertion sites per individual were significantly higher in the DK6971 and Benzon strains than in the other six strains (Turkey's HSD test, $p < 0.05$; Table 2). The equilibrium copy number of *Hztransib* per individual ($\bar{n}$ in Table 2) ranged from 5.83 copies in the Benson strain to 14.21 copies in the Pop2 population (Table 2). These data suggest that *Hztransib* might be still capable of increasing its copy number in its host genome.

Insertion site polymorphism and site-occupancy frequency of *Hztransib* in *H. zea*

The 39 sites detected from the 8 populations were numbered/named from 1 to 39 according to the sizes of the corresponding bands on transposon display gels (Figs 3–4), with site 1 and 39 representing the largest (2.5 kb) and smallest (0.3 kb) bands, respectively. For the *Hztransib* copies that have no *MspI* site within their left TIR, the sizes of the resulted bands should be at least equal to 407 bp. Site 1 to 37 belonged to the *Hztransib* copies without an *MspI* site within their left TIR, whereas site 38–39 were the *Hztransib* copies with an *MspI* site within their left TIR (Figs 3–4).

None of the 39 insertion site is fixed at the species level, since no single site was occupied in all the 128 individuals (Fig. 4). The site-occupancy (or -occurrence) frequency of all the 39 insertion sites in any of the 8 populations (Fig. 5) or the whole species (Fig. 4), and the overall site occupancy of each population (Table 2), defined as the ratio of the sum of occurrences of all the insertion sites to the total insertion sites (Arensburger *et al.*, 2005) in a population, can be used to infer the activity and mobility of a transposable element (Charlesworth & Charlesworth, 1983; Arensburger *et al.*, 2005; Subramanian *et al.*, 2007). Only 4 sites (site 5, 24, 25, 37) were occupied in more than 40% of the 128 larvae (Fig. 4); these were presumably ancient and less mobile insertion sites. The remaining 35 sites were not frequently occupied, with 1 (unique site 4, 6, 11, and 31) to 33 (site 30) occurrences in the 128 individuals (Fig. 4). Twenty-two sites occurred less than 10 times in the 128 individuals. The unique and low-frequent sites presumably represent recently-occupied sites. All these data implicate that *Hztransib* might be currently active at the species level.

At the population level, there were great variations in the number of total and unique insertion sites and the site-occupancy frequency among the eight populations (Table 2; Fig. 5). The Bt corn and USDA populations had the lowest number of total insertion sites (14), whereas the Pop 2 strain had the highest total insertion sites (24). In terms of the number of

unique sites within each population, the Pop 2, field corn, and NCSU populations had more unique sites (8 or 9) than the other populations (Table 2; Fig. 5). By contrast, the Benzon and Bt cotton populations had only 2 and 3 unique sites, respectively. Moreover, both the Benzon and Bt cotton populations had three high frequency sites with a site-occupancy frequency of ≥ 13 (Fig. 5). These data imply that *Hztransib* is probably most active and mobile in the Pop 2, field corn, and NCSU populations, but least active in the Benzon and Bt cotton populations.

Frequency of the intact copy of *Hztransib* in *H. zea*

The structural integrity of *Hztransib* copies in all the 128 larvae from the eight *H. zea* populations used for DNA transposon display were individually analyzed by the full-length, ORF-flanking, and ORF PCR gel analyses described above (see Figs 6–7 for the corresponding representative gel pictures). All the three PCR gel analyses revealed the 100% presence of the putative intact *Hztransib* copy and the complete ORF in every individual larva of the 8 populations (Table 3; Figs 6–7). The frequency of the internally-deleted copies, however, varied greatly among the 8 populations (Table 3). Extremely noticeable difference in the frequency of the internally-deleted copies was the contrast between the field populations and the laboratory strains. In 3 out of the 4 laboratory strains, 93% (NCSU strain) or 100% (Benzon and USDA strains) larvae tested had internally-deleted copies in the genome (Table 3). By contrast, the frequency of the internally-deleted copies ranged from 0% (Bt corn population) to 16.9% (Field corn population) in the 4 field populations (Table 3). Furthermore, when both the putative intact and internally-deleted copies existed in the same individuals, there were always more putative intact copies than deleted copies regardless of populations. This is because the bands representing the putative intact copies were always stronger than the bands representing the deleted copies (Figs 6 and 7A). These data indicate that *Hztransib* has a low deletion rate, particularly in the field populations.

Discussion

Almost all members of the *Transib* superfamily of TEs identified so far are inactive degenerate TE (Chen & Li, 2008). *Hztransib* characterized from *H. zea*, however, is structurally intact and transcriptionally active (Chen & Li, 2008). To address its evolutionary history and intragenomic dynamics, genomic PCR-gel analyses and transposon display were conducted to reveal its presence in the genomes of three closely-related *Helicoverpa* species and its insertion polymorphism in *H. zea* field and laboratory populations. Our genomic PCR-gel analyses shown that the putative intact ORF of *Hztransib* was present not only in *H. zea*, but also in *H. armigera*, from which *H. zea* seem to have been established via a founder event at around 1.5 million years ago (Mallet *et al.*, 1993; Li *et al.*, 2002c; Behere *et al.*, 2007), and *H. assulta* (Fig. 1), which is basal to *H. armigera* and *H. zea* (Mitter *et al.*, 1993; Behere *et al.*, 2007). However, the two types of *Hztransib* (full-length and internally-deleted) detected in *H. zea* by the full-length and ORF-flanking PCR were shared only by *H. zea*'s closest relative *H. armigera*, not by their common ancestor *H. assulta* (Fig. 1). This suggests that *Hztransib* had most likely invaded *H. armigera* genome at least 1.5 million years ago and then passed to *H. zea* genome via vertical transmission when *H. zea* was established from a founder population of *H. armigera* in the New world. The fact that the full-length intact form of *Hztransib* has survived for at least 1.5 million years in *H. zea* and *H. armigera* suggests that *Hztransib* may be less prone to vertical inactivation/removal as it has been seen with *Transib* elements in *Drosophila* spp., *Bombyx mori*, *Tribolium castaneum*, mosquitoes, and other organisms (Kapitonov & Jurka, 2003, 2005; Chen & Li, 2008).

Although at least approximately 1.5 million years have been passed since *H. zea* inherited *Hztransib* from its *H. armigera* founder stock, *Hztransib* was still present in all the 128 individuals from the 4 laboratory and 4 natural populations of *H. zea*. This indicates that *Hztransib* has become fixed in *H. zea*. However, the loci where *Hztransib* copies inserted within the *H. zea* genome are not fixed because none of the 39 insertion sites documented in this study was shared by the all 128 individuals (Fig. 4). In fact, most of the 39 genomic sites were low frequency (1.5% – 26%) or unique sites, with only four of them having an occurrence frequency of >40% (Fig. 4). Such a high insertion polymorphism and low site-occupancy frequency across populations and the entire species (Figs 4–5, Table 2) strongly support that *Hztransib* probably is still active and has recently remobilized in *H. zea*. Otherwise, many of the 39 insertion sites should be high frequency or even fixed sites, rather than low frequency or unique sites.

The transposition activity of *Hztransib* is also supported by its variances in the number of insertion sites and equilibrium copy number per individual among the 8 populations (Table 2). The number of *Hztransib* insertion sites per individual varied greatly within and among populations and ranged from 1 to 11 across entire species' samples (Table 2). On average, each individual had 4.50–6.56 insertion sites (bands) and 5.83 to 14.21 copies, depending on populations (Table 2). According to the theory on the population dynamics of TEs (Charlesworth & Charlesworth, 1983; Montgomery *et al.*, 1987; Mackay, 2007), the extra copies carried by some individuals can only occur through transposition of replicates of pre-existing elements to novel insertion sites. The wide range of variance in *Hztransib* copy number per individual seen among the 8 populations (5.83 to 14.21 copies, Table 2) implies its replication potential for further increase of its copy number in *H. zea* genome.

Two types of *Hztransib* copies—the full-length intact ones and the internally-deleted ones—were detected previously (Chen & Li, 2008) and in this study. The full-length intact *Hztransib* copies have become fixed in *H. zea* as all the 128 individuals had at least one putative full-length intact copy (Table 3). Furthermore, no matter how many copies were present in any of the 128 individuals, the putative full-length intact *Hztransib* was always more abundant than the internally-deleted *Hztransib* (Figs 6–7). Such a relative abundance of the intact vs. deleted copies of *Hztransib* in the 8 populations of *H. zea* makes it an ideal candidate as a gene drive in terms of linkage between the drive and the refractory genes.

In summary, our data suggest that *Hztransib* probably still retains the capacity to maintain its structural integrity, increase its copy number, and remobilize in its natural host *H. zea*. Further transposition and post-integration studies are needed to directly demonstrate the excision, integration, post-integration duplication, and remobilization of *Hztransib* in insects and insect cells. If such studies demonstrate its capacity to transpose, duplicate, and remobilize in insects, particularly in mosquitoes, *Hztransib* could be developed as a gene drive for replacing a wild vector population with a refractory one to control mosquito-borne diseases.

Experimental procedures

Insects and DNA extraction

Larvae samples of *H. armigera* and *H. assulta* used in this study were generously provided by Dr. Weihua Ma of Huazhong Agricultural University and Dr. Sufen Bai of Henan Agricultural University, respectively. 4 laboratory strains (Benzon, USDA, Pop 2 and NCSU) and 4 field populations of *H. zea* (Table 1) were used in this study. The 4 *H. zea* laboratory strains were maintained in an insectary kept at 28°C with a photoperiod of 16 h light: 8 h dark on a semi-synthetic diet containing wheat germ (Waldbauer *et al.*, 1984). The 4 *H. zea* field populations were established from approximately 150 larvae collected from

either corn or cotton fields in Indiana or Georgia, USA (Table 1). These larvae and their progeny were maintained on the same semi-synthetic diet as the four laboratory strain. In F1 or F2 generation (Table 1), at least 16 5th instar larvae were randomly picked from each of the 4 field populations, individually flash-frozen in liquid nitrogen and stored in -80°C for subsequent genomic DNA isolation and transposon display. We also froze at least 16 5th instar larvae from each of the 4 laboratory strains for transposon display. For PCR detection of the putative full-length and internally-deleted *Hztransib* copies in the three *Helicoverpa* species, a pool of 16 larvae were used for extraction of a pool DNA sample for each species using the procedure described in Li *et al.* (2002c).

PCR detection of *Hztransib* in the three *Helicoverpa* species

Three types of genomic PCR [full-length, open reading frame (ORF), and ORF-flanking] were conducted to detect the presence of *Hztransib* and its structural integrity in the three *Helicoverpa* species. The single primer for the full-length PCR was *Hztransib*IR (5'-CACGGTGGATCGAAAATCGGC-3'). Since *Hztransib*IR anneals to the left and right terminal invert repeat (TIR) of *Hztransib* (see its annealing location in Fig. 2), the full-length PCR was employed to detect the presence of the putative full-length intact (3518 bp) and internally-deleted (<3518 bp) copies of *Hztransib*. The ORF PCR with the primers *transib*-F (5'-ATGCTCCAAAATCCAGAAGCAGCAG-3') and *transib*-R (5'-TCAGAAC TCTGTCGGAGCTTCCTGA-3') (see Fig. 2 for their annealing locations) was used to detect the presence of the ORF-intact (1524 bp) and ORF-internally deleted (<1524 bp) copies of *Hztransib*. In case substitutions may occur at the regions to which the ORF PCR primers *transib*F and *transib*R anneal, the ORF-flanking PCR with the primers *flanking*-F (5'-GATGAATATAAAGGCGATGACCAAT-3') and *flanking*-R (5'-GTTAAAAGGATACGTGCCATAATG-3') (see Fig. 2 for their annealing locations) was used to detect the presence of the intact (1809 bp), ORF-internally deleted (<1524 bp) and untranslated region-(UTR-) or ORF-internally deleted (>1524 bp, but <1809 bp) copies of *Hztransib*.

The PCR set up and cycling conditions were the same for the three types of PCR except for the extension time of each cycle. PCR amplifications were carried out in a 25 μl reaction mixture containing 2.5 μl of $10\times$ PCR buffer, 0.5 μl of dNTP (10 mM), 0.25 μl of Taq polymerase (5 U/ μl), 0.5 μl of the specific primer pairs (20 μM), and 0.2 μg of the template DNA. PCR cycling conditions were as follows: 5-min initial denaturation at 94°C , followed by 35 cycles of 1-min denaturation at 94°C , 0.5-min annealing at 60°C , 3-min (full-length PCR) or 2-min (ORF and ORF-flanking PCR) extension at 72°C , and a 10-min final extension at 72°C . PCR products were run on a 1% agarose gel in $1\times$ TAE buffer and visualized by staining with ethidium bromide.

Hztransib transposon display

The 5'-flanking sequences of all *Hztransib* insertions/copies within the genome of each of the 128 larvae randomly picked from the 8 *H. zea* populations (16 larvae per population, $N=16\times 8=128$) were PCR-amplified and revealed by transposon display (Fig. 2). In brief, 2.5 μg of genomic DNA from each larva was digested by 8 U of *Msp* I in a 100 μl reaction mixture at 37°C for 16–18 h and then purified using PureLink PCR Purification Kit (Invitrogen, CA, USA). Approximately 500 ng of the purified *Msp* I-digested fragments were ligated to 60 pmol of *Msp* I adapters, which were prepared by annealing equal amounts of the oligonucleotides *Msp*I-a (5'-CAGACATGAGTCCTGAGA-3') and *Msp*I-b (5' (p)-CGTCTCAGGACTCAT-NH₂-3') (heating in a boiling water bath for 10 min, then gradually cooling down to room temperature), in a 20 μl reaction mixture containing $1\times$ ligation buffer, 1 U of T4 DNA ligase and 4 U of *Msp* I (for cutting fragment-fragment ligations but not fragment-adaptor ligations) at 22°C for 16–18 h.

The resulting DNA fragments were used as templates to PCR-amplify the 5'-flanking sequences of all *Hztransib* copies using the general forward primer MspI-a complementary to the *Msp I* adapter sequence and the two *Hztransib* left ITR-specific reverse primer Gsp1 (5'-GAAAATCACAAAGTGGGTAATACGACT-3') and Cy5-labeled Gsp2 (5'-GACT TTGTGATTTCTTAGAATATTTTGGTC-3') (Fig. 1). The pre-selective PCR reaction system (25 µL) contained 2.5 µL of 10 × PCR buffer, 0.5 µL of dNTP (10 mM), 0.25 µL of Taq polymerase (5 U/µL), 0.5 µL of MspI-a (20 µM), 0.5 µL of Gsp1 (20 µM) and 1 µL of 10-fold dilution of the ligation products. PCR amplifications began with 94°C denaturation for 5 min, followed by 25 cycles of 94°C denaturation for 5 min, 55°C annealing for 0.5 min, and 72°C extension for 2.5 min, and a final 72°C extension for 10 min. The selective PCR set up was the same with the pre-selective PCR except that 1) 1 µL of 20-fold dilution of the pre-selective PCR products, rather than 1 µL of 10-fold dilution of the ligation products, was used as the template; and 2) the reverse primer was the Cy5-labeled Gsp2 (HPLC-grade pure). The selective PCR was carried out according to the following cycling program: initial denaturation at 94 °C for 5 min, followed by 5 touchdown cycles of 94 °C denaturation for 1 min, annealing for 1 min at 65–60 °C going down by 1 °C/cycle, and 72 °C extension for 2.5 min, and 25 cycles of 94 °C denaturation for 1 min, 60 °C annealing for 1 min and 72 °C extension for 2.5 min, and a final extension of 10 min at 72 °C.

The resultant 25 µL selective PCR products were mixed with 20 µL of formamide loading dye (98% formamide, 0.005% bromophenol blue, 0.005% xylene cyanol, 10 mM EDTA). After denaturation by heating 10 min at 95 °C and then chilling more than 5 min on ice, 5 µL of the 45 µL sample was electrophoresed (120 volts, 0.2 A) for 12 h at 4 °C under dark on a 6% denaturing polyacrylamide gel, which had been pre-run for 30 min in the same conditions. After electrophoresis, the gel was fixed with 10 % ethanol for 30 min, transferred to 3 mm filter paper, and dried at room temperature for 20 min. The dried gel was scanned on a storm 860 phosphorimager (GE Healthcare/Molecular Dynamics, Piscataway NJ).

Data analysis

All the bands detected on the transposon display gels of the 128 *H. zea* larvae were numbered and named based on their size. Since denaturing polyacrylamide gels can effectively separate DNA fragments that differ by as little as fewer than 10 base pairs, each band was considered to represent an insertion site. Presences (recorded as 1) or absences (recorded as 0) of these bands (sites) in the 16 individuals of each of the 8 *H. zea* populations were recorded in Excel files to calculate the intragenomic dynamics parameters of *Hztransib* including its total, unique, and average insertion sites, site-occupancy frequency, and equilibrium copy number. The equilibrium copy number per individual ($\hat{n}$)

was estimated by the formula $\hat{n} = \sum_{i=1}^m (1 - \sqrt{1 - Z_i})$, where Z_i is equal to the frequency of the 'null' allele for the i th insertion site (Wright *et al.*, 2001). Differences in average insertion sites per individual among populations were evaluated by Turkey's HSD (Honestly Significant Difference) test at $P < 0.05$ using JMP 8.0 statistical software.

Acknowledgments

We thank Dr. Chonglie Ma for technical assistance and Dr. Sufen Bai and Dr. Weihua Ma for providing the *H. assulta* and *H. armigera* samples. We also thank Dr. Fred Gould for kindly providing the NCSU *H. zea* laboratory strain. This work was supported by NIH grant 1R21AI083680-01 to X. Li and China Scholarship Council scholarship to E. Du.

References

- Abd-Elghafar SF, Knowles CO, Wall ML. Pyrethroid resistance in two field strains of *Helicoverpa zea* (Lepidoptera: Noctuidae). *J Econ Entomol*. 1993; 86:1651–1655.
- Agrawal A, Eastman Q, Schatz D. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature*. 1998; 394:744–751. [PubMed: 9723614]
- Ali MI, Luttrell RG, Young SY III. Susceptibilities of *Helicoverpa zea* and *Heliothis virescens* (Lepidoptera: Noctuidae) populations to Cry1Ac insecticidal protein. *J Econ Entomol*. 2006; 99:164–175. [PubMed: 16573337]
- Arensburger P, Orsetti J, Kim Y-J, Aluvihare C, O'Brochta DA. A new active transposable element, *Herves*, from the African malaria mosquito *Anopheles gambiae*. *Genetics*. 2005; 169:697–708. [PubMed: 15545643]
- Behere GT, Tay WT, Russell DA, Heckel DG, Appleton BR, Kranthi KR, Batterham P. Mitochondrial DNA analysis of field populations of *Helicoverpa armigera* (Lepidoptera: Noctuidae) and of its relationship to *H. zea*. *BMC Evol Biol*. 2007; 7:117. [PubMed: 17629927]
- Biever KD, Thomas GD, Boldt PE. Effects of *Heliothis zea* (Lepidoptera: Noctuidae) on soybean yield and quality. *J Econ Entomol*. 1983; 76:762–765.
- Brandon, H. Cotton pests claim 1.17 million bales. 2004. http://southwestfarmpress.com/news/farming_cotton_pests_claim_3/
- Brown TM, Bryson PK, Brickle DS, Pimprale S, Arnette F, Roof ME, Walker JT, Sullivan MJ. Pyrethroid-resistant *Helicoverpa zea* and transgenic cotton in South Carolina. *Crop Protect*. 1998; 17:441–445.
- Burd AD, Gould F, Bradley JR, Van Duyn JW, Moar WJ. Estimated frequency of nonrecessive Bt Resistance genes in bollworm, *Helicoverpa zea* (Boddie) (Lepidoptera: Noctuidae) in eastern North Carolina. *J Econ Entomol*. 2003; 96:137–142. [PubMed: 12650356]
- Charlesworth B, Charlesworth D. The population dynamics of transposable elements. *Genet Res*. 1983; 42:1–27.
- Chen S, Li X. Transposable elements are enriched within or in close proximity to xenobiotic-metabolizing cytochrome P450 genes. *BMC Evol Biol*. 2007; 7:46. [PubMed: 17381843]
- Chen S, Li X. Molecular characterization of the first intact transib transposon from *Helicoverpa zea*. *Gene*. 2008; 408:51–63. [PubMed: 18031956]
- Feschotte C, Pritham EJ. DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet*. 2007; 41:331–368. [PubMed: 18076328]
- Fitt DP. The ecology of *Heliothis* species in relation to agroecosystems. *Annu Rev Entomol*. 1989; 34:17–53.
- Kapitonov VV, Jurka J. Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome. *Proc Natl Acad Sci USA*. 2003; 100:6569–6574. [PubMed: 12743378]
- Kapitonov VV, Jurka J. RAG1 Core and V (D) J Recombination signal sequences were derived from *Transib* transposons. *PLoS Biol*. 2005; 3:e181. [PubMed: 15898832]
- Kogan, JD.; Sell, K.; Stinner, RE.; Bradley, JR.; Kogan, M. INTSOY Series number 17. International Agricultural Publication; Urbana, Illinois: 1978. A bibliography of *Heliothis zea* (Boddie) and *H. virescens* (F.) (Lepidoptera: Noctuidae).
- Li X, Berenbaum MR, Schuler MA. Molecular cloning and expression of CYP6B8: a xanthotoxin-inducible cytochrome P450 cDNA from *Helicoverpa zea*. *Insect Biochem Mol Biol*. 2000; 30:75–84. pp. 3-20. [PubMed: 10646973]
- Li X, Schuler MA, Berenbaum MR. Jasmonate and salicylate induce expression of cytochrome P450 genes in insects. *Nature*. 2002a; 419:712–715. [PubMed: 12384696]
- Li X, Berenbaum MR, Schuler MA. Plant allelochemicals differentially regulate *Helicoverpa zea* cytochrome P450 genes. *Insect Mol Biol*. 2002b; 11:343–351. [PubMed: 12144700]
- Li X, Berenbaum MR, Schuler MA. Cytochrome P450 and actin genes expressed in *Helicoverpa zea* and *Helicoverpa armigera*: paralogy/orthology identification, gene conversion and evolution. *Insect Biochem Mol Biol*. 2002c; 32:311–320. [PubMed: 11804803]

- Li X, Baudry J, Schuler MA, Berenbaum MR. Structural and functional evolution of insect CYP6B proteins: from specialist to generalist P450. *Proc Natl Acad Sci USA*. 2004; 101:2939–2944. [PubMed: 14981232]
- Li X, Berenbaum MR, Schuler MA. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annu Rev Entomol*. 2007; 52:231–253. [PubMed: 16925478]
- Li, X.; Ni, X. Deciphering the plant–insect phenotypic arms race. In: Liu, T–X.; Kang, L., editors. *Recent Advances in Entomological Research: From Molecular Biology to Pest Management*. Higher Education Press; Beijing, China: 2009. in press
- Luttrell RG, Wan L, Knighten K. Variation in susceptibility of noctuid (Lepidoptera) larvae attacking cotton and soybean to purified endotoxin proteins and commercial formulations of *Bacillus thuringiensis*. *J Econ Entomol*. 1999; 92:23–32.
- Mallet J, Korman A, Heckel D, King P. Biochemical genetics of *Heliothis* and *Helicoverpa* (Lepidoptera: Noctuidae and evidence for a founder event in *Helicoverpa zea*. *Ann Entomol Soc Am*. 1993; 86:189–197.
- Mackay TFC. What prevents transposable elements from taking over the genome? a commentary on ‘A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster* by Elizabeth Montgomery, Brian Charlesworth and Charles H. Langley. *Genet Res*. 2007; 89:433–434. [PubMed: 18976532]
- McCaffery AR. Resistance to insecticides in Heliothine Lepidoptera: A Global View. *Phil Trans Biol Sci*. 1998; 353:1735–1750.
- McLeod PJ, Katayama RK, Sweeden MB, Burleigh JG, Giese WG, Burks RA. Insect pests and impact on late–season tomato in Arkansas. *J En Sci*. 1996; 31:152–165.
- Metcalf, RL.; Metcalf, RA. *Destructive and useful insects: Their habits and Control*. New York: McGraw–Hill, Inc; 1993.
- Messier T, O’Neill J, Hou S, Nicklas J, Finette B. In vivo transposition mediated by V (D) J recombinase in human T lymphocytes. *EMBO J*. 2003; 22:1381–1388. [PubMed: 12628930]
- Mitter C, Poole RW, Matthews M. Biosystematics of the *Heliothinae* (Lepidoptera: Noctuidae). *Ann Rev Entomol*. 1993; 38:207–225.
- Montgomery E, Charlesworth B, Langley CH. A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. *Genet Res*. 1987; 49:31–41. [PubMed: 3032743]
- Niu G, Wen Z, Rupasinghe SG, Zeng RS, Berenbaum MR, Schuler MA. Aflatoxin B1 Detoxification by CYP321A1 in *Helicoverpa zea*. *Arch. insect Biochem Phys*. 2008; 69:32–45.
- Ring DR, Benedict JH. Comparison of insect injury–cotton yield response functions and economic injury levels for *Helicoverpa zea* and *Heliothis virescens* (Lepidoptera: Noctuidae) in the lower Gulf Coast of Texas. *J Econ Entomol*. 1993; 86:1228–1235.
- Robinson, E. Insects take tenth of 2000 cotton crop. 2001.
http://southwestfarmpress.com/mag/farming_insects_tenth_cotton/
- Rupasinghe SG, Wen Z, Chiu T–L, Schuler MA. *Helicoverpa zea* CYP6B8 and CYP321A1: different molecular solutions to the problem of metabolizing plant toxins and insecticides. *Protein Eng Des Sel*. 2007; 20:615–624. [PubMed: 18065401]
- Sakano H, Hüppi K, Heinrich G, Tonegawa S. Sequences at the somatic recombination sites of immunoglobulin light–chain genes. *Nature*. 1979; 280:288–294. [PubMed: 111144]
- Sasabe M, Wen Z, Berenbaum MR, Schuler MA. Molecular analysis of CYP321A1, a novel cytochrome P450 involved in metabolism of plant allelochemicals (furanocoumarins) and insecticides (cypermethrin) in *Helicoverpa zea*. *Gene*. 2004; 388:163–175. [PubMed: 15315820]
- Subramanian RA, Arensburg P, Atkinson PW, O’Brochta DA. Transposable element dynamics of the *hAT* element *Herves* in the human malaria vector *Anopheles gambiae* s.s. *Genetics*. 2007; 176:2477–2487. [PubMed: 17603116]
- Tabashnik BE, Gassmann AJ, Crowder DV, Carrière Y. Insect resistance to Bt crops: evidence versus theory. *Nat Biotechnol*. 2008; 26:199–202. [PubMed: 18259177]
- Thompson C. New insights into V (D) J recombination and its role in the evolution of the immune system. *Immunity*. 1995; 3:531–539. [PubMed: 7584143]

- Waldbauer GP, Cohen RW, Friedman S. An improved procedure for laboratory rearing of the corn earworm, *Heliothis zea* (Lepidoptera:Noctuidae). Great Lakes Entomol. 1984; 17:113–118.
- Wen Z, Zeng R, Niu G, Berenbaum MR, Schuler MA. Ecological Significance of Induction of Broad-Substrate Cytochrome P450s by Natural and Synthetic Inducers in *Helicoverpa zea*. J Chem Ecol. 2009; 35:183–189. [PubMed: 19198946]
- Wright S, Le Q, Schoen D, Bureau T. Population dynamics of an Ac-like transposable element in self- and cross-pollinating Arabidopsis. Genetics. 2001; 158:1279–1288. [PubMed: 11454774]
- Zeng R, Wen Z, Niu G, Schuler MA, Berenbaum MR. Allelochemical Induction of Cytochrome P450 Monooxygenases and Amelioration of Xenobiotic Toxicity in *Helicoverpa zea*. J Chem Ecol. 2007; 33:449–461. [PubMed: 17216359]
- Zhang C, Luo X, Ni X, Zhang Y, Li X. Functional characterization of cis-acting elements mediating flavone-inducible expression of *CYP321A1*. Insect Biochem Mol Biol. 201010.1016/j.ibmb.2010.09.003

Figure 1.

PCR-based detection of *Hztansib*-like elements in three *Helicoverpa* species. Three types of genomic PCR using the primers *Hztansib*IR (A. Full-length PCR), flanking-F/flanking-R (B. ORF-flanking PCR) and transib-F/transib-R (C. ORF PCR) were conducted to attempt amplifications of the putative full-length and internally-deleted *Hztansib* copies as well as the ORF of *Hztansib*-like elements in *H. zea*, *H. armigera*, and *H. assulta*. The annealing positions of the 5 primers are indicated in Fig. 2. The white arrows indicate the putative intact (3518 bp for the full-length PCR, 1809 bp for the ORF-flanking and 1524 bp for ORF PCR) and internally-deleted copies of *Hztansib*. M = 1kb DNA ladder.

Figure 2.

Schematic presentation of transposon display and three types of genomic PCR for detection of *Hztransib* in the *Helicoverpa* genomes. A hypothetical full-length copy of *Hztransib* and part of its 5' flanking sequence surrounding the first *Msp I* restriction enzyme site (depicted as two short parallel lines) upstream of the copy are drawn to scale with its ORF (open reading frame) depicted as an open box of dot lines, 5' and 3' UTR (untranslated region) as open boxes, promoter and 5' flanking sequences as filled black boxes, and left and right TIR (terminal inverted repeat) as filled black arrowheads. This copy is flanked by its 5-bp (CGTCC) TSD (target site duplication) sequence. The 54-bp insertion sequence (depicted as an inverted triangle) present only in its left TIR is shown above the copy. The *Msp I* adapter that will be ligated to the fragment containing the *Hztransib* copy at the first *Msp I* restriction enzyme site upstream of it is shown as an open box (not drawn to scale) above the 5'-flanking sequence of the copy. The annealing positions of the *Msp I* adapter- and *Hztransib*-specific primers used for transposon display (MspI-a, Gsp1, Gsp2), full-length (HztransibIR), ORF-flanking (flanking-F, flanking-R) and ORF (transib-F, transib-R) PCR are indicated by arrows.

Figure 3.

Transposon display of *Hzttransib* in the 8 populations of *H. zea*. The 5'-flanking sequences of all *Hzttransib* insertions within the genome of each of the 128 larvae randomly picked from the 8 *H. zea* populations (16 larvae per population, $N=16 \times 8 = 128$) were PCR-amplified and revealed by transposon display. Shown here is a representative transposon display gel image of 2 individuals from each of the 8 populations. Each lane displays the *Hzttransib* elements harbored in a single individual larva. All the different bands (i.e. sites, 39 in total) detected from 8 populations are numbered 1–39 from the largest to the smallest in size. As an example, 5, 37, and 39 at the right of the gel refer to the names of the corresponding bands and insertion sites. M= 100 bp DNA Marker (Invitrogen, USA).

Figure 4.

Occupancy frequency of the 39 *Hztansib* insertion sites in *H. zea*. Data were derived from transposon display gels of the 128 *H. zea* larvae ($N = 16 \times 8 = 128$)

Figure 5.

Site-occupancy frequency distribution of *Hztansib* from the 8 populations (A-H) of *H. zea*. The times of occurrence in a sample is plotted on the X-axis, whereas the number of sites that occurred in a sample exactly “x” times is plotted on the Y-axis. Data were derived from transposon display gels of the 16 larvae of each population.

Figure 6.

Detection of the putative intact and internally-deleted copies of *Hztransib* in the 8 populations of *H. zea*. The full-length PCR with the primer *Hztransib*IR spanning the whole element (see Fig. 2 for its annealing positions) was conducted to attempt amplifications of the putative full-length (3518 bp) and internally-deleted (<3518 bp) copies of *Hztransib* present in a single individual larva. Shown here is a representative gel picture of 1 individual from each population. The arrows at the right of the gel picture indicate the putative intact and internally-deleted copies of *Hztransib*. M = 1kb DNA ladder.

Figure 7.

ORF-flanking and ORF PCR detection of *Hztansib* in the 8 populations of *H. zea*. A representative ORF-flanking PCR gel of 2 individuals from each population is shown in **A**. The ORF PCR gel of the same set of individuals from the 8 populations is shown in **B**. The arrows at the right of the two gel pictures indicate the intact (1809 bp for ORF-flanking PCR and 1524 bp for ORF PCR) and internally-deleted (smaller than 1809 bp and 1524 bp, respectively) copies of *Hztansib*. M = 1kb DNA ladder.

The eight populations of *H. zea*

Table 1

Name	Origin	Host/diet	Collection date	Generation	Frozen date
Field Populations					
Field corn	Lafayette, Indiana	Field corn	07-14-2006	F1	08-04-2006
Bt corn	Tifton, Georgia	Bt corn	07-03-2008	F2	11-05-2008
Bt cotton	Tifton, Georgia	Bt cotton	07-30-2008	F2	11-05-2008
DK 6971	Tifton, Georgia	Bt corn DK6971	10-01-2008	F1	11-05-2008
Lab Populations					
Benzon	Benzon Research	Diet	-	-	11-09-2008
USDA	ARS, Georgia	Diet	-	-	11-09-2008
Pop 2	UA	Diet	-	-	08-02-2006
NCSU	NCSU lab colony	Diet	-	-	11-20-2008

Table 2
Insertion site, copy number and site occupancy of *Hztransib* in the eight populations of *H. zea*

Population	Sample size	Total insertion site	Unique sites	Sites Occupancy*	Insertion sites per individual			n#
					Min	Max	Mean ± SE§	
Field Corn	16	18	8	4.22	2	7	4.7±0.3b	9.77
Benzon	16	14	2	7.43	2	11	6.5±0.6a	5.83
Bt cotton	16	17	3	5.76	4	9	6.13±0.36ab	7.77
DK6971	16	19	5	5.53	2	11	6.56±0.83 a	8.78
Bt corn	16	14	4	5.21	1	8	4.56±0.42 b	6.75
USDA	16	14	4	5.14	2	7	4.50±0.38 b	6.75
Pop 2	16	24	9	3.25	3	9	4.87±0.48 b	14.21
NCSU	16	21	8	3.48	1	10	4.56±0.52 b	12.01

* Sites occupancy = the sum of the occurrence times of all the sites of a population/the total insertion sites of a population.

§ Different letters indicate significant difference at P <0.05 (Turkey's HSD test, P <0.05).

The equilibrium copy number per individual (*n*) was estimated by the formula
$$\hat{n} = \sum_{i=1}^m (1 - \sqrt{1 - Z_i})$$
, where *Zi* is equal to the frequency of the 'null' allele for the *i*th insertion site (Wright *et al.*, 2001).

Table 3

Relative frequencies of the putative intact and internally-deleted *Hztransib* copies in the eight populations of *H. zea*

Population	Sample size	Frequency of full-length copy (%)	Frequency of Complete ORF (%)	Frequency of deleted copies (%)
Field corn	16	100	100	16.8
Benzon	16	100	100	100.0
Bt cotton	16	100	100	9.0
DK 6971	16	100	100	5.0
Bt corn	16	100	100	0.0
USDA	16	100	100	100.0
Pop 2	16	100	100	13.5
NCSU	16	100	100	93.0