

## ORIGINAL ARTICLE

# Fine mapping of $E^{kp-1}$ , a locus associated with silkworm (*Bombyx mori*) proleg development

H Xiang<sup>1</sup>, M Li<sup>1,2</sup>, F Yang<sup>1</sup>, Q Guo<sup>1</sup>, S Zhan<sup>1</sup>, H Lin<sup>1</sup>, X Miao<sup>1</sup> and Y Huang<sup>1</sup>

<sup>1</sup>Shanghai Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Center of Insect Science, The Chinese Academy of Sciences, Shanghai, P.R. China and <sup>2</sup>Sericultural Research Institute, Chinese Academy of Agricultural Sciences, Zhenjiang, P.R. China

The silkworm homeotic mutant  $E^{kp}$  has a pair of rudimentary abdominal legs, called prolegs, in its A2 segment. This phenotype is caused by a single dominant mutation at the  $E^{kp-1}$  locus, which was previously mapped to chromosome 6. To explore the possible association of Hox genes with proleg development in the silkworm, a map-based cloning strategy was used to isolate the  $E^{kp-1}$  locus. Five  $E^{kp-1}$ -linked simple sequence repeat markers on chromosome 6 were used to generate a low-resolution map with a total genetic distance of 39.5 cM. Four additional cleaved amplified polymorphic sequence markers were developed based on the initial map. The closest marker to  $E^{kp-1}$  was at a genetic distance of 2.7 cM. A high-resolution genetic map was constructed using nine BC1 segregating

populations consisting of 2396 individuals. Recombination suppression was observed in the vicinity of  $E^{kp-1}$ . Four molecular markers were tightly linked to  $E^{kp-1}$ , and three were clustered with it. These markers were used to screen a BAC library. A single bacterial artificial chromosome (BAC) clone spanning the  $E^{kp-1}$  locus was identified, and  $E^{kp-1}$  was delimited to a region less than 220 kb long that included the Hox gene *abdominal-A* and a non-coding locus, *iab-4*. These results provide essential information for the isolation of this locus, which may shed light on the mechanism of proleg development in the silkworm and possibly in Lepidoptera.

Heredity (2008) 100, 533–540; doi:10.1038/hdy.2008.10; published online 26 March 2008

**Keywords:**  $E^{kp}$ ; proleg; simple sequence repeat (SSR); cleaved amplified polymorphic sequence (CAPS); BAC *abd-A*

## Introduction

Insects show a wide range of anatomical diversity in appendage morphology. The larvae of many holometabolous insects possess abdominal appendages called prolegs (Nagy and Grbic, 1999). There is considerable diversity in proleg distribution on the insect body, including wide variation in segmental arrangement and number. For example, hymenopteran species have prolegs on each abdominal segment, whereas the prolegs of some dipteran species may be a useful adaptation for certain larval habits or predatory needs (Suzuki and Palopoli, 2001). Lepidopteran larvae have prolegs in segments A3–A6. The mechanism underlying proleg development is unclear. The body segments of many insects are specified by Hox, or homeotic, genes (Ouweneel, 1976). Changes in the functions of these genes and/or in the downstream genes under their control may result in these morphological differences.

In the silkworm *Bombyx mori*, a homeotic gene complex known as the E complex specifies the identities of the body segments. The E complex genes are part of the silkworm Hox gene cluster, which is located on

chromosome 6 (Yasukochi *et al.*, 2004). These genes include Bm *Ultrabithorax* (*Ubx*), Bm *abdominal-A* (*abd-A*) and Bm *Abdominal-B* (*Abd-B*; Ueno *et al.*, 1992), which are homologs of the *Drosophila* bithorax complex (BX-C; Bender *et al.*, 1983). In *Drosophila*, these Hox genes, especially *abd-A*, are important for determinative decisions in the anterior abdomen (Karch *et al.*, 1985). The Hox protein BX-C represses the development of ventral limbs by preventing the transcription of *Distal-less* (*Dll*), a limb-promoting gene. In Lepidoptera, it is thought that E complex genes also repress *Dll* since in the abdominal segments, circular holes lacking Hox protein expression showed *Dll* expression and proleg development (Warren *et al.*, 1994). *Dll* seemed to escape from repression by the Hox genes in these special holes. What produces this pattern of Hox gene expression in the abdominal segments of Lepidoptera? Do changes in the regulation of the Hox genes or their downstream targets play a role in the occurrence of prolegs? The answers to these questions are largely unknown (Averof, 2002; Hughes and Kaufman, 2002).

In the silkworm, more than 30 types of homeotic mutants, called E mutants, have been identified by phenotype and analyzed for pseudoallelism. Most of the mutations in these mutants are dominant and induce the expression of extra markings or supernumerary legs/prolegs (that is, rudimentary abdominal legs; Itikawa, 1943; Tazima, 1964). Until recently, the molecular nature of the E mutants was unknown, except for the mutants  $E^N$  and  $E^{Ca}$ . Ueno *et al.* (1992) reported that deletions

Correspondence: Professor Y Huang, Shanghai Institute of Plant Physiology and Ecology, The Chinese Academy of Sciences, 300 Fenglin Road, Shanghai 200032, P.R. China.

E-mail: yphuang@sibs.ac.cn

Received 19 June 2007; revised 29 December 2007; accepted 2 January 2008; published online 26 March 2008

within Bm *Ubx* may be responsible for the phenotypes of these two mutants and that the function of Bm *Ubx* might be similar to that of Dm *Ubx* in repressing leg and proleg development. Although the mechanism of proleg development remains unknown, changes in Hox gene function (Warren *et al.*, 1994) or in the expression of genes downstream of the Hox genes (Panganiban *et al.*, 1994), including novel target genes (Ueno and Suzuki, 1997), have been implicated. The analysis of mutations will help to explain how segmental developmental mechanisms may have been modified during evolution.

The *E<sup>kp</sup>* mutant is an ideal subject for such studies. Larvae homozygous for *E<sup>kp</sup>* (*E<sup>kp</sup>/E<sup>kp</sup>*) exhibit various shifts on their dorsal and ventral sides (Figure 1), including extra crescents on the A3 segment (anterior transformation) and small supernumerary prolegs on the A2 segment (posterior transformation). In contrast, larvae heterozygous for *E<sup>kp</sup>* (*E<sup>kp</sup>/+*) show only the A2 segment phenotype. Regardless of the crescent phenotype of the *E<sup>kp</sup>* mutant, its supernumerary prolegs can be considered a typical phenotype for studying the mechanism of proleg development. We have named this phenotype *E<sup>kp</sup>-1*. The various phenotypes of the *E<sup>kp</sup>* mutant are thought to be associated with mutations in E complex loci (Goldsmith, 1995), but experimental evidence is still lacking.

In this study, we used simple sequence repeat (SSR) and cleaved amplified polymorphic sequence (CAPS) markers to restrict *E<sup>kp</sup>-1* to within a 220-kb region in the Hox gene cluster, thereby providing the starting point for molecular analysis of the locus, which may play a role in silkworm proleg development. This is the first study to finely map a phenotypic locus in the silkworm. Our results demonstrate that it may be feasible to locate new genes in the silkworm by map-based cloning.

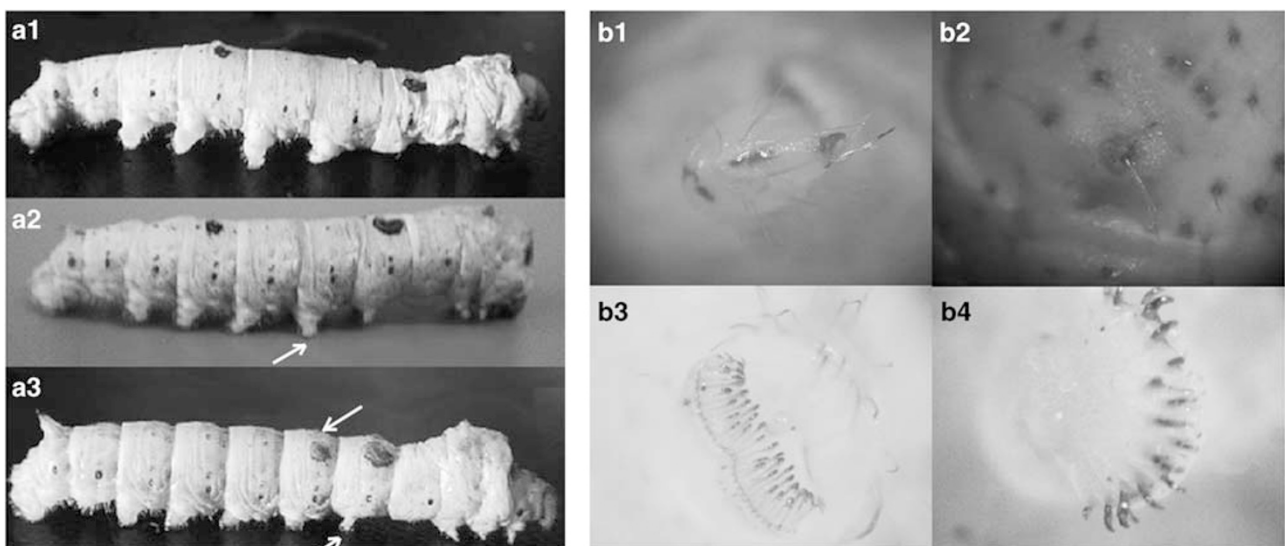
## Materials and methods

### Silkworm material and mapping populations

*E<sup>kp</sup>* (*E<sup>kp</sup>/E<sup>kp</sup>*) is maintained by single-pair mating in a stable silkworm strain at the Chinese Academy of Agriculture Sciences' Sericulture Research Institute. *E<sup>kp</sup>* and the wild-type strain p50 are the parental insects. *E<sup>kp</sup>-1* was previously placed within the sixth chromosomal linkage group based on SSR linkage analysis (Miao *et al.*, 2005). To map the *E<sup>kp</sup>-1* locus on a fine scale, we used nine male F1 individuals from a single-pair cross of an *E<sup>kp</sup>* female with a p50 male and mated them with a p50 female to generate backcrossed F1 (BC1♂) progeny (designated p50 × (*E<sup>kp</sup>* × p50)); thus, a total of 2396 BC1♂ offspring were constructed. The larvae were reared at 25 °C until the fifth instar.

### Genetic mapping of *E<sup>kp</sup>-1*

Total DNA was extracted from the hind parts of 2396 BC1♂ larvae according to the method of Li *et al.* (2006) with the modification that we added SDS to the extraction buffer at a final concentration of 2%. We used the p50 and *E<sup>kp</sup>* parents to test for polymorphisms at each of the SSR markers developed by Miao *et al.* (2005). In total, 192 BC1♂ backcross progeny were then genotyped using the polymorphic SSR markers. SSR polymorphism determination and genotyping were performed using an Applied Biosystems 377 DNA sequencer (ABI PRISM) for 2 h with 96 lanes run per gel. Next, published sequence-tagged site (STS) markers that map to the Hox cluster on chromosome 6 (Yasukochi *et al.*, 2004) were amplified using p50 DNA as the template and sequenced to develop CAPS markers for the BC1 population. BLAST searches using the sequences of the STS markers



**Figure 1** Phenotype and appendage morphology of p50 and the *E<sup>kp</sup>* mutant. (a1) *E<sup>+</sup>/E<sup>+</sup>* (p50 strain). There is one crescent on the A2 segment and another on the A5 segment. The A2 segment lacks ventral appendages. (a2) *E<sup>+</sup>/E<sup>kp</sup>* (F1 progeny of p50 and *E<sup>kp</sup>*). The dorsal morphology is the same as that of the *E<sup>+</sup>/E<sup>+</sup>*, except the *E<sup>+</sup>/E<sup>kp</sup>* individual has small supernumerary prolegs on the dorsal side of its A2 segment (that is, the *E<sup>kp</sup>/E<sup>kp</sup>* phenotype). The arrow shows the supernumerary prolegs. (a3) *E<sup>kp</sup>/E<sup>kp</sup>* (*E<sup>kp</sup>* strain). The mutant has extra crescents on its A3 segment and small supernumerary prolegs on its A2 segment; on the dorsal side, it is missing a crescent on the A5 segment. The arrows show the extra crescents and supernumerary prolegs. (b1) Thoracic leg. (b2) Ventral side of a normal A2 segment. (b3) Abdominal leg (that is, proleg). (b4) Supernumerary proleg, which is similar to a proleg but distinct from a thoracic leg.

and silkworm SSR sequences were performed to identify linkages between the two sets of markers. Using this approach, we developed a CAPS marker, B422251, within Scaffold000028 that bears both an SSR marker (S0608) and a published STS marker (B4C5Ga; Yasukochi *et al.*, 2004). We also developed three polymorphic STS markers, B3110Gb, B9B9Ba (Yasukochi *et al.*, 2004) and ab17577 (an *Abd-B*-derived marker in Scaffold003414) on either side of B422251. Genotyping of the same BC1 ♂ population was conducted using these markers.

To determine the relationship between  $E^{kp-1}$  and the molecular markers linked to it, MAPMAKER (version 3.0) was used to calculate genetic map distances (Li and Garvin, 2003). Two-point linkage was determined at log of odds (LOD) = 5.0.

On the basis of our low-resolution mapping results (see Results), a total of 2396 BC1 ♂ larvae were genotyped using two closely linked markers, ab17577 and B422251, to produce a fine map of the  $E^{kp-1}$  locus.

### Candidate gene cloning

The epidermis and fat body were dissected from fifth instar p50 and  $E^{kp}$  larvae in sterile phosphate-buffered saline (50 mM sodium phosphate, pH 7.5 and 150 mM NaCl), frozen rapidly in liquid nitrogen and stored at  $-70^{\circ}\text{C}$  until use. Total RNA was extracted using TRIzol (Invitrogen, Salt Lake City, Utah, USA) and digested with DNase I (Takara, Tokyo, Japan) to remove any remaining DNA. cDNA was synthesized using a first-strand cDNA synthesis kit (Toyobo, Osaka, Japan).

To clone the complete coding region of *Ubx*, we designed the primer *Ubx1f* (5'-ATGAACCTTACTTC GAC-3') based on a conserved motif in the translation initiation region. Similarly, the primer *Ubx3'ORFr* (5'-TTAATGTTCTGGGGTGTCCCTG-3') was designed based on a motif close to the terminal coding region that included the predicted ending code. cDNA from p50 and  $E^{kp}$  was used as the template for PCR using the primers *Ubx1f* and *Ubx3'r*.

To clone full-length *abd-A*, we screened a silkworm gonad cDNA library constructed in our laboratory (Liu WB *et al.*, unpublished) by PCR using the primers *abd-Aboxf* (5'-GCCAGGACGACAACTTA-3') and *abd-Aboxr* (5'-TGCTGCTGCTTCATTCTG-3'). Positive clone pools were subsequently amplified using either universal primer M13 forward with *abd-Aboxf* or universal primer M13 reverse with *abd-Aboxr*.

The PCR products were cloned into pMD18-T (Takara) and sequenced using an ABI 3700 DNA Analyzer (ABI PRISM). The sequences were then BLAST-searched against the silkworm genome database. The matching scaffolds were then used to develop several CAPS markers.

### Marker development

For high-resolution mapping, seven CAPS markers were developed based on published data (Yasukochi *et al.*, 2004), and sequence information was inferred from the candidate genes. These CAPS markers include *Ubx3'A*, which is located about 29 kb downstream of the *Ubx* 3'ORF, designed from Scaffold000017; B12J2Aa, which is in the Hox cluster (Yasukochi *et al.*, 2004); B910325 and B923511, which are in Scaffold005660 (B910325 is about 12 kb from the published STS marker B9B9Ba (Yasukochi

*et al.*, 2004), while B923511 is actually a B9B9Ba-derived CAPS marker); *abd-A5'*, which is about 9 kb upstream of the *abd-A* 5' open reading frame (ORF), designed from Scaffold000405 and *abd-A3'A* and *abd-A3'B*, which are about 9 and 19 kb downstream of the *abd-A* 3' ORF, respectively, designed from Scaffold000075.

### Construction of BAC pools

The bacterial artificial chromosome (BAC) library was kindly provided by Kazuei Mita and Chengang Wu (Wu *et al.*, 1999). The library has 36 864 clones, which are in 96 384-well microtiter plates. For each plate, we constructed four pools, labeled A, B, C and D, each with 96 BAC clones. In total, 384 pools were constructed. The BAC DNA from each pool was isolated using a conventional alkali method (Sambrook *et al.*, 1989), diluted 200-fold and then put into the corresponding well in one of four 96-well microtiter plates marked pool-A, pool-B, pool-C and pool-D.

### PCR-based BAC library screening and construction of a physical map

The markers used for fine mapping plus additional monomorphic STS markers, including the *Ubx*-derived markers *Ubx3'B* and *Ubx5'*, were used to screen the BAC library. We discovered that the predicted non-coding locus *iab-4* (Tong *et al.*, 2006) was located in Scaffold005660 (where B9B9Ba is located); thus, *iab-4*-derived markers were designed and used in the screening. All markers developed in this study are listed in Table 1.

The first round of PCR was performed using the CAPS markers known to be closely linked to  $E^{kp-1}$  with the four pools containing all BAC clones in the library. For each positive pool, we constructed a row pool and a column pool. The BAC DNA of each pool was isolated using the method described above. A second round of PCR was performed on the row and column pools using the corresponding markers to confirm the positive BAC clones. To prepare the DNA of the positive clones for further analysis, single colonies were cultured for 24 h in 20 ml of Luria-Bertani containing  $12.5\ \mu\text{g ml}^{-1}$  chloramphenicol. The BAC DNA was then prepared as described by Yin *et al.* (2006). The positive BAC DNA was digested with *NotI* (NEB, Ipswich, MA, USA) at  $37^{\circ}\text{C}$  for 3 h and analyzed by pulsed-field gel electrophoresis in 1.0% agarose gel with  $0.5 \times$  tris-borate-EDTA buffer under the following conditions:  $6.0\ \text{V cm}^{-1}$ ,  $14^{\circ}\text{C}$ , pump setting 80, switch time 1–12 s,  $120^{\circ}$  for 16 h. The end sequences of the positive BAC clones were determined using the primers T7 and Sp6. One BAC-end derived STS marker was designed based on the sequence (Table 1).

## Results

### Low-resolution mapping of the $E^{kp-1}$ locus on chromosome 6

Among the 2396 BC1 individuals, 1230 had super-numerary prolegs and 1166 were normal, giving a segregation ratio of 1:1. The  $\chi^2$  analysis confirmed segregation from a single gene ( $\chi^2 = 0.45$ ), proving that the  $E^{kp-1}$  mutation affects a single locus.  $E^{kp-1}$  was previously mapped to the sixth chromosome using BC1 ♀ offspring (Miao *et al.*, 2005). Four SSR markers on chromosome 6 (S0601, S0603, S0606 and S0609) were

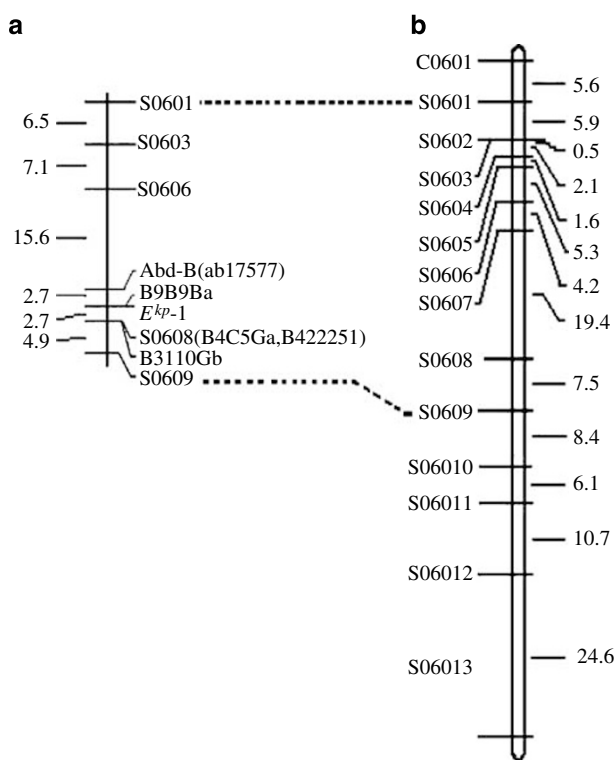
**Table 1** Markers used in this study

Name	Forward primer	Reverse primer	Size <sup>a</sup> (bp)	Polymorphism <sup>b</sup> (restriction enzyme)
<i>Published STS marker-derived CAPS markers</i>				
B422251	TCCCTGTCGCTCTGTTGA	GCATTCCTCTGGTCGGTC	927	P ( <i>Hinf</i> I)
B923511	ACTTCAACCTGCTTATCGG	GCTGTAATCGCTTGGAGA	783	P ( <i>Ssp</i> I)
B910325	AACCTGGCTCAATACTCG	CGTGCGTTGTGATTTGTA	1032	P ( <i>Dra</i> I)
<i>Hox-gene-derived CAPS and STS markers</i>				
ab17577	GCTCCGTTGTCAITTTGTT	GTGTGGCGCTCACTTTAT	995	P
abd-A5'	GTCGGTAGGTTTGTAGTTGA	TCTCGTCTGTCTTATGTT	895	P ( <i>Ssp</i> I)
abd-A3'A	CTCCTCTGACGAGCCAAGC	TGCCGATGTCTGCCACGA	949	P ( <i>Fok</i> I)
abd-A3'B	AAGCAACTTGACTAAGCGAATG	CTGGATGGAATGCCGATG	1172	P ( <i>Fok</i> I)
Ubx5'	ATGACATAAGTGTTCCTCG	AGTTGTGATCGAGATAATT	547	M
Ubx3'A	GCTATTCTCCACGATTIG	TCCTTGTTGGGTCACCTGT	897	P ( <i>Hinf</i> I)
Ubx3'B	GGTCGGAATGCTCAAAGG	ACTACGGTACACCTACGCACTA	789	M
iab-4	TGCGTTCCGTGTTCTGT	GATGTCCTAACTTATGCCTTCT	1114	M
<i>BAC-derived STS marker</i>				
083DE9end	ATTGTCATCTTTTCTGAG	GTATTCATAATCTCACTTGC	896	M

Abbreviations: CAPS, cleaved amplified polymorphic sequence; STS, sequence-tagged site; BAC, bacterial artificial chromosome.

<sup>a</sup>Sizes in p50 are listed above.

<sup>b</sup>M, monomorphic; P, polymorphic.



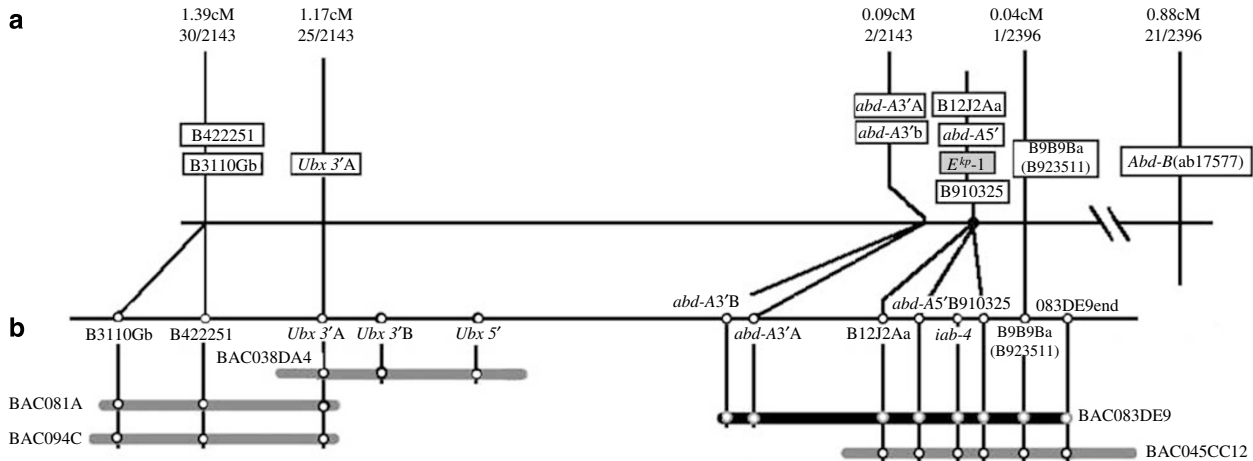
**Figure 2** Low-resolution mapping of the *E<sup>kp</sup>-1* locus. The genetic distance between the markers S0601 and S0609 in the p50 × (*E<sup>kp</sup>* × p50) population was 39.5 cM (a), which is shorter than but comparable to that in the C108 × (p50 × C108) population (46.5 cM; (b)) (Miao *et al.*, 2005).

polymorphic between p50 and *E<sup>kp</sup>*. We found that S0608 colocalized with the B4C5Ga-derived polymorphic CAPS marker B422251; thus, S0608 was also mapped. The recombination data are shown in Figure 2a. The order of the SSR markers in the linkage group was S0601, S0603, S0606, S0608 and S0609, which is consistent with previously published data (Miao *et al.*, 2005). Four STS

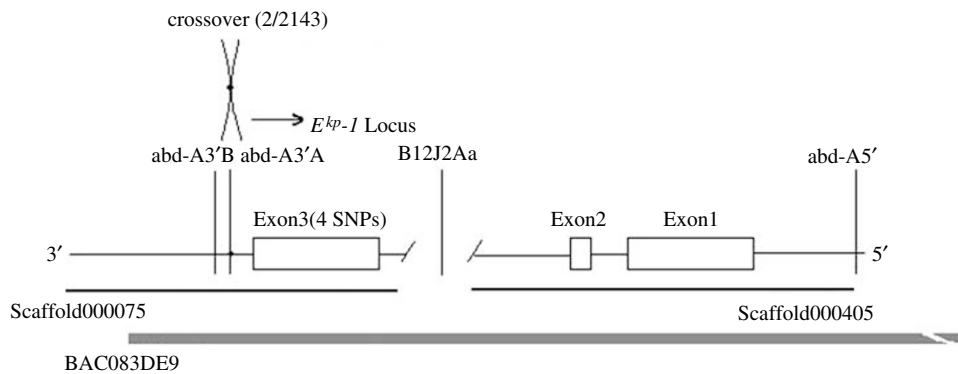
markers, *Abd-B* (represented by its derived marker ab17577), B9B9Ba, B4C5Ga (represented by its derived marker B422251) and B3110Gb were also mapped. The linkage map was 39.5 cM in length (Figure 2a), which is shorter than that determined for the C108 × (p50 × C108) population (46.5 cM, Figure 2b; Miao *et al.*, 2005). *E<sup>kp</sup>-1* mapped at 29.2 cM, between the SSR markers S0606 and S0609. The genetic distance between these two markers was 25.9 cM, which is shorter than the distance found using p50- and C108-derived BC1♂ progeny (32.2 cM; Miao *et al.*, 2005). *Abd-B* (ab17577) mapped at a distance of 2.7 cM from *E<sup>kp</sup>-1*. B4C5Ga (B422251) and B3110Gb were clustered on the other side, 2.7 cM away. B9B9Ba was clustered with the *E<sup>kp</sup>-1* locus.

#### High-resolution genetic mapping of the *E<sup>kp</sup>-1* locus

In the low-resolution genetic map, B422251 and B3110Gb were clustered (Figure 2a). As the corresponding physical map indicated that B422251 might be 1 Mb nearer to the *E<sup>kp</sup>-1* locus than B3110Gb (Yasukochi *et al.*, 2004), we chose B422251 and ab17577 for fine mapping of the *E<sup>kp</sup>-1* locus. We identified 21 recombinants for ab17577 and 30 for B422251, placing *E<sup>kp</sup>-1* between the two markers (Figure 3a). These recombinants were used to map our previous and newly developed CAPS markers (B3110Gb, *Ubx3'A*, *abd-A3'A*, *abd-A3'B*, *abd-A5'*, B12JaAa, B910325 and B923511). The segregation data were then used to generate a high-resolution genetic map. Markers B3110Gb, B422251, *Ubx3'A*, *abd-A3'A* and *abd-A3'B* were mapped to one side of the *E<sup>kp</sup>-1* locus. B3110Gb and B422251 were clustered, as were *abd-A3'A* and *abd-A3'B*. *abd-A3'A* and *abd-A3'B* were the closest to *E<sup>kp</sup>-1*, while B3110Gb and B422251 were the farthest from *E<sup>kp</sup>-1*. The other two markers, ab17577 and B923511, flanked *E<sup>kp</sup>-1* on the other side (Figure 3a). B9B9Ba (represented by its derived marker B923511), which was initially clustered with *E<sup>kp</sup>-1*, was spaced farther away in the high-resolution genetic map. The markers B910325, *abd-A5'* and B12J2Aa, which were previously mapped to the Hox cluster (Yasukochi *et al.*, 2004), were clustered



**Figure 3** Genetic map and BAC contig of the *E<sup>kp</sup>-1* locus. (a) High-resolution genetic map of the *E<sup>kp</sup>-1* region. Genetic distances are given in centiMorgan, and the number of recombinants/number of segregants tested is indicated. Vertical lines denote the position of each marker or gene. The black dot indicates the genetic position of the *E<sup>kp</sup>-1* locus. (b) BAC contig spanning the *E<sup>kp</sup>-1* locus. Dotted vertical lines mark the relative positions of the corresponding markers. The *E<sup>kp</sup>-1*-containing BAC insert (BAC083DE9) is highlighted in black.



**Figure 4** Structure of the Bm *abd-A* gene. Open boxes indicate the exons. The polymorphic markers are indicated on top of the vertical lines. Scaffold000405 contains exons 1 and 2, and scaffold000075 contains exon 3. There are four single nucleotide polymorphisms in exon 3. The grey bar indicates BAC083DE9, which contains the Bm *abd-A* gene.

with *E<sup>kp</sup>-1*. These markers spanned a total distance of 2.28 cM, with *E<sup>kp</sup>-1* in the middle, 0.88 cM from ab17577 and 1.39 cM from G422251; in comparison, a 5.4-cM interval was previously estimated for the same markers in the low-resolution genetic map. The order and distance of some of the markers in the high-resolution genetic map were different from those of the physical map of the Hox gene cluster (Yasukochi *et al.*, 2004); this was especially true for B3110Gb, B4C5Ga, *Ubx* and *abd-A* (*abd-A3'A* and *abd-A3'B*).

#### PCR-based BAC assembly to narrow the *E<sup>kp</sup>-1* region to a single BAC clone

To develop a BAC contig encompassing the *E<sup>kp</sup>-1* locus, we screened the BAC library using B9B9Ba and B4C5Ga. In total, 11 BAC pools were positive for B9B9Ba, and 20 were positive for B4C5Ga. No pools included both B9B9Ba and B4C5Ga. These positive pools were confirmed using all other markers that were used for fine genetic mapping (Figure 3b). Among the 20 BAC pools that were positive for B4C5Ga, 16 were positive for B3110Gb, among which two were also positive for

*Ubx3'A*; none was positive for *abd-A5'*, *abd-A3'A* or *abd-A3'B*.

Two B9B9Ba-positive BAC pools and their corresponding clones, 083DE9 and 045CC12, were also positive for the markers closely linked to and clustered with the *E<sup>kp</sup>-1* locus (Figure 3b). In particular, 083DE9 was positive for *abd-A3'A* and *abd-A3'B*. The end sequence of this clone, determined using the primer Sp6, indicated that the left end of the BAC is located 26 kb downstream of the *abd-A* 3' ORF within Scaffold000075; thus, 083DE9 harbors the region corresponding to *E<sup>kp</sup>-1*. The insert size, which was determined by CHEF electrophoresis after digestion with *NotI*, was about 220 kb. The BAC-end-derived marker 083DE9end, which was developed from the BAC clone BAC083DE9, was used to confirm the overlap between 083DE9 and 045CC12 (Figure 3b).

#### Candidate gene analysis

Among the silkworm Hox genes, only Bm *abd-A* is located in the interval between the markers flanking the *E<sup>kp</sup>-1* locus (Figure 4). We detected T-C substitutions at four sites within the *abd-A* cDNA from *E<sup>kp</sup>* compared to

that from p50, but they did not cause amino-acid changes and thus were considered to be single nucleotide polymorphisms. The Bm *abd-A* ORF contains three exons (Figure 4). The four single nucleotide polymorphisms were all in exon 3. Two recombination events were detected about 9 kb downstream, indicating that the *E<sup>kp</sup>-1* locus is in the upstream direction. No recombinants were detected in intron 2 or the 5' flanking region. The coding sequences of *Ubx* and *abd-A* from the p50 and *E<sup>kp</sup>* strains have been deposited in GenBank (accession numbers EU365396–365399).

#### Bm *iab-4* is clustered with the *E<sup>kp</sup>-1* locus

In the *E<sup>kp</sup>-1* region, we also found a predicted non-coding locus, Bm *iab-4*. Since no polymorphic markers were developed from *iab-4*, we were unable to map this locus to the genetic map. On the basis of the physical map, however, the locus is between the markers *abd-A*5' and B910325, which are clustered with the *E<sup>kp</sup>-1* locus. The *iab-4* locus was also found to be clustered with the *E<sup>kp</sup>-1* locus. The sequence of the fragment containing putative *iab-4* was deposited in GenBank under accession number EU365400.

## Discussion

The location of a gene on a genetic linkage map provides a starting point for map-based cloning. Advances in the silkworm DNA molecular map and genomic sequencing (Xia *et al.*, 2004; Miao *et al.*, 2005) have made it possible to identify the chromosomal location of a gene and to quickly screen a large number of molecular markers and identify those closely linked to the target gene for further fine mapping.

The *E<sup>kp</sup>-1* locus was previously placed on the sixth chromosome (Miao *et al.*, 2005). In the silkworm, owing to a lack of crossing over in females, BC1 progeny are usually used for linkage analysis (Miao *et al.*, 2005; Li *et al.*, 2006). The genetic distance of the linkage map constructed here was smaller than that constructed by Miao *et al.* (2005), which affected the genetic distances of the SSR markers flanking the *E<sup>kp</sup>-1* locus. This indicates a difference in the number of recombination events between the two pairs of parents (Ganal and Tanksley, 1996; Chetelat *et al.*, 2000). The colocalization of S0608 and the STS marker B4C5Ga allowed us to link the map of *E<sup>kp</sup>-1* with the physical map of the Hox gene cluster (Yasukochi *et al.*, 2004); thus, we were able to finely map the *E<sup>kp</sup>-1* locus in a relatively short time. With regard to the physical map produced by Yasukochi *et al.* (2004), *Ubx* and *abd-A* were the only Hox genes in the region, so they were the first candidate genes to be considered. Cloning and analysis of these genes allowed us to create new closely linked markers for fine mapping.

A large-insert silkworm BAC library, derived from the p50 and C108 strains, has been established that can be used for physical mapping of a silkworm gene or locus of interest (Wu *et al.*, 1999). PCR analysis confirmed that the marker B4C5Ga, located downstream of the *Ubx* 3' UTR, was close to the marker B3110Gb. This is quite different from the previously published position of B4C5Ga, which was mapped between B12J2Aa and B9B9Ba in the Hox gene cluster (Yasukochi *et al.*, 2004).

The physical and genetic mapping data for *E<sup>kp</sup>-1* suggest that the region shows reduced recombination.

The genetic distance between *abd-A*3'A and B923511, the innermost loci that flank the target, was 0.13 cM; however, the physical size of this region is estimated to be less than 220 kb. Therefore, the average physical/genetic (P/G) distance ratio of the target interval is about 1.24 Mb cM<sup>-1</sup>. This is much higher than the estimated average ratio for the entire silkworm genome (about 0.25 Mb cM<sup>-1</sup>; Goldsmith *et al.*, 2005). This phenomenon has also been reported in plant studies. For example, the rice blast resistance locus *Pib* (Wang *et al.*, 1999) showed a ratio of 1.34 Mb cM<sup>-1</sup>, while in the rice lesion mimic gene *Spl11*, the ratio was even higher (2.46 Mb cM<sup>-1</sup>; Zeng *et al.*, 2002). In plants, the offspring used for mapping, which may include the F2, F6 or even F7 generation, might be one reason for the apparent reduction in recombination (Zeng *et al.*, 2002); however, that does not apply in this study. Recombination suppression can also be caused by a high degree of sequence divergence between the two parents (Ganal and Tanksley, 1996; Chetelat *et al.*, 2000). p50 and *E<sup>kp</sup>* are two strains of the species *B. mori* and thus can mate normally, so there is no high degree of sequence divergence between them. Another reason for recombination suppression is the proximity of the locus to the centromere (van Daelen *et al.*, 1993; Haupt *et al.*, 2001). Some insects, such as the Hemiptera and Lepidoptera, which include the silkworm, are believed to have diffuse centromeres (Murakami and Imai, 1974; Pimpinelli and Goday, 1989), resulting in holocentric chromosomes. Chromosomes with full-length kinetochores are not sensitive to breakage (Okazaki *et al.*, 1993); thus, recombination is not easy to occur. Whether *E<sup>kp</sup>-1* is near a kinetochore region remains unknown. Fluorescence *in situ* hybridization using marker-specific probes as well as heterochromatic centromere probes may be used to test this possibility. Moreover, fine mapping of the genomic sequence in the *E<sup>kp</sup>-1* region will provide a powerful tool with which to predict the possible heterochromatic zones therein.

Our physical mapping data pinpointed *E<sup>kp</sup>-1* to a region less than 220 kb in length and containing an unknown number of genes. However, there is at least one gene, *abd-A*, and a non-coding locus, *iab-4*, for microRNA (Tong *et al.*, 2006) in this region. By comparing the ORF of the *abd-A* gene from *E<sup>kp</sup>* to that from p50, we found mutations in exon 3, but they were synonymous. Thus, these mutations cannot be the cause of the mutant phenotype. Moreover, because we did not explore the UTR region, the promoter, or other upstream regulatory elements, *abd-A* cannot be excluded as a candidate. In *Drosophila*, *abd-A* represses limb development in the abdomen (Vachon *et al.*, 1992) by preventing expression of the limb-promoting gene *Dll*. However, in Lepidoptera the effect of *abd-A* on *Dll* is unclear. In the larvae of the butterfly *Junonia coenia*, *abd-A* is absent from areas of *Dll* accumulation in abdominal segments A3–A6 (that is, where larval prolegs form). The otherwise uniform domain of Hox gene expression in this species shows 'holes' where *Dll* accumulates (Warren *et al.*, 1994). These holes imply that the *Ubx/abd-A* proteins in Lepidoptera have repressive effects on *Dll* (Warren *et al.*, 1994; Suzuki and Palopoli, 2001). Zheng *et al.* (1999) have addressed the possible difference between *Ubx* and *abd-A* in regulating *Dll*. In the silkworm, the *E<sup>Ca</sup>/E<sup>Ca</sup>* mutant embryo, which has a deletion near the homeobox region of *abd-A*, does not develop prolegs (Ueno *et al.*, 1992).

One possibility is that *Ubx* represses proleg development, whereas *abd-A* may have a distinct regulatory role and might have associations to promote proleg development. Using RNAi to prevent Hox gene expression, Xiang (2008) also found that *Ubx* but not *abd-A* repressed ventral limb development. Here, we localized *E<sup>kp</sup>-1* to the *abd-A* region. Detailed analysis of this gene will hopefully allow the identification of the causal mutation.

Studies in *Drosophila* have shown that in this homeotic gene complex, many different classes of *cis*-regulatory elements, such as insulators, silencers, enhancers and promoters, work together to control gene expression during development (Sanchez-Herrero, 1991; Drewell *et al.*, 2002), and they have begun to unravel the intricate nature of these regulatory interactions, including the suggested activation of *cis*-sequences at the complex under the control of non-genic transcription, providing an additional level of regulatory complexity to regions of non-protein coding DNA (Akbari *et al.*, 2006; Chopra and Mishra, 2006). In particular, a recent study showed that in *Drosophila*, microRNA transcribed from the *iab-4* locus regulates *Ubx* and can result in transformation of the haltere into a wing (Ronshaugen *et al.*, 2005). In the present study, the putative Bm *iab-4* locus was in the *E<sup>kp</sup>-1* region, although no difference in sequence was found between p50 and *E<sup>kp</sup>* at the *iab-4* locus. This might be due to the very limited number of sequences compared between the two strains in this region. As it is only a predicted locus, further research is needed to isolate Bm *iab-4* and to characterize its functions. In addition, the *abd-A* region in silkworm is larger than that in *Drosophila*, so this region, which is about 220 kb in length, may have diverse regulatory elements such as *cis*-regulatory elements and non-coding loci. In the near future, the sequencing of the silkworm genome will be completed, allowing for more informative annotations. This may help identify the *E<sup>kp</sup>-1* locus and further our understanding of the mechanism of proleg development in Lepidoptera. Further, cloning of *E<sup>kp</sup>-1* may reveal key elements that respond to proleg development in Lepidoptera, such as non-coding loci.

## Acknowledgements

We thank Fei Liu, Yun-Po Zhao, Lin He and Peng-Guang Liang for their help with the DNA extractions and BAC pool construction, and Wen-Bin Liu for providing the cDNA library. We also thank Ming-Hui Li for key suggestions. We are grateful to Dr Zhong-Jun Qin for providing the CHEF equipment. This study was supported by the National Basic Research Program of China (2005CB121000), the National High-Tech R&D Program (2006AA10A119), and a Grant (06JC14079) from the Shanghai Commission of Science.

## References

Akbari OS, Bousum A, Bae E, Drewell ER (2006). Unraveling *cis*-regulatory mechanisms at the *abdominal-A* and *Abdominal-B* genes in the *Drosophila* bithorax complex. *Dev Biol* **293**: 294–304.  
Averof M (2002). Arthropod Hox genes: insights on the evolutionary forces that shape gene functions. *Curr Opin Gene Develop* **12**: 386–392.

Bender W, Akam M, Karch F, Beachy PA, Peifer M, Spierer P *et al.* (1983). Molecular genetics of the Bithorax complex in *Drosophila melanogaster*. *Science* **221**: 23–29.  
Chetelat RT, Meglic V, Cisneros P (2000). A genetic map of tomato based on BC(1) *Lycopersicon esculentum* × *Solanum lycopersicoides* reveals overall synteny but suppressed recombination between these homologous genomes. *Genetics* **154**: 857–867.  
Chopra VS, Mishra RK (2006). Miracles in Hox gene regulation. *Bioessays* **28**: 445–448.  
Drewell RA, Bae E, Burr J, Lewis EB (2002). Transcription defines the embryonic domains of *cis*-regulatory activity at the *Drosophila* bithorax complex. *Proc Natl Acad Sci USA* **99**: 16853–16858.  
Ganal MW, Tanksley SD (1996). Recombination around the Tm2a and Mi resistance genes in different crosses of *Lycopersiconperuvianum*. *Theor Appl Genet* **92**: 101–108.  
Goldsmith MR (1995). The genetics of the silkworm: revisiting an ancient model system. In: Goldsmith MR, Wilkins AS (eds). *Molecular Model Systems in the Lepidoptera*. Cambridge University Press: New York. pp 21–76.  
Goldsmith MR, Shimada T, Abe H (2005). The genetics and genomics of the silkworm, *Bombyx mori*. *Annu Rev Entomol* **50**: 71–100.  
Haupt W, Fischer TC, Winderl S, Fransz P, Torres-Ruiz RA (2001). The CENTROMERE1 (CEN1) region of *Arabidopsis thaliana*: architecture and functional impact of chromatin. *Plant J* **27**: 285–296.  
Hughes CL, Kaufman TC (2002). Hox genes and the evolution of the arthropod body plan. *Evol Dev* **4**: 459–499.  
Itikawa N (1943). Genetic and embryological studies of a dominant mutant, ‘new additional crescent’, of the silkworm. *Jap J Genet* **19**: 182–188.  
Karch F, Weiffenbach B, Peifer M, Bender W, Duncan I, Celniker S *et al.* (1985). The abdominal region of the Bithorax complex. *Cell* **43**: 81–96.  
Li L, Garvin DF (2003). Molecular mapping of Or, a gene inducing beta-carotene accumulation in cauliflower (*Brassica oleracea* L. var. *botrytis*). *Genome* **46**: 588–594.  
Li MW, Guo QH, Hou CX, Miao XX, Xu AY, Guo XJ *et al.* (2006). Linkage and mapping analyses of the densonucleosis non-susceptible gene *nsd-Z* in the silkworm *Bombyx mori* using SSR markers. *Genome* **49**: 397–402.  
Miao XX, Xu SJ, Li MH, Li MW, Huang JH, Dai FY *et al.* (2005). Simple sequence repeat-based consensus linkage map of *Bombyx mori*. *Proc Natl Acad Sci USA* **102**: 16303–16308.  
Murakami A, Imai HT (1974). Cytological evidence for holocentric chromosomes of the silkworm, *Bombyx mori* and *B. mandarina* (Bombycidae, Lepidoptera). *Chromosoma* **80**: 167–178.  
Nagy LM, Grbic M (1999). Cell lineages in larval development and evolution of holometabolous insects. In: Hall BK, Wake MH (eds). *The Origin and Evolution of Larval Forms*. Academic Press: San Diego. pp 275–300.  
Okazaki S, Tsuchida K, Maekawa H, Ishikawa H, Fujiwara H (1993). Identification of a pentanucleotide telomeric sequence, (TTAGG)*n*, in the silkworm *Bombyx mori* and in other insects. *Mol Cell Biol* **13**: 1424–1432.  
Ouweneel WJ (1976). Developmental genetics of homeosis. *Adv Genet* **18**: 179–248.  
Panganiban G, Nagy L, Carroll SB (1994). The role of the *Distal-less* gene in the development and evolution of insect limbs. *Curr Biol* **4**: 671–675.  
Pimpinelli S, Goday C (1989). Unusual kinetochores and chromatin diminution in *Parascaris*. *Trends Genet* **5**: 310–315.  
Ronshaugen M, Biemar F, Piel J, Levine M, Lai MC (2005). The *Drosophila* microRNA *iab-4* causes a dominant homeotic transformation of halteres to wings. *Genes Dev* **19**: 2947–2952.  
Sambrook J, Fritsch EF, Maniatis T (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press: New York. pp 2–50.

- Sanchez-Herrero E (1991). Control of the expression of the bithorax complex genes *abdominal-A* and *Abdominal-B* by *cis*-regulatory regions in *Drosophila* embryos. *Development* **111**: 437–449.
- Suzuki Y, Palopoli MF (2001). Evolution of insect abdominal appendages: are prolegs homologous or convergent traits? *Dev Genes Evol* **211**: 486–492.
- Tazima Y (1964). E-group as a tool of developmental genetics. In: Tazima Y (ed). *The Genetics of the Silkworm*. Logos Press: London. pp 60–75.
- Tong CZ, Jin YF, Zhang YZ (2006). Computational prediction of microRNA genes in silkworm genome. *J Zhejiang Univ Sci B* **7**: 806–816.
- Ueno K, Hui CC, Fukuta M, Suzuki Y (1992). Molecular analysis of the deletion mutants in the E homeotic complex of the silkworm *Bombyx mori*. *Development* **114**: 555–563.
- Ueno K, Suzuki Y (1997). p260/270 expressed in embryonic abdominal leg cells of *Bombyx mori* can transfer palmitate to peptides. *J Biol Chem* **272**: 13519–13526.
- Vachon G, Cohen B, Pfeifle C, McGuffin ME, Botas J, Cohen SM (1992). Homeotic genes of the bithorax complex repress limb development in the abdomen of the *Drosophila* embryo through the target gene *Distal-less*. *Cell* **71**: 437–450.
- Van Daelen RA, Gerbens F, van Ruissen F, Aarts J, Hontelez J, Zabel P (1993). Long-range physical maps of two loci (*Aps-1* and *GP79*) flanking the root-knot nematode resistance gene (*Mi*) near the centromere of tomato chromosome 6. *Plant Mol Biol* **23**: 185–192.
- Wang ZX, Yano M, Yamanouchi U, Iwamoto M, Monna L, Hayasaka H *et al.* (1999). The *Pib* gene for rice blast resistance belongs to the nucleotide binding and leucine rich repeat class of plant disease resistance genes. *Plant J* **19**: 55–64.
- Warren RW, Nagy L, Selegue J, Gates J, Carroll S (1994). Evolution of homeotic gene regulation and function in flies and butterflies. *Nature* **372**: 458–461.
- Wu C, Asakawa S, Shimizu N, Kawasaki S, Yasukochi Y (1999). Construction and characterization of bacterial artificial chromosome libraries from the silkworm, *Bombyx mori*. *Mol Gen Genet* **261**: 698–706.
- Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Li B *et al.* (2004). A draft sequence for the genome of the domesticated silkworm (*Bombyxmori*). *Science* **306**: 1937–1940.
- Xiang H (2008). Preliminary study of the structure and function of the E-complex genes in the silkworm (*Bombyx mori*). *Insect Gut Bacterial Community—Diversity Census and Functional Exploration*. PhD thesis, Shanghai Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, The Chinese Academy of Sciences, Shanghai, P.R. China. pp 96–109.
- Yasukochi Y, Ashakumary LA, Wu C, Yoshido A, Nohata J, Mita K *et al.* (2004). Organization of the Hox gene cluster of the silkworm, *Bombyxmori*: a split of the Hox cluster in a non-*Drosophila* insect. *Dev Genes Evol* **214**: 606–614.
- Yin JM, Guo WZ, Yang LM, Liu LW, Zhang TZ (2006). Physical mapping of the *Rf1* fertility-restoring gene to a 100 kb region in cotton. *Theor Appl Genet* **112**: 1318–1325.
- Zeng L, Yin Z, Chen J, Leung H, Wang GL (2002). Fine genetic mapping and physical delimitation of the lesion mimic gene *Spl11* to a 160-kb DNA segment of the rice genome. *Mol Genet Genomics* **268**: 253–261.
- Zheng Z, Khoo A, Fambrough Jr D, Garza L, Booker R (1999). Homeotic gene expression in the wild-type and a homeotic mutant of the moth *Manduca sexta*. *Dev Genes Evol* **209**: 460–472.