Insect Molecular Biology (1999) **8**(4), 491–500

A *P* element-homologous sequence in the house fly, *Musca domestica*

Seung Hoon Lee, Jonathan B.

© 1999 Blackwell Science Ltd

© 1999 Blackwell Science Ltd, *Insect Molecular Biology*

```
                                                                          ORF 0 ->
1    GAACTTCTGCTTTCCCCGGAGCACTATGCGGATAAAAATATCCAATTACAGTACTATTATTTACCAAAGAATCTGCACGCCACCGTGAAAAGCCCCTTTACA
                                             --> TRANSLATION START
101  CGAGCCTTGGGGTAAACAAATAAAAAGATTTATGAAGTTTATACATAGCAGTACTCTGTTATTCCCACTATTTACGAAATTATTATAACTTTTTTCTCAT
                                         M  K  F  I  H  S  S  T  L  L  F  P  L  F  T  K  L  L  *  L  F  S  H
201  ACTCATAAGTTGTAAAGGCACAGCGGGAATAAGGGAAAAAACGCCGAAAACGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTCATTTTCTACGAC
      T  H  K  L  *  R  H  S  G  N  K  G  K  N  A  E  N  E  G  Q  K  R  V  G  L  C  D  H  F  H  F  L  R
301  TTCTAGTCAACCCAACGTGCTCACCCAACAGCAAATGCTTACATACGTATCGTTTCTATTCGTTATGTGATCGTTGCCCATGTACACCAGCGAAAGCGGT
     L  L  V  N  P  T  C  S  P  N  S  K  C  L  H  T  Y  R  F  Y  S  L  C  D  R  C  P  C  T  P  A  K  A  V
401  CCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCCTACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCAAAGTATCGGCGTCAACTGTATGT
       L  A  E  N  D  F  E  R  C  R  H  L  S  Y  E  L  H  D  K  E  D  S  H  K  C  K  V  S  A  S  T  V  C
                                                                               ORF 1 ->
501  CGTTTGACCAAAACAGTGTAAAAATGTATACTTGTTATTATATAATTTATTCAGCTATAATAAAGTAGACGTTCATATACGGTACAAGAGAGTAGTAAAT
      R  L  T  K  T  V                                                           R  S  Y  T  V  Q  E  S  S  K
601  TAAGTAAAAATAAAGAGTGTAACAATAAACCATAATTGATATATATAAATTATTGCCAACGTAAATCTGATTACTAGGTAAGGCGTCCAAGTGGATGCTTTG
     L  S  K  N  K  E  C  N  N  K  P  *  L  I  Y  K  L  L  P  T  *  I  *  L  L  G  K  A  S  K  W  M  L  W
701  GAACAATTTCAAGGAATATCAAACAGACCAGGCCAAAAAAAAATGCCTAAAAAAATACAGTTACATGTTGGGATATATCGGGAGCTATTTGTTTCCACACC
       N  N  F  K  E  Y  Q  T  D  Q  A  K  K  N  A  *  K  N  T  V  T  C  W  D  I  S  G  A  I  C  F  H  T
801  GCTGGACCGAGGGCATACAATCATTTGTATGAAAAGGGGTTTCGTTTGCCTTCCCCTTTCCACATTGCAACGTTGGTGCTCAAAAATAGAAATAAAAGAAC
      A  G  P  R  A  Y  N  H  L  Y  E  K  G  F  R  L  P  S  L  S  T  L  Q  R  W  C  S  K  I  E  I  K  E
901  CTTTAATACAAGTTTCGTTGGACTTTATGGGCAATGCAACGAATATGTCAGAACAAGAAAAAATTTGTATTCTTGTATTCGACGAAATGAAGATTTTGGA
     P  L  I  Q  V  S  L  D  F  M  G  N  A  T  N  M  S  E  Q  E  K  I  C  I  L  V  F  D  E  M  K  I  L  E
1001 AACCCATGAATATGATAATTCTGGCGACTTCGTAAGAAAACCGACAAACTTTGTTCATGTCTTAATGGCCAGGGGACTGTGGAAATCCTGGAAACACCCA
       T  H  E  Y  D  N  S  G  D  F  V  R  K  P  T  N  F  V  H  V  L  M  A  R  G  L  W  K  S  W  K  H  P
1101 [illegible]
```

Figure 3. Nucleotide and derived amino acid sequences for the *M. domestica P* element. The underlined sequences identify the additional sequences present in the intron between exons 1 and 2 and the extra two introns present within exon 2 in the house fly. The TAT (tyrosine [Y]) codon that ends the putative exon 3 reading frame corresponds to the termination codons found at this position in *D. melanogaster* and *L. cuprina*. The precise ends of Md-P1 remain undefined.

cuprina, and from *M. domestica*. Figure 3 gives the nucleotide and derived amino acid sequences for the *M. domestica P* element and Fig. 4 shows the amino acid sequence alignments for exons 1 and 2 of *P* elements from *M. domestica*, *L. cuprina* and *D. melanogaster*.

As is the case with the Lu-P1 element from *L. cuprina*, the ends of Md-P1 are not well-defined in *M. domestica*. An additional 661 bp of this clone were sequenced in an attempt to find the inverted repeats. However, no candidate sequences were identified. Since Md-P1 is so degenerate

in sequence and structure, this is not surprising. The inverted repeats of Lu-P1 from *L. cuprina* are also elusive. Thus, no good evidence for the presence of *P* element inverted repeat termini has been found in either species. In contrast, most drosophilid *P* element coding regions are flanked by short inverted repeat sequences. However, these sequences are not particularly well conserved when *P*

simply diverged to a much greater extent than exons 1 and 2 when compared to sequences from *Drosophila* and the blow fly.

The *P* element sequence from *M. domestica* has lost its coding capacity. The reading frame is interrupted by numerous indels and nonsense mutations. Furthermore, conserved sequences found in the *P* element sequence from *D. melanogaster* are divergent in the house fly sequence. Three leucine zipper motifs have been identified in the canonical *P* element from *D. melanogaster* (Rio, 1990). The sequence alignment of the first, located between positions 101 and 136 in the *P*

© 1999 Blackwell Science Ltd, *Insect Molecular Biology*, **8**, 491–500

Discussion

Following the earlier description of *P* elements in the blow fly, *L. cuprina* (Perkins & Howells, 1992) this report represents only the second instance of *P* element sequence identification outside of the family Drosophilidae. In neither species is there clear evidence for recent activity of *P* elements in these genomes. This is not surprising because, to date, complete sequences have been obtained from only six species in the genus

this TE may be more widely distributed than previously thought. The use of additional PCR primers, designed to reflect the diversity of known sequences, may extend this distribution to other flies and perhaps other insects.

With the early success of *P* element germline transformation in *D. melanogaster* (Spradling & Rubin, 1982) and the development of efficient *P* element-based gene vectors (Steller & Pirrotta, 1985), there was initially considerable optimism that this element could be used as the basis of a generalized gene transformation system in a wide variety of insects. The development of *P* vectors carrying dominant selectable markers (Steller & Pirrotta, 1985) allowed this notion to be tested. In contrast to positive results in a number of Drosophilid species, negative results were obtained from several nondrosophilid species (Handler & O'Brochta, 1991). Furthermore, O'Brochta & Handler (1988) and O'Brochta *et al*. (1991) adapted an excision assay developed by Rio *et al*. (1986) to assess *P* functionality in the soma of both drosophilids and non-drosophilids. They found that although *P* elements could be mobilized in all the drosophilids tested, the *P* excision frequency decreased as a function of relatedness to *D. melanogaster* and *P* element mobility was not detected in tephritids, sphaerocerids, muscids or phorids (Handler & O'Brochta, 1991).

In light of the present results demonstrating that members of the

in these lineages. Such evidence provides another reason for confidence that overcoming the barriers to movement may not be as difficult as might have been thought. Despite several recent successes in developing new vector systems based on transposable elements other than *P* (e.g. Coates *et al*., 1998; Loukeris *et al*., 1995a, b; Jasinskiene *et al*., 1998), it may be necessary to develop an array of transformation systems with varying properties for use in different insects and under different conditions.

Experimental procedures

House fly strains

The five *M. domestica* strains used are listed in Table 1. They were kindly supplied by F. W. Plapp Jr, Department of Entomology, University of Arizona.

Primer design/PCR

Genomic DNA was prepared from wild-type *M. domestica* larvae as described (Cockburn & Seawright, 1988) and used as a template in PCR amplification. Degenerate primers (2684:GCTATTTGNY-TNCAYACCGCNGG, 2687:CCCAATGNATWGCANCGTCTKAT) were designed to correspond to two regions of the most conserved nucleic acid sequences located between nucleotides 703 (exon 1) and 1530 (exon 2) of the *P* elements. Primer 2684 is 256-fold degenerate; primer 2687 is 64-fold degenerate. Amplification reactions were carried out in 50 µL volumes, with 100 ng template DNA and 0.25 units of *Taq* polymerase (Gibco-BRL, Gaithersburg, MD). The reaction conditions were template denaturation for 1 min, 94 °C; primer annealing for 1 min, 50 °C; and primer extension for 1 min, 72 °C (with 2 s. added for each cycle), for a total of thirty cycles.

Southern hybridization

Genomic DNA was extracted from ten adult flies of each strain and 10 µg DNA was used for restriction digestion. Gel electrophoresis, Southern blotting and filter hybridizations were performed as described in (Sambrook *et al*., 1989). The filters were probed with the 1 kb PCR fragment generated from *M. domestica.* Filters were probed and washed under conditions of high stringency (65 °C, 0.1 × SSC final wash).

Library screening

A genomic DNA library of the *M. domestica* R-Diazinon (Rutgers) strain was kindly supplied by Rene Feyereisen. 150 000 plaques from the genomic library were screened by standard protocols (Sambrook *et al*., 1989) using the PCR fragment as a probe under stringent hybridization conditions.

Cloning and sequencing

Lambda DNA was isolated by the high yield method for isolation of lambda DNA (Lee & Clark, 1997). After mapping, the DNA fragment that was selected by Southern hybridization was subcloned into pBluscript vectors (Stratagene) by standard ligation and transformation techniques, using *E. coli* host strain DH5-alpha. The *P* element DNA sequence was obtained from both strands using both manual and automated sequencing techniques (the latter employed an ABI 377 automated DNA sequencer in the Laboratory of Molecular Systematics and Evolution, University of Arizona).

Sequence alignment

The conserved portion of the *P* element was aligned by eye. The more divergent ends of the *P* element sequence were aligned with the aid of computer programs Clustal W (Thompson *et al*., 1994) and GeneStream <http://eerie.fr/bin/align-guess.cgi>. Discrepancies between the programs were resolved by eye to maximize sequence similarity. The sequence of the *P* element from *M. domestica* has been submitted to GENBANK (accession number AF 183396).

Acknowledgements

We thank Rene Feyereisen for supplying the *M. domestica* genomic DNA library and F. W. Plapp Jr for supplying

Table 1. Description of *Musca domestica* strains used.

Name of strain	Abbreviation	Mutants	Origin	Description
Old Rutgers	R^W	Wild-type	Unknown	Has metabolic resistance to insecticides
New Rutgers	R^M	Wild-type	From New Jersey dairy barns in the early 1960s	Has metabolic resistance to insecticides
Cornell-R	C^R	Wild-type	Mid 1960s, NY state poultry houses	Combines metabolic and target site resistance to organophosphate insecticide
Cornell-R-H	C^{RH}			

M. domestica strains. We are indebted to Jake Tu, Becky Wattam, Joana Silva and three anonymous reviewers for comments on the manuscript. This work was supported by a postdoctoral fellowship to J.B.C. and a grant from the John D. and Catherine T. MacArthur Foundation for the sport of research on vector biology at the University of Arizona.

References

Anxolabéhère, D., Kidwell, M.G. and Périquet, G. (1988a) Molecular characteristics of diverse populations are consistent with the hypothesis of a recent invasion of *Drosophila melanogaster* by mobile P elements. *Mol Biol Evol* **5**: 252–269.

Anxolabéhère, D., Nouaud, D. and Périquet, G. (1985) Séquences homologues a l'élément *P* chez des espèces de *Drosophila* du groupe *obscura* et chez *Scaptomyza pallida* (*Drosophilidae*). *Génét Sél Evol* **17**: 579–584.

Anxolabéhère, D., Nouaud, D. and Périquet, G. (1988b) Evolutionary genetics of the *P* transposable elements in *Drosophila melanogaster* and in the *Drosophilidae family*. *Genet (Life Sci Adv)*

Rubin, G.M. and Spradling, A.C. (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**: 348–353.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning, a Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Simonelig, M. and Anxolabehere, D. (1991) A *P* element of Scaptomyza pallida is active in Drosophila melanogaster. *Proc Natl Acad Sci USA* **88**: 6102–6106.

Simonelig, M. and Anxolabéhère, D. (1994) P-elements are old components of the *Scaptomyza pallida* genome. *J Mol Evol* **38**: 232–240.

Spradling, A.C. and Rubin, G.M. (1982) Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* **218**: 341–347.

Steller, H. and Pirrotta, V. (1985) Fate of DNA injected into early Drosophila embryos. *Dev Biol* **109**: 54–62.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acids Res* **22**: 4673–4680.