Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 9216–9220, December 1990
Genetics

Evidence for horizontal transmission of the mobile element *jockey* between distant *Drosophila* species

(retroposon/internal promoter/*Drosophila funebris*)

LEV J. MIZROKHI*† AND ALEXANDER M. MAZO‡

*Laboratory of Biochemistry, National Cancer Institute, and ‡Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892

Communicated by Maxine Singer, August 31, 1990 (received for review July 13, 1990)

ABSTRACT **We addressed the possibility of the horizontal transfer of long interspersed element (LINE)-like mobile elements by studying the distribution of the *Drosophila melanogaster* LINE-like element *jockey* in different *Drosophila* species. Outside the *D. melanogaster* group *jockey* was detected only in the distantly related species *Drosophila funebris*. Cloning and sequencing of this element from *D. funebris* revealed the existence of the two open reading frames highly similar to those of *jockey* from *D. melanogaster*. Elements from both species are transcriptionally active and contain in their promoter regions a conserved sequence important for its activity. The high degree of similarity between the *D. melanogaster* and the *D. funebris jockey* and the absence of *jockey* from other sibling species of the *D. funebris* group provide evidence for the horizontal transmission of *jockey* into *D. funebris*.**

Transposition of long interspersed element (LINE)-like mobile elements (1–3) is an important source of mutations, at least in *Drosophila* (4) and humans (5, 6). Previously we have cloned the original copy of *jockey*, a member of this family of mobile elements, from *Drosophila melanogaster* ct^{MRpN} strain, where it is involved in the generation of the mutant phenotype (7). The sequence of *jockey* contains two open reading frames (ORFs) similar, respectively, to the *gag* and *pol* genes of retroviruses (8). There are two types of *jockey* in the genome of *D. melanogaster*: the full-size elements of about 5 kilobases (kb) and the copies with an internal deletion, which are approximately 2.5 kb long (8). Two polyadenylylated transcripts of *jockey* were detected at different stages of *Drosophila* ontogenesis and in cell culture. These transcripts have the same length as the genomic copies of *jockey* and correspond to the coding strand. Experiments with α-amanitin showed that *jockey* is transcribed by RNA polymerase II. We demonstrated also that the transcription of *jockey* is due to an internal promoter located downstream of the site of initiation of the transcription (9). Such an internal location of the promoter allows it to be preserved in the course of replication by means of reverse transcription and accounts for the distribution of *jockey* and probably other LINEs throughout the genome. However, nothing is known about the origin of LINEs or the modes of mechanisms of their spreading through different organisms. The fact that the putative reverse transcriptase of *jockey* is much more similar to the transcriptases from mammalian and *Trypanosoma* LINEs than to those from *Drosophila* long terminal repeat (LTR)-containing retroelements (8) suggests a close evolutionary relationship between LINEs from different species. A possible explanation for such a close relationship might be that this type of mobile element is capable of horizontal transmission between different organisms. We have chosen *jockey* as an active and well-studied LINE-like mobile element to test this hypothesis. In this paper, we report results that provide the evidence for horizontal transmission of *jockey* between distantly related *Drosophila* species.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

***Drosophila* Strains.** *Drosophila* strains were obtained from the Bowling Green, Ohio, and Umëa, Sweden, stock centers.

DNA Preparation, Electrophoresis, and Hybridization. DNA from adult flies was prepared as described previously (7), digested with *Hin*dIII, electrophoresed in 1% agarose, and transferred to a nylon filter. Hybridization was performed in 6× SSPE/10× Denhardt's solution/0.5% SDS containing sheared salmon sperm DNA at 100 μg/ml for 24 hr at 55°C. The filter was washed three times for 15 min in 2× SSC/0.5% SDS at 37°C. (1× SSPE = 0.18 M NaCl/10 mM sodium phosphate, pH 7.4/1 mM EDTA; 1× Denhardt's solution = 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone; 1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0.)

Library Construction and Screening. Genomic DNA from adult *Drosophila funebris* flies (strain 1911.1 from Bowling Green) was partially digested with *Sau*3A and ligated into phage EMBL3 (Stratagene). Hybridization conditions while screening were the same as described above.

DNA Sequencing. Nucleotide sequence was determined for both strands by using the dideoxynucleotide chain-termination method (10) and Sequenase (United States Biochemical). For the sequence comparisons, the programs from the University of Wisconsin Genetics Computer Group (UWGCG) were used (11).

RNA Preparation and Hybridization. Poly(A)$^+$ RNA from adult flies was purified as described earlier (9). Northern hybridization was conducted under the same conditions as Southern (see above). Washing conditions are specified in the legend to Fig. 3.

Constructs for Transfection and Chloramphenicol Acetyltransferase (CAT) Assay. Transfection of the *Drosophila* cell culture (Schneider 2), control of the efficiency of transfection, and CAT assay were performed as described earlier (9). To prepare plasmids for transfection, oligonucleotides corresponding to the first 41 base pairs (bp) (CAT 1) or 33 bp (CAT 2) of *D. melanogaster jockey* (8) were synthesized and inserted into the *Sal* I–*Bam*HI sites of pUC19 upstream of the previously inserted *Kpn* I–*Sac* I CAT fragment from pDM26 (12). Both constructs were checked by sequencing.

Abbreviations: LINE, long interspersed element; LTR, long terminal repeat; ORF, open reading frame; CAT, chloramphenicol acetyltransferase.

†To whom reprint requests should be addressed.

§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M38437).

RESULTS AND DISCUSSION

Screening of genomic DNA from a number of different species of the genus *Drosophila* with probes derived from different regions of *jockey* from *D. melanogaster* revealed decreasing hybridization within the *D. melanogaster* group which correlates with the phylogenetic relationships of the species (Fig. 1). A similar pattern of distribution has been described for other mobile elements, including the LINE-like *I* element (15, 16). No reproducible hybridization was detected with DNA from almost all tested species outside the *D. melanogaster* species group. The only exception was the signal seen with *D. funebris* DNA (Fig. 1, lane FU) which remained even after higher-stringency washes. This observation suggests that *jockey* may have been transferred horizontally between *Drosophila* species. To investigate this possibility we studied the structure of this element and its distribution in species related to *D. funebris*.

We prepared a λ phage library from *D. funebris* DNA and screened it with DNA of *jockey* from *D. melanogaster*. Two clones containing presumably full-length copies were selected for subcloning and sequencing. The complete sequence of one of the copies (Fig. 2) revealed two ORFs with 61% and 69% overall protein identity to the corresponding ORFs of *jockey* from *D. melanogaster*. The 5′ ORF-1 (*gag*) encodes a conserved region in the C-terminal half of its predicted protein that includes the proposed zinc-finger-like structure (8). The N-terminal part of this ORF-1 has little similarity to that of the element from *D. melanogaster*; this suggests the absence of conserved functional domains in this part of the protein. The protein homology of the ORF-2 (*pol*) is rather uniform, comparing the *D. funebris* and *D. melanogaster* elements, and does not reveal any conserved structures in addition to those identified previously (8). Sequencing of 5′ and 3′ nontranslated regions of a second cloned copy has shown that both copies are identical in these regions and that their sizes coincide with those of the full-length copy of *jockey* from *D. melanogaster*.

It was of particular interest to compare the DNA sequences from the 5′ nontranslated region of *jockey* because the *D. melanogaster* element contains an internal RNA polymerase II promoter in that region (9). In addition to the first few nucleotides, which are absolutely necessary for the transcriptional activity of the promoter (9) and show homology in both elements (Fig. 3*B*), there is a fully conserved 6-bp sequence, GGACGT, containing a short palindrome. To determine whether this sequence is important for promoter function, we constructed two plasmids based on the sequence of the promoter of the *D. melanogaster jockey*. One construct contained the 6-bp core (CAT1) and one did not (CAT2). In both clones the *jockey* segment was fused to the reporter CAT gene of *Escherichia coli* (Fig. 3*B*). The results of the CAT assay after transfection of *D. melanogaster* tissue

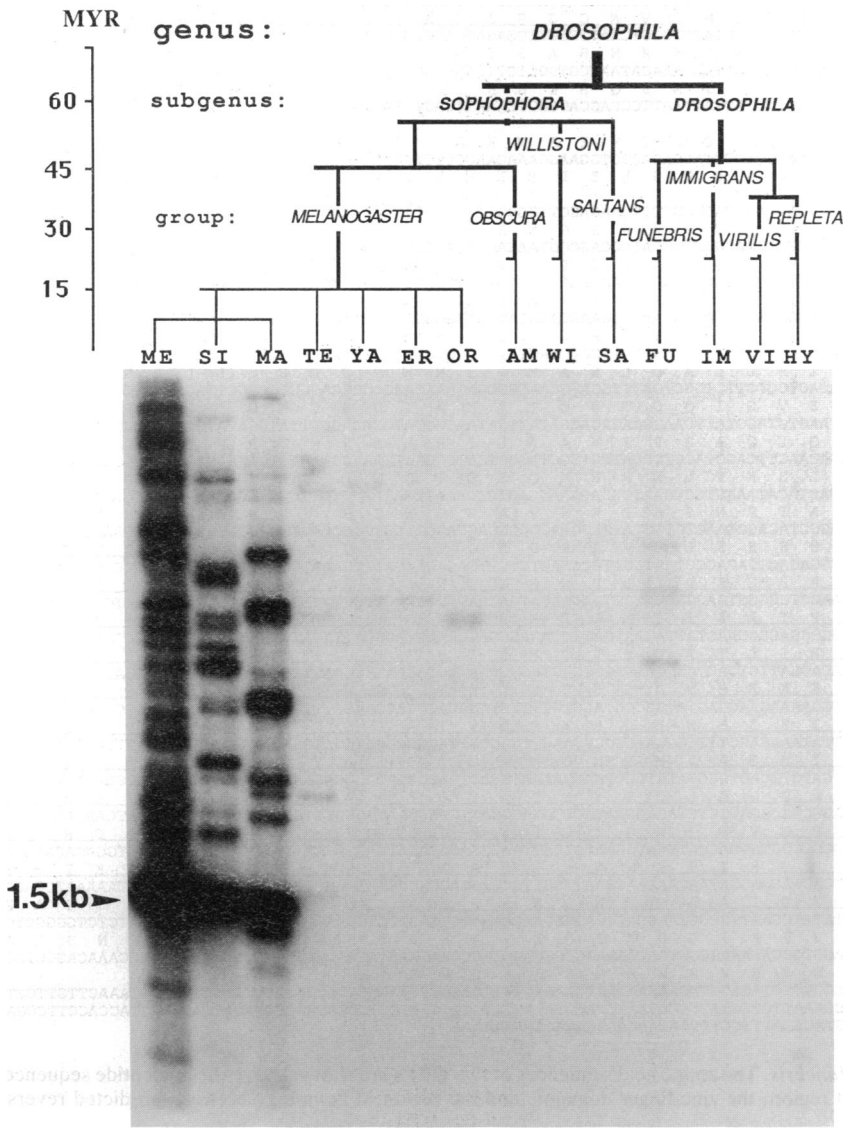

FIG. 1. Hybridization of *D. melanogaster jockey* DNA to genomic DNAs from different species of the genus *Drosophila*. The names of the species and their positions in the simplified phylogenetic tree are indicated above each lane with the following abbreviations: ME, *D. melanogaster*; SI, *D. simulans*; MA, *D. mauritiana*; TE, *D. teissieri*; YA, *D. yakuba*; ER, *D. erecta*; OR, *D. orena* (all from *melanogaster* group); AM, *D. ambigua* (*obscura* group); WI, *D. willistoni* (*willistoni* group); SA, *D. saltans* (*saltans* group); FU, *D. funebris* (*funebris* group); IM, *D. immigrans* (*immigrans* group); VI, *D. virilis* (*virilis* group); and HY, *D. hydei* (*repleta* group). DNAs of the following species of the subgenus *Sophophora* were tested and did not reveal any homology to *jockey* (data not shown): *D. takahashii*, *D. kikkawai*, *D. eugracilis* (different subgroups of the *melanogaster* group); *D. nebulosa* (*willistoni* group); *D. sturtevanti* (*saltans* group); and *D. pseudoobscura* (*obscura* group). The time scale given (MYR, millions of years) is approximate; for details of the phylogenetic relationships and time scale see refs. 13 and 14. A mixture of two *Pst* I fragments (1.4 and 2.2 kb) from ORF-1 and ORF-2 of *D. melanogaster jockey* (for map see ref. 8) was randomly labeled and used as a probe. The 1.5-kb internal fragment of *D. melanogaster jockey* is indicated by an arrowhead.

culture cells (Fig. 3C) showed that deletion of the 6-bp core abolished the transcriptional activity of the internal promoter. *jockey* is also actively transcribed in *D. funebris* (Fig. 3A), and we believe that this promoter element is involved in the regulation of *jockey* transcription in *D. funebris* as well, and this would explain the conservation of the core. C. McLean and D. Finnegan noticed (personal communication) that an almost identical 6-bp sequence, at the same position with respect to the site of initiation, is present in the downstream regulatory elements of engrailed, the P2 promoter of Antennapedia, and in several LINE-like mobile elements of *Drosophila*. For these mobile elements the existence of an internal promoter has been suggested (4, 9). It seems likely that the same protein(s) are required for transcription of these mobile elements and genes in *D. melanogaster*, and of *jockey* and other genes in *D. funebris*. We also found two long regions of conservation in the right part of the 5′ nontranslated region (Fig. 2), but their functional activity remains to be determined.

Southern hybridization experiments (Fig. 4) with a probe derived from the newly cloned element showed that *jockey* is present in all four strains of *D. funebris* tested in approxi-

```
   1  CAGTCTCTGTCCAGCATTCATTGCAGTC GGACGTACGCGCTTTGTAAAGTGAAACAACAACAACAACAACTCAACTCAACTCTTTCGCGGTGTGCCAAGTTAGGCTCGTGATCTATTTTT
 121  AGCTACTGCTGAGCGGTGGCGGCAGCTTATTACGGTACTTGCCATCGTTAAGCCAGTCTTAAGCGCAGCTTCGTTTGCCCCTGCCCGCTTTGCGCAAGCGAGAGGTTGATAC CGTTCTAA
                                                                                I  E  T  I  V  I  M  N  Y  G  N  S  S  T  R  S  L  P  S  D
 241  TTTCCCTTATCTACATTTTGATCTGGTGAAGTTGTTTACAATTGCCTATTTAACTATTAAATAGAAACAATTGTAATTATGAATTATGGAAACTCATCGACTCGCTCGCTACCTAGTGAT
      R  S  A  K  L  E  E  M  R  K  I  A  G  A  D  P  A  T  I  R  S  Q  I  I  A  S  G  A  S  T  S  N  Q  A  A  S  L  L  A  N
 361  AGGAGTGCAAAGCTTGAGGAGATGAGGAAAATAGCTGGTGCTGACCCTGCAACTATACGCTCCCAAATTATAGCTAGCGGTGCGTCAACGTCTAACCAAGCTGCGAGTTTATTGGCGAAC
      S  A  F  N  Y  P  S  V  T  T  S  S  D  T  S  F  T  C  T  V  T  T  S  Q  A  S  C  A  S  T  S  T  Y  T  V  A  A  R  V  G
 481  TCTGCGTTCAATTACCCATCGGTAACGACTAGCAGTGATACCAGCTTTACCTGCACTGTCACTACGTCTCAAGCAAGCTGTGCATCTACATCAACGTACACTGTGGCGGCACGTGTGGGG
      P  N  I  C  S  T  T  P  V  G  N  N  S  V  S  D  T  M  A  Q  M  A  G  K  A  A  R  K  G  K  C  S  F  Q  A  P  G  L  S  S
 601  CCGAACATCTGCTCAACTACGCCGGTTGGTAACAATTCGGTGAGTGATACAATGGCTCAAATGGCTGGAAAGGCTGCTCGTAAGGGAAAGTGCTCGTTTCAAGCACCTGGTCTATCTAGT
      S  K  K  R  K  V  V  V  V  N  A  C  Q  T  I  R  K  V  K  M  Q  A  T  D  L  R  K  A  T  P  K  V  A  T  A  S  A  T  A  T  T
 721  AGCAAGAAACGCAAAGTAGTTGTGAATGCTTGCCAAACCATTCGCAAAGTAAAGATGCAGGCGACTGACCTACGCAAGGCGACACCTAAAGTTGCTACAGCCTCGGCGACAGCAACGACC
      M  S  N  N  R  F  A  A  L  A  L  E  A  D  N  V  M  E  Q  D  L  A  E  N  C  S  N  G  E  E  D  V  P  V  V  N  M  N  R  Q
 841  ATGAGCAACAACCGCTTTGCAGCACTAGCTTTAGAAGCGGATAACGTTATGGAGCAAGACTTGGCAGAAAATTGCTCGAACGGAGAAGAAGATGTACCGGTTGTCAACATGAACAGGCAA
      S  T  N  G  P  S  K  P  P  A  I  C  V  P  N  V  D  N  P  H  S  L  E  Q  A  L  N  S  C  V  G  D  N  S  Y  S  I  R  T  S
 961  TCAACAAATGGTCCATCAAAGCCTCCAGCTATATGCGTTCCAAATGTTGACAATCCCCACTCATTGGAACAGGCATTAAATTCGTGCGTCGGTGACAATAGTTACAGTATTCGTACCTCC
      K  F  G  V  S  R  I  Y  T  I  N  A  D  A  F  R  A  V  V  K  H  L  T  S  L  N  C  Q  F  W  H  H  Q  L  R  E  D  K  P  Y
1081  AAATTCGGTGTGTCCCGAATATACACTATCAACGCCGATGCTTTCCGTGCAGTGGTGAAACACTTAACATCGTTGAACTGTCAATTCTGGCATCACCAGCTTAGGGAAGATAAGCCTTAT
      T  V  A  V  K  G  M  H  A  N  V  P  K  A  Q  I  E  H  A  F  A  D  H  G  F  E  A  L  N  I  Y  C  P  R  K  A  D  W  R  N
1201  ACTGTCGCAGTCAAGGGAATGCACGCTAACGTGCCCAAGGCACAAATCGAGCATGCTTTTGCTGACCACGGTTTCGAAGCCCTCAACATATATTGCCCGAGAAAGGCTGATTGGCGCAAT
      A  Q  V  S  E  D  D  D  E  A  T  V  N  F  K  T  R  Q  N  M  F  F  V  N  L  K  Q  G  P  K  I  A  E  A  L  K  I  T  Q  L
1321  GCCCAGGTTAGCGAGGATGACGATGAAGCGACCGTTAATTTCAAGACAAGGCAAAATATGTTTTTCGTCAACCTCAAACAAGGACCGAAGATCGCTGAAGCGTTGAAAATAACTCAACTA
      G  R  Y  R  V  T  V  E  R  A  P  R  R  R  E  V  L  Q  C  L  R  C  H  I  F  G  H  S  K  N  Y  C  V  R  D  P  I  C  A  K
1441  GGCAGATACAGAGTCACTGTGGAGCGTGCTCCTCGCAGAAGAGAGGTGCTACAGTGCCTGAGATGTCACATTTTTGGACATTCTAAAAATTATTGTGTCCGGGACCCCATATGCGCAAAG
      C  A  G  S  H  M  T  G  S  L  L  C  T  S  D  I  C  M  C  V  N  C  G  G  D  H  A  S  T  D  K  D  C  P  V  R  I  E  K  L
1561  TGCGCTGGCTCTCATATGACTGGCTCGCTACTTTGCACCAGTGACATTTGTATGTGTGTCAACTGTGGTGGAGATCATGCCTCGACTGATAAAGACTGCCCAGTCAGGATTGAAAAATTA
      R  K  M  K  P  S  P  R  L  P  V  P  D  N  A  T  N  N  K  H  N  R  A  S  S  A  R  G  F  I  P  A  E  A  V  R  G  N  M  S
1681  AGGAAAATGAAGCCGAGCCCGAGGCTCCCTGTTCCTGACAACGCAACCAATAATAAACATAATCGTGCCTCTTCAGCACGCGGCTTTATCCCCGCTGAAGCTGTCAGAGGCAATATGTCG
      Y  A  D  I  V  R  P  K  G  S  Q  F  R  A  T  A  T  I  N  S  Q  H  E  S  N  P  L  H  D  L  G  S  G  N  K  F  L  T  L  E
1801  TACGCTGACATTGTTCGACCTAAAGGATCTCAATTTAGAGCAACTGCTACTATTAATTCCCAGCATGAGTCAAATCCACTGCATGACTTAGGATCTGGCAATAAGTTCCTGACGTTGGAA
                                                                                                                              * M
      K  S  I  R  E  I  N  G  R  M  D  Q  L  F  K  L  V  Q  E  T  V  E  A  N  K  A  F  R  E  L  V  Q  V  L  I  S  R  M  P  K
1921  AAATCCATTAGAGAGATCAACGGCAGAATGGACCAACTATTCAAACTTGTGCAGGAGACTGTGGAAGCAAACAAGGCTTTCCGAGAACTGGTCCAGGTTCTCATCTCACGGATGCCTAAA
        T  Q  P  T  I  K  I  G  L  W  M  H  R  G  L  T  R  G  S  E  E  L  R  L  F  L  S  V  H  N  I  D  V  M  L  V  T  E  S  H
      *
2041  TGACTCAACCAACCATAAAAATCGGATTGTGGATGCACCGAGGACTAACTAGGGGCTCTGAAGAGCTTCGGCTTTTTCTCAGTGTACATAATATAGATGTCATGCTTGTCACTGAGTCAC
        M  R  E  G  Q  R  I  F  L  P  G  Y  S  T  Y  H  V  F  H  P  S  G  N  S  R  G  G  A  S  V  I  V  R  S  S  I  S  H  S  P
2161  ACATGCGGGAGGGTCAGCGGATCTTTCTACCTGGATATTCCACGTATCACGTCTTCCATCCCAGCGGTAACAGCAGAGGTGGGGCCTCAGTCATTGTGCGCTCTAGCATTAGCCACAGTC
        Q  P  P  I  S  T  N  D  R  Q  I  A  Y  I  Q  L  Q  T  A  E  G  P  V  V  L  A  A  V  Y  L  P  P  R  E  R  W  I  R  A  E
2281  CCCAGCCTCCCATCTCTACCAACGACAGGCAGATAGCATACATTCAGCTGCAGACAGCGGAGGGGCCTGTTGTTCTCGCGGCTGTGTATCTCCCACCTAGAGAGCGATGGATACGAGCTG
        F  E  S  L  F  A  V  L  G  N  K  F  I  A  G  G  D  Y  N  A  K  H  A  W  W  G  N  S  R  A  C  A  R  G  K  V  L  Q  E  V
2401  AATTTGAGTCGCTATTTGCCGTGCTAGGCAACAAATTCATTGCAGGAGGAGACTACAATGCCAAGCACGCATGGTGGGGCAACTCAAGAGCCTGTGCAAGAGGCAAGGTATTGCAGGAAG
        V  A  N  G  Q  Y  Q  I  L  A  T  G  E  P  T  F  Y  S  Y  N  P  L  V  S  P  S  A  L  D  F  F  V  V  N  G  Y  D  M  R  R
2521  TCGTTGCGAATGGACAGTACCAAATTTTAGCAACTGGAGAGCCGACCTTCTACTCCTACAATCCTCTAGTATCTCCGTCTGCGCTAGACTTCTTTGTCGTCAATGGGTACGACATGAGAA
        L  N  V  Q  T  L  H  E  L  S  S  D  H  T  P  L  L  A  D  L  H  A  M  P  I  N  K  P  P  R  S  C  L  L  A  R  G  A  D  I
2641  GGCTTAATGTACAAACGCTTCATGAGCTTTCATCAGACCACACTCCTCTTCTGGCTGATTTGCATGCAATGCCGATAAACAAGCCCCCACGCTCCTGCTTACTCGCTCGAGGTGCTGATA
        E  R  F  K  A  Y  L  T  Q  H  I  D  L  S  V  G  I  Q  G  T  D  D  I  D  N  A  I  D  S  F  M  D  I  L  K  R  A  A  I  R
2761  TAGAGAGATTCAAGGCTTATCTGACGCAACATATCGACCTAAGTGTAGGTATTCAAGGAACAGACGATATTGATAACGCCATCGACAGCTTTATGGATATCCTGAAGCGAGCAGCAATTC
        S  A  P  S  H  Q  Q  N  V  E  S  S  R  Q  L  Q  L  P  P  I  V  A  S  L  I  R  L  K  R  K  V  R  R  E  Y  A  R  T  G  D
2881  GATCGGCACCAAGTCACCAACAAAATGTGGAATCCTCTCGACAACTTCAGCTACCTCCCATTGTGGCTAGCCTCATCCGACTCAAACGTAAAGTCAGAAGAGAATATGCGAGAACTGGAG
        A  R  I  Q  Q  I  H  S  R  L  A  N  R  L  H  K  V  L  N  R  R  K  Q  S  Q  I  D  N  L  L  E  N  L  D  T  D  G  S  T  N
3001  ATGCCCGAATCCAGCAGATTCACAGTAGGCTTGCCAATCGACTACATAAAGTGCTGAATCGCAGAAAGCAGTCCCAAATTGATAATTTGTTGGAAAACTTGGACACTGATGGCTCCACAA
        F  S  L  W  R  I  T  K  R  Y  K  T  Q  A  T  P  N  S  A  I  R  N  P  A  G  G  W  C  R  T  S  R  E  K  T  E  V  F  A  N
3121  ATTTCTCTCTATGGCGTATCACAAAACGGTATAAGACCCAGGCTACACCGAACTCGGCTATAAGGAACCCGGCAGGAGGCTGGTGCCGCACATCACGGGAAAAAACAGAAGTGTTTGCGA
        H  L  E  Q  R  F  K  A  L  A  F  A  P  E  S  H  S  L  M  V  A  E  S  L  Q  T  P  F  Q  M  A  L  P  A  D  P  V  T  L  E
3241  ACCATTTAGAGCAACGCTTCAAAGCGCTCGCCTTTGCGCCCGGAGAGTCACAGCCTAATGGTTGCGGAGTCTCTTCAGACACCATTCCAAATGGCATTGCCCGCCGATCCTGTCACACTGG
        E  V  K  E  L  V  S  K  L  K  P  K  K  A  P  G  E  D  L  L  D  N  R  T  I  R  L  L  P  D  Q  A  L  L  Y  L  V  L  I  F
3361  AGGAAGTAAAGGAGCTTGTATCAAAATTGAAGCCCAAGAAAGCTCCGGGTGAGGACCTTCTTGACAACAGGACAATACGGCTCCTTCCAGATCAGGCTCTGCTATATCTGGTGCTAATAT
        N  S  I  L  R  V  G  Y  F  P  K  A  R  P  T  A  S  I  I  M  I  L  K  P  G  K  Q  P  L  D  V  D  S  Y  R  P  T  S  L  L
3481  TTAACAGCATCCTGCGAGTGGGATACTTTCCGAAAGCTCGGCCGACGGCAAGTATTATAATGATTCTAAAGCCAGGAAAGCAGCCATTAGATGTGGACTCATACAGGCCGACCAGCCTTC
        P  S  L  G  K  M  L  E  R  L  I  L  N  R  I  L  T  S  E  E  V  T  R  A  I  P  K  F  Q  F  G  F  R  L  Q  H  G  T  P  E
3601  TACCTTCACTGGGCAAGATGCTGGAGAGATTAATCCTTAACAGGATTCTCACTAGCGAGGAGGTCACTCGGGCGATCCCAAAATTCCAATTTGGTTTCCGGCTGCAGCATGGCACGCCAG
        Q  L  H  R  V  V  N  F  A  L  E  A  L  E  K  K  E  Y  A  G  S  C  F  L  D  I  Q  Q  A  F  D  R  V  W  H  P  G  L  L  Y
3721  AACAACTGCATAGGGTTGTTAATTTCGCGCTAGAGGCCCTGGAAAAGAAGGAGTATGCCGGTAGCTGTTTTTTAGACATCCAGCAAGCTTTCGATAGAGTGTGGCACCCTGGTCTTCTTT
        K  A  K  S  L  L  S  P  Q  L  F  Q  L  I  K  S  F  W  E  G  R  K  F  S  V  T  A  D  G  C  R  S  S  V  K  F  I  E  A  G
3841  ACAAGGCCAAGAGCTTGCTATCTCCCCAGCTGTTCCAGCTTATAAAGAGCTTTTGGGAAGGGCGCAAGTTCAGCGTGACGGCCGATGGATGCAGATCCTCTGTAAAATTTATTGAAGCCG
        V  P  Q  G  S  V  L  G  P  T  L  Y  S  I  F  T  A  D  M  P  N  Q  N  A  V  T  G  L  A  E  G  E  V  L  I  A  T  Y  A  D
3961  GCGTCCCCCAAGGCAGTGTTCTAGGCCCCACCCTGTACTCCATCTTCACTGCGGATATGCCTAACCAAAATGCAGTGACCGGACTCGCAGAAGGAGAGGTCCTAATCGCGACATATGCCG
        D  I  A  V  L  T  K  S  T  C  I  V  E  A  T  D  A  L  Q  E  Y  L  D  A  F  Q  E  W  A  V  K  W  N  V  S  I  N  A  G  K
4081  ACGACATCGCAGTGCTCACTAAAAGCACCTGCATTGTCGAGGCTACGGATGCTCTACAGGAATACTTGGACGCATTCCAGGAATGGGCTGTGAAGTGGAACGTGAGCATCAATGCCGGCA
        C  A  N  V  T  F  T  N  A  I  R  D  C  P  G  V  T  I  N  G  S  L  L  S  H  T  H  E  Y  K  Y  L  G  V  I  L  D  R  S  I
4201  AATGCGCAAACGTTACTTTTACAAACGCCATTCGCGACTGCCCCGGTGTCACGATCAATGGGAGCCTATTAAGCCACACACATGAGTACAAATACCTTGGTGTCATCCTGGACAGAAGCC
        T  F  R  R  H  I  T  S  L  Q  H  S  F  R  T  R  I  T  K  M  N  W  L  L  A  A  R  N  K  L  S  L  D  N  K  V  K  I  Y  K
4321  TGACCTTCAGAAGACATATTACGTCGTTGCAGCACTCATTCAGGACAAGGATAACAAAAATGAATTGGTTGCTTGCAGCACGCAATAAGCTGTCCTTGGACAACAAGGTAAAAATCTACA
        C  I  V  A  P  G  L  F  Y  A  I  Q  V  Y  G  I  A  A  R  T  H  L  N  K  I  R  V  L  Q  A  K  M  L  R  K  I  S  G  A  P
4441  AATGCATAGTAGCTCCAGGGCTTTTCTATGCCATCCAAGTATATGGAATTGCGGCAAGGACTCACCTGAACAAAATCCGAGTTCTGCAAGCTAAAATGCTGAGGAAGATCTCTGGGGCTC
        W  Y  M  R  T  R  D  I  E  C  D  L  K  V  P  K  I  G  D  K  I  R  E  V  A  K  K  Y  H  E  R  L  D  S  H  P  N  S  L  A
4561  CTTGGTACATGAGAACCAGGGACATCGAGTGCGATCTCAAGGTGCCAAAAATCGGTGACAAAATCCGTGAGGTTGCCAAGAAATATCACGAGAGACTGGATTCTCACCCAAACAGCCTGG
        R  R  L  G  V  A  A  Q  R  S  A  S  P  R  S  R  V  R  R  R  L  K  R  H  H  P  Q  D  L  L  D  R  A  L  T  *
4681  CCAGGAGACTCGGTGTTGCAGCACAGAGGTCTGCCAGCCCCCGCTCTAGAGTTAGAAGAAGACTCAAACGTCACCACCCTCAAGACCTGCTGGACCGGGCCTTGACGTAAACTTGTTCCT
4801  ATATTCCCAAAAGTAAAATTGTCTGTGATATCTACACCCCCAAATTCTTGTATCCCCTAAACTATGTATACTATGCCCGTTTCTTCCCATAGCCCTACATTATAAACTACCACCTTCCGA
4921  TTAGATTAAGCTGCCTTAAAGGTAGCAGGGCACTAATCTGTACATATTTCCTTTAATAAAAAAAAAAAAAAAAAAAA
```

FIG. 2. Sequence of the *jockey* element from *D. funebris*. The amino acid sequences of two ORFs are shown above the nucleotide sequence. The conservative sequences in the 5′ nontranslated region, the zinc-finger domains, and the region of homology between predicted reverse transcriptases of different LINEs are underlined.

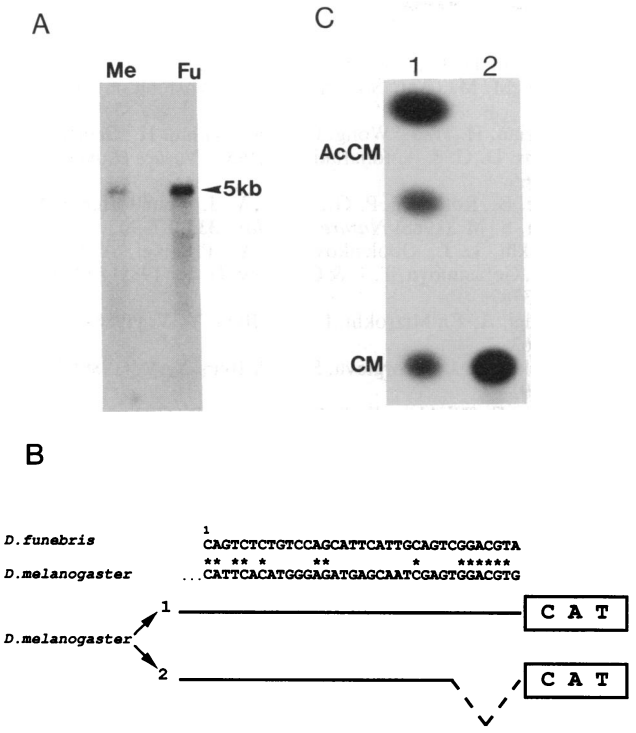

FIG. 3. *jockey* is transcribed in *D. funebris* and contains conserved promoter elements. (*A*) Northern blot containing 5 μg of poly(A)$^+$ RNA from *D. melanogaster* (lane Me) and *D. funebris* (lane Fu) was hybridized with the randomly labeled 1.4-kb *Pst* I fragment of *D. funebris jockey* (Fig. 4, fragment A) and washed in 0.1× SSC/0.5% SDS at 50°C. The region homologous to this probe is absent from copies of *D. melanogaster jockey* with internal deletion (8), and therefore only full-length 5-kb transcript is detected in lane Me. Increasing the stringency of washing (0.1× SSC, 68°C) leads to disappearance of the signal in lane Me (not shown). (*B*) Comparison of the promoter regions of *jockey* from *D. melanogaster* and *D. funebris*. The first nucleotide of the *D. funebris jockey* is indicated above its sequence. The design of the constructs used in the transfection experiments is shown on the bottom. (*C*) A 6-bp conserved sequence is essential for the activity of *jockey* promoter. CAT was assayed in extracts of *D. melanogaster* cultured cells (Schneider 2) transfected with the above constructions. Lanes 1 and 2 correspond to constructs CAT1 and CAT2, respectively. CM, chloramphenicol; AcCM, acetyl derivatives of CM.

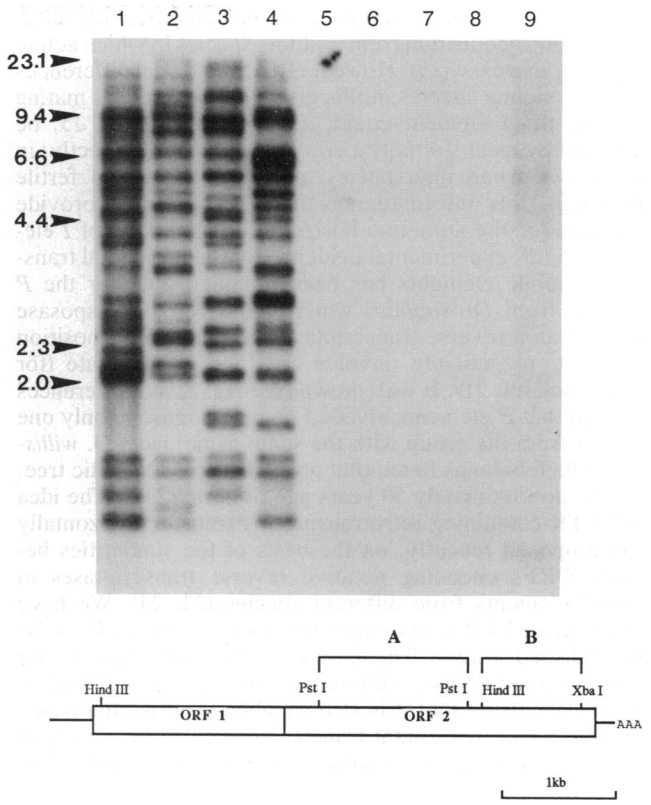

FIG. 4. Map of *jockey* from *D. funebris* (*Lower*) and autoradiogram of the Southern blot with DNA from different strains of the *D. funebris* group probed with a *jockey* fragment from *D. funebris* (*Upper*). The 900-bp *Hin*dIII–*Xba* I fragment (indicated as B on the map) was used as a probe. Because this fragment is not an internal *Hin*dIII fragment, it gives different bands for different copies of the element when hybridized to *Hin*dIII digests of genomic DNAs. The following strains from the Bowling Green collection were used (stock number and origin are in parentheses): 1–4, *D. funebris* (1911.1, Mexico; Umëa *Drosophila* Stock Center, Sweden; 1911.0, Kentucky; 1911.6, Alaska); 5, *D. subfunebris* (1951.51, California); 6, *D. microspina microspina* (1931.2, Mexico); 7, *D. microspina limpiensis* (1921.0, Arizona); 8, *D. multispina* (1941.0, Japan); 9, *D. melanogaster* (Oregon R). In addition, no signal was detected in lanes 5–8 when other fragments of *jockey* were used as probes (data not shown).

mately the same number of copies. The differences in hybridization patterns in the different strains of *D. funebris* is most likely caused by different localizations of the copies of *jockey* rather than by DNA polymorphism, because the same diversity was observed when a number of other restriction endonucleases were used (data not shown). This observation, together with the finding of the full-length copies and ORFs, suggests that *jockey* represents a family of active mobile elements in *D. funebris*. This conclusion has been supported by the analysis of the *jockey* RNA from *D. funebris* (Fig. 3*A*). Northern hybridization of the poly(A)$^+$ RNA of *D. funebris* flies with a probe derived from *D. funebris jockey* revealed the transcript that corresponds to the full-length copy of *jockey*.

We analyzed the distribution of *jockey* in other species within the *D. funebris* group. Surprisingly, no hybridization of *jockey* under low-stringency conditions was detected with DNA from *D. subfunebris*, *D. microspina microspina*, *D. microspina limpiensis*, or *D. multispina*—i.e., all the species of the *D. funebris* group described to date (Fig. 4, lanes 5–8). This result provides evidence for the horizontal transfer of *jockey* into the genome of *D. funebris* after its separation from the above-mentioned species.

The absence of *jockey* in sibling species from the same *D. funebris* group makes the explanation of this distribution by means of a vertical transmission from common ancestor unacceptable. Polyspermy should also be excluded because of the significant morphological differences between *D. melanogaster* and *D. funebris*. This leaves the horizontal transmission model as the only one that is consistent with the observed distribution of *jockey*. Southern hybridization of DNA of different species from within the *D. melanogaster* group with a probe from *jockey* from *D. funebris* indicates that the homology with DNA from *D. melanogaster* is the highest (data not shown)—i.e., *D. melanogaster* is the most likely species from which *jockey* was transferred to *D. funebris*.

The possibility of the horizontal transmission of the LINE-like *I* element in *Drosophila* has been discussed earlier (15). The distribution of *I* elements correlates with the phylogenetic relationship between species and is confined to the *D. melanogaster* group (15, 16). On the basis of population studies, Kidwell suggested (17) that the complete and active *I* factor appeared in natural populations of *D. melanogaster* in the 1930s–1940s. It was suggested (18) that it may have occurred through the reactivation of some inactive element(s)

which is present in all strains of *D. melanogaster* or, alternatively, by acquisition from a sibling species in which active elements also exist (19). However, morphological differences between sibling species in this group do not prevent mating (20), so the *I* element could, as suggested in ref. 15, be acquired by means of polyspermy or, even more directly, in rare cases when interspecies matings give rise to fertile offspring. Thus, information available now does not provide evidence for the authentic horizontal transmission of *I* element. So far, experimental evidence for the horizontal transfer of mobile elements has been obtained only for the *P* element from *Drosophila*, which codes for a transposase rather than a reverse transcriptase and whose transposition does not, presumably, involve an RNA intermediate (for review see ref. 21). It was shown (see ref. 22 and references therein) that *P* elements invaded *D. melanogaster* (only one species from the group with the same name) from *D. willistoni*, which belongs to another group (see phylogenetic tree, Fig. 1), approximately 50 years ago or earlier (17). The idea that LTR-containing retrotransposons transfer horizontally was proposed recently, on the basis of the similarities between ORFs encoding putative reverse transcriptases in mobile elements from different species (23, 24). We have found active LTR-containing retroelement *gypsy* in *D. virilis* (unpublished results). The features of the *gypsy* element and its distribution between different *Drosophila* species lead us to the conclusion that it is also capable of horizontal transmission. Thus, horizontal transfer appears to be a general feature of mobile elements and could play a significant role in eukaryotic evolution.

We thank M. F. Singer for advice and help throughout; M. M. Green for discussions; and M. F. Singer, S. Haynes, and I. B. Dawid for critical reading of the manuscript. L.J.M. thanks V. Corces for support while working in his laboratory.

1. Fanning, T. G. & Singer, M. F. (1987) *Biochim. Biophys. Acta* **910**, 203–212.
2. Boeke, J. D. & Corces, V. G. (1989) *Annu. Rev. Microbiol.* **43**, 403–434.
3. Hutchison, C. A., III, Hardies, S. C., Loeb, D. D., Shehee, W. R. & Edgell, M. H. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington), pp. 593–617.
4. Finnegan, D. J. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington), pp. 503–521.
5. Kazazian, H. H. J., Wong, C., Youssoufian, H., Scott, A. F., Phillips, D. G. & Antonarakis, S. (1988) *Nature (London)* **332**, 164–166.
6. Morse, B., Rothberg, P. G., South, V. J., Spandorfer, J. M. & Astrin, S. M. (1988) *Nature (London)* **333**, 87–90.
7. Mizrokhi, L. J., Obolenkova, L. A., Priimägi, A. F., Ilyin, Y. V., Gerasimova, T. I. & Georgiev, G. P. (1985) *EMBO J.* **4**, 3781–3787.
8. Priimägi, A. F., Mizrokhi, L. J. & Ilyin, Y. V. (1988) *Gene* **70**, 253–263.
9. Mizrokhi, L. J., Georgieva, S. G. & Ilyin, Y. V. (1988) *Cell* **54**, 685–691.
10. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
11. Devereux, J., Heaberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
12. Mismer, D. A. & Rubin, G. M. (1989) *Genetics* **121**, 77–87.
13. Throckmorton, L. H. (1975) in *Handbook of Genetics*, ed. King, R. C. (Plenum, New York), Vol. 3, pp. 421–669.
14. Beverley, S. M. & Wilson, A. C. (1984) *J. Mol. Evol.* **21**, 1–13.
15. Bucheton, A., Simonelig, M., Vaury, C. & Crozatier, M. (1986) *Nature (London)* **322**, 650–652.
16. Stacey, N. S., Lansman, R. A., Brock, H. W. & Grigliatti, T. A. (1986) *Mol. Biol. Evol.* **3**, 522–534.
17. Kidwell, M. G. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1655–1659.
18. Bucheton, A., Paro, R., Sang, H. H., Pelisson, A. & Finnegan, D. J. (1984) *Cell* **38**, 153–163.
19. Abad, P., Vaury, C., Pellison, A., Chaboissier, M.-C., Busseau, I. & Bucheton, A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8887–8891.
20. Sturtevant, A. H. (1920) *Genetics* **5**, 488–500.
21. Engels, W. R. (1989) in *Mobile DNA*, eds. Berg, D. E. & Howe, M. M. (Am. Soc. Microbiol., Washington), pp. 437–484.
22. Daniels, S. B., Peterson, K. B., Strausbaugh, L. D., Kidwell, M. G. & Chovnick, A. (1990) *Genetics* **124**, 339–355.
23. Doolittle, R. F., Feng, D.-F., Johnson, M. S. & McLure, M. A. (1989) *Q. Rev. Biol.* **64**, 1–30.
24. Smith, D. R., Kalitsis, P., Joseph, J. L. & Sentry, J. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5015–5019.