Phylogenetic Analysis Supports Horizontal Transfer of *P* Transposable Elements

Jonathan B. Clark, Wayne P. Maddison, and Margaret G. Kidwell

Department of Ecology and Evolutionary Biology, University of Arizona

Nucleotide sequence comparisons were used to investigate the evolution of *P* transposable elements and the possibility that horizontal transfer has played a role in their occurrence in natural populations of *Drosophila* and other Diptera. The phylogeny of *P* elements was examined using published sequences from eight dipteran taxa and a new, partial sequence from *Scaptomyza elmoi*. The results from a number of different analyses are highly consistent and reveal a *P*-element phylogeny that contradicts the phylogeny of the species. At least three instances of horizontal transfer are necessary to explain this incongruence, but other explanations cannot be ruled out at this time.

Introduction

P elements are mobile genetic sequences that were first described in *Drosophila melanogaster*. Their mobility in the germ line of offspring from crosses between certain strains is responsible for a number of phenotypic abnormalities known collectively as "hybrid dysgenesis" (Kidwell et al. 1977). Early work on hybrid dysgenesis indicated that *P* elements are not found in all strains of *D. melanogaster;* there is a correlation between the presence or absence of *P* elements and the length of time since the strain was collected from the wild (Kidwell 1983). In general, flies collected prior to ~40 years ago were found to be M strains, which were subsequently shown to completely lack these sequences, while those species collected more recently were shown to contain *P* elements (P strains) (Anxolabéhère et al. 1988). These results can best be explained by a recent invasion of natural populations of *D. melanogaster*, during the past half-century (Kidwell 1983).

An extensive survey of the genus *Drosophila*, using Southern blot hybridizations, has revealed that *P* elements are widely distributed; however, *P* elements have not been detected in those species most closely related to *D. melanogaster* (Daniels et al. 1990). This observation is consistent with the recent-invasion hypothesis, as are other observations, including the uneven worldwide geographical distribution of *P* elements in *D. melanogaster* (Anxolabéhère et al. 1988), the high sequence similarity among *D. melanogaster P* elements from diverse sources (O'Hare and Rubin 1983; Sakoyama 1985), and the highly invasive nature of *P* elements when introduced into a susceptible population (M strain) via germ-line transformation in the laboratory (Good et al. 1989).

There is a growing body of evidence (summarized in Daniels et al. 1990) that suggests that *P* elements entered the *D. melanogaster* genome by horizontal transfer from a species in the willistoni group of *Drosophila*. Most striking is the fact that the nucleotide sequence of a *P* element from *D. willistoni* differs from the canonical *D. melanogaster* sequence by a single nucleotide substitution. *P*-element nucleotide sequences have also been obtained from *D. nebulosa* (Lansman et al. 1987), *D. subobscura* (Paricio et al. 1992), *D. bifasciata* (Hagemann et al. 1992), and *D. guanche* (Miller et al. 1992), the drosophilid *Scaptomyza pallida* (Simonelig and Anxolabéhère 1991) and the blowfly, *Lucilia cuprina* (Perkins and Howells 1992). Here, we provide a phylogenetic analysis of *P*-element sequences, including published sequences from eight species of Diptera and a new, partial sequence from a ninth species. The results reveal a *P*-element phylogeny that is inconsistent with the phylogeny of the species themselves.

Key words: *P* elements, horizontal gene transfer, *Drosophila* evolution, mobile DNA, phylogenetic analysis.

Address for correspondence and reprints: Margaret G. Kidwell, University of Arizona, Department of Ecology and Evolutionary Biology, 310 Biosciences West, Tucson, Arizona 85721.

Mol. Biol. Evol. 11(1):40–50. 1994.

0737-4038/94/1101-0005$02.00

Materials and Methods

Nucleotide Sequences

Nucleotide sequences from the following species were obtained from the literature: *Drosophila melanogaster* (pπ25.1) (O'Hare and Rubin 1983), *D. bifasciata* (IbifM3), *D. guanche* (G1), *D. nebulosa* (N10), *D. subobscura* (DsA1; DsG2), *Scaptomyza pallida* PS2;

PS18), and *Lucilia cuprina* (Lu-P1). The *D. willistoni* sequence (P13) was determined previously in this laboratory (Daniels et al. 1990), while that from *S. elmoi* (Bowling Green Stock Center #31000-2651.3) was generated directly from a fragment amplified with the polymerase chain reaction (PCR). Genomic DNA was prepared from adult *S. elmoi* flies (Daniels and Strasbaugh 1986) and was used as a template in PCR with two degenerate primers: #2684 (GCTATTTGNC/TTNCAC/TACCGCNGG) is complementary to *D. melanogaster* *P*-element positions 703–725, and #2687 (CCCAATGNATA/TGCANCGTCTT/GAT) to positions 1508–1531 (O'Hare and Rubin 1983). Reaction conditions were four cycles, each of 94°C for 2 min; 37°C for 45 s, and 74°C for 2 min, followed by 25 cycles, each of 94°C for 2 min, 50°C for 45 s, and 72°C for 2 min. Reactions used the GeneAmp Kit (Cetus, Norwalk, Conn.) and followed the conditions recommended by the manufacturer, except that 15 mM $MgCl_2$ was used. After PCR, the product was separated on a 1.2% agarose gel, and the fragment was isolated using glass milk (GeneClean; Bio 101, La Jolla, Calif.). The fragment was cloned by ligating into pCRII (Invitrogen, San Diego) and subsequently was sequenced with T7 polymerase (Sequenase; U.S. Biochemical, Cleveland). Each nucleotide was resolved at least twice, either from the opposite or the same strand.

Phylogenetic Analysis

Nucleotide sequences of exons 1 and 2 were aligned manually after identification of reading frames and intron boundaries in the *P* elements. Sequences were then translated into amino acids, and the nucleotide sequence alignments were refined by considering the codons. This alignment is available from the authors by request. The alignment was also examined using the CLUSTAL V program (Higgins et al. 1992). The two alignments differ significantly only at the beginning of exon 1, where there is considerable divergence of the *L. cuprina* molecule and four extra codons in the sequences from *D. subobscura* and *D. guanche*. This region, of ~30 nt, was not used in the sequence analyses. The phylogenetic affiliations of *P* elements from nine species were determined by nucleotide sequence comparisons. Relationships were analyzed using PAUP 3.0 (Swofford 1990) (parsimony), PHYLIP 3.4 (Felsenstein 1991) (maximum likelihood), and NJBOOT2 (K. Tamura) (neighbor joining). Data editing and some tree comparisons were performed using MacClade 3.0 (Maddison and Maddison 1992).

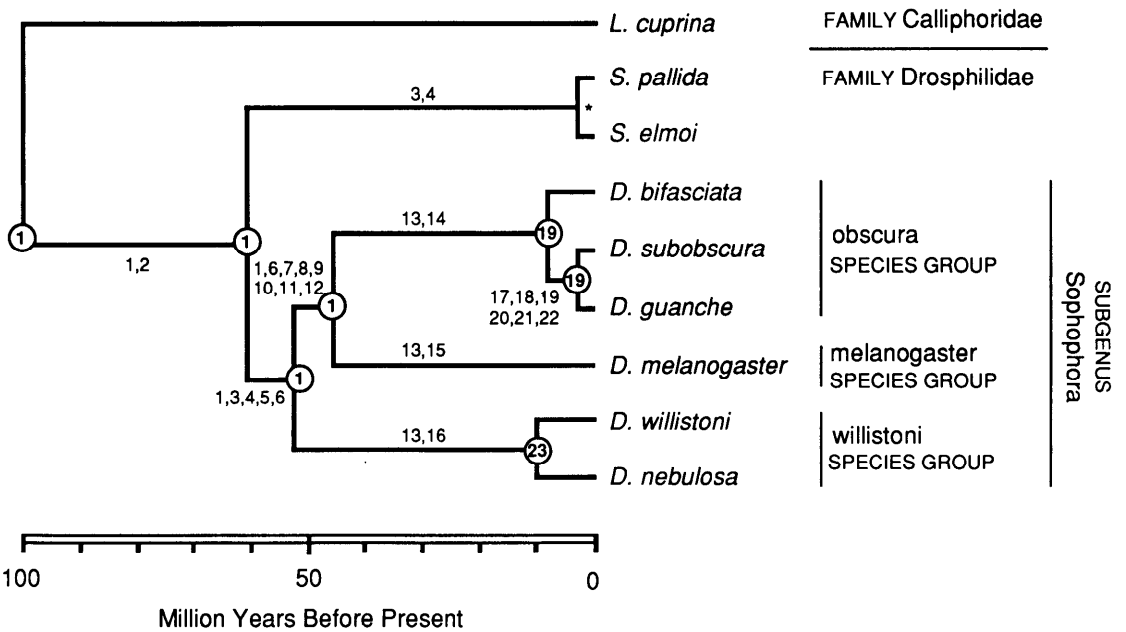

FIG. 1.—Phylogeny of nine species of Diptera that are the subject of this study. All organisms belong to the dipteran suborder Muscomorpha (= Cyclorrhapha) and, within this suborder, the division Schizophora. The Drosophilidae belongs to the section Acalyptratae, while Calliphoridae is a family in the section Calyptratae. Numbers represent particular studies, listed in table 1, that support the branching pattern depicted here. The references to synapomorphies that describe the three species groups of the subgenus *Sophophora* include the original description of the subgenera, based primarily on morphology, and more accessible reviews. Numbers in circles are those studies, listed in table 1, used to provide approximate divergence times for particular nodes (see scale bar). The asterisk indicates that there are no estimates available for divergence times within the genus *Scaptomyza*.

Table 1
List of Phylogenetic Studies Supporting Various Nodes and Branches in the Phylogeny Presented in Figure 1

Number[a]	Type of Study	Reference
1	Microcomplement fixation	Beverley and Wilson 1984
2	Morphology	McAlpine 1989
3	Morphology	Grimaldi 1990
4	Morphology	Carson et al. 1970
5	Fossils	Grimaldi 1987
6	*Adh* nucleotide sequences	C. Russo and M. Nei, personal communication
7	Mitochondrial DNA sequences	DeSalle 1992
8	28S rDNA sequences	Pélandakis et al. 1991
9	Complementary DNA hybridization	Caccone et al. 1992
10	*Adh* nucleotide sequences	Anderson et al. 1993
11	Biogeography; morphology	Throckmorton 1975
12	DNA-DNA hybridization	Zwiebel et al. 1982
13	Morphology	Sturtevant 1942
14	Review	Buzzati-Traverso and Scossirolli 1955
15	Review	Bock 1980
16	Review	Dobzhansky and Powell 1975
17	Enzyme electrophoresis	Cairou et al. 1988
18	Nuclear DNA RFLP	Loukas et al. 1986
19	Enzyme electrophoresis	Loukas et al. 1984
20	Mitochondrial DNA RFLP	González et al. 1990
21	Mitochondrial DNA RFLP	Barrio et al. 1992
22	Enzyme elecrophoresis	Lakovaara and Saura 1982
23	DNA-DNA hybridization	J. Powell, personal communication

[a] Numbers correspond to those given in fig. 1.

Results

Most of the species in this study belong to the *Sophophora* subgenus of *Drosophila,* whose members and the associations among them are well-studied; these relationships are summarized in figure 1. This phylogeny is supported by numerous molecular and classical studies, as listed in table 1. Within the subgenus *Sophophora,* the melanogaster species group contains ~80 named species, the saltans-willistoni species group ~50, and the obscura species group ~25 species. Two outlying drosophilid species from the genus *Scaptomyza* were also included in this study, as was one species, the sheep blowfly, *Lucilia cuprina,* from the calyptrate family, Calliphoridae.

Figure 2 provides a schematic representation of the structure of the *P* elements from the species examined here and summarizes the variation in the structure and nucleotide composition. The complete *P* element from *D. melanogaster* comprises four exons (0–3). In germline tissue these exons are spliced correctly, producing a messenger RNA (mRNA) that encodes an 87-kD transposase necessary for *P*-element mobility (Rio et al. 1986). In somatic tissue, the intron separating exons 2 and 3 is not removed, and a 66-kD repressor protein is produced from an mRNA that ends within exon 3 (Misra and Rio 1990). The ends of the complete element are defined by perfect 31-bp inverted repeats (IRs), which are adjacent to transposase-binding sites and are also necessary for mobility.

The *P* element from *D. bifasciata* (IbifM3) has all of the structural features of the *D. melanogaster P* element. Although mobility has not been demonstrated, it is suggested by the differences in copy number observed among different geographic isolates (Hagemann et al. 1992). Exons 0–2 of the *D. guanche P* element share high similarity with those of *D. melanogaster.* Exon 3, however, has diverged considerably and has no coding potential, because of stop codons in all three reading frames. Like elements from other obscura-group species, the *D. guanche* element could encode a putative repressor protein (Miller et al. 1992). While the overall structure of the *D. nebulosa P* element is intact, all four exons are blocked by termination codons (Lansman et al. 1987). In the nucleotide sequence alignment used here, the *D. nebulosa* sequence contains a single stop codon, two single-base deletions, and two single-base insertions in exon 1. In the exon 2 alignment, the sequence contains an insertion of 4 nt, as well as two contiguous stop codons. Both *P*-element copies from *D. subobscura* lack exon 3; exons 0–2 of both DsA1 and DsG2 could, how-

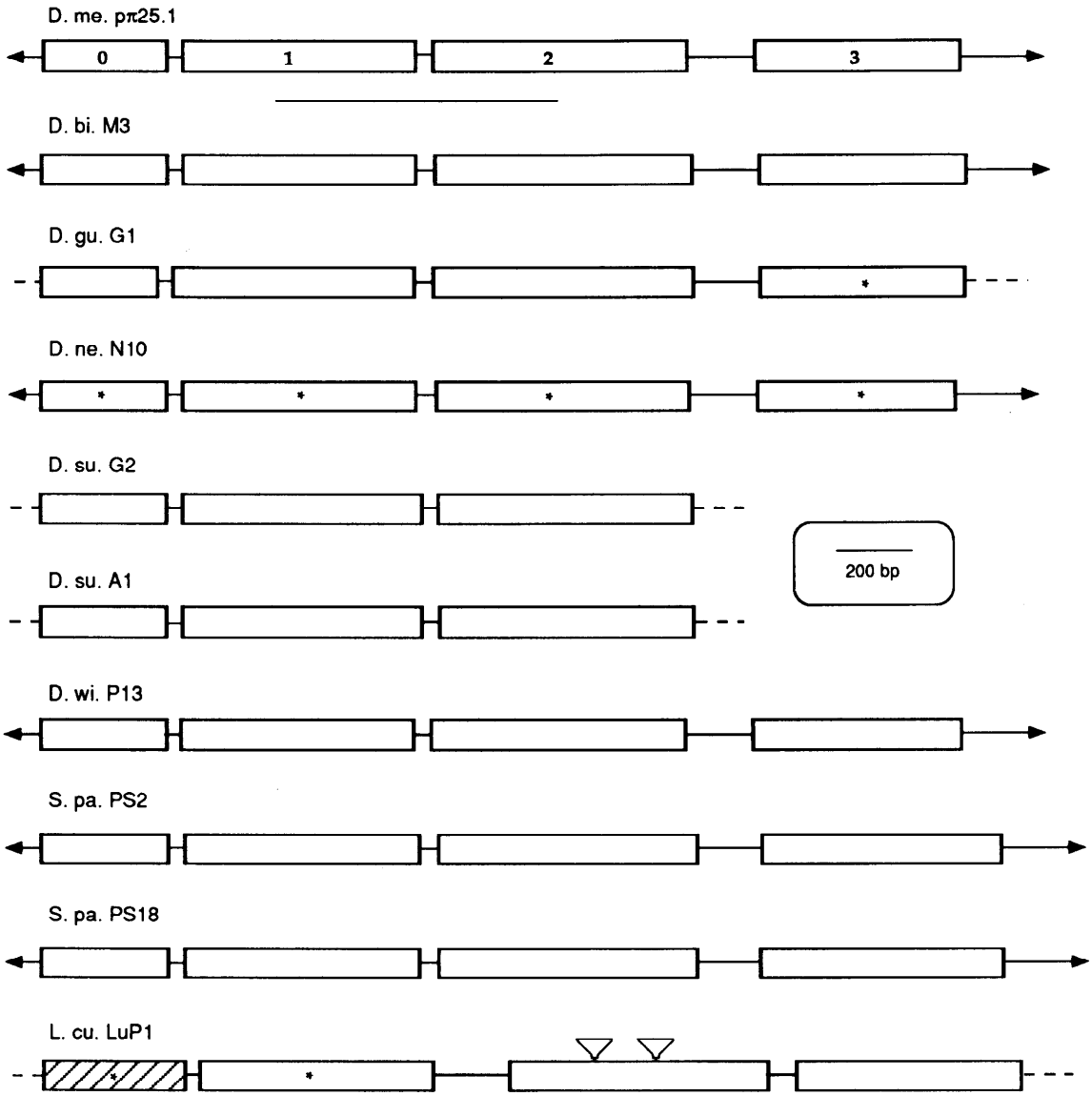

FIG. 2.—Schematic representation of the *P*-element structure in eight species. The prototype element, that from *Drosophila melanogaster*, is 2,907 bp in length and encodes an 87-kD transposase. The inverted repeats that flank the complete *P* element are denoted by arrowheads. The absence of arrowheads in some species indicates that no canonical inverted repeats are found in those elements. A dashed line indicates that the 5′ and 3′ ends of the element are not well-defined. The demarcations of the *P* element as indicated are, from left to right, 5′ inverted repeat, 5′ untranslated region, exon 0, intron, exon 1, intron, exon 2, intron, exon 3, 3′ untranslated region, and 3′ inverted repeat. The length of the various segments can be determined from the scale bar (*inset*). An asterisk within an exon indicates that all three reading frames are interrupted by stop codons. The *Lucilia cuprina* exon 0 is shaded because it lacks significant nucleotide or amino acid similarity with exon 0 of the other species. Exon 2 of the *L. cuprina* element is interrupted uniquely by two small introns as indicated. The relative location of the 838-bp fragment (777 bp without the intron sequence) amplified by PCR in *Scaptomyza elmoi* is denoted by the line under the *D. melanogaster* *P* element.

ever, encode a repressor-like protein, and a corresponding mRNA has been detected in this species (Paricio et al. 1992). The *P*-element nucleotide sequence from *D. willistoni* is identical to that of *D. melanogaster*, except for a single transition substitution at position 32, just outside the IR region (Daniels et al. 1990). This element is active when microinjected into appropriate strains of *D. melanogaster*. In spite of 24% sequence divergence from the *D. melanogaster P* element, at least one of the *P* elements from *S. pallida* is active when microinjected into *D. melanogaster* embryos (Simonelig and Anxolabéhère 1991). Exons 1–3 of the *L. cuprina P* element share obvious sequence similarity with all of the drosophilid *P* elements; however, there is no amino acid or

nucleotide similarity in exon 0. In the alignment used here, there is a single stop codon in exon 1 of the *L. cuprina* *P*-element sequence.

Because of (1) both the high degree of divergence within exon 3 in the *D. guanche* element and the absence of this exon in the *D. subobscura* elements, and (2) the divergence of exon 0 of the *L. cuprina* element, the only regions that show obvious similarity in all eight species are those that include exons 1 and 2, representing 1,404 positions, including gaps necessary for correct alignment. Because of divergence of the *P*-element sequence from *L. cuprina* and those from *D. subobscura* and *D. guanche* at the beginning of exon 1, the sequence analyses begin with codon 7 (GAA) of exon 1 of the *D. melanogaster* *P* element, where the obvious homology begins. Both the nucleotide differentiation over this two-exon region and a genetic distance matrix based on Kimura's two-parameter method (Kimura 1980) are presented in table 2.

Figure 3 summarizes the results of phylogenetic analysis limited to exons 1 and 2, with *L. cuprina* used as the outgroup. The cladogram (fig. 3*A*) is the single most-parsimonious reconstruction and requires 1,246 steps. The next most-parsimonious tree (not shown) requires 1,257 steps and differs only in the placement of *D. guanche* with respect to the two *D. subobscura* sequences: it clusters *D. guanche* with *D. subobscura* G2, while a reconstruction of 1,258 steps places *D. guanche* with *D. subobscura* A1. Bootstrap values are not given for the maximum-likelihood tree (fig. 3*B*), because they are not computationally practical. However, additional searches with differing transition:transversion ratios and a variety of other options offered by DNAML, including global or nonglobal searches, as well as randomized or nonrandomized sequence additions, do not yield results different from those in this tree. Neighbor-joining analyses with various genetic distance measures also yield trees that do not differ from that shown in figure 3*C*.

The consistent branching patterns illustrated in figure 3 are clearly incongruent with the species phylogeny of figure 1. The specific points of disagreement are the close relationship among the *P* elements from *D. melanogaster* and the two willistoni-group species, *D. nebulosa* and *D. willistoni,* the distinction of the clade that includes *D. subobscura* and *D. guanche* from the other *Drosophila P* elements, and the close affiliation of the sequences from *D. bifasciata* and *S. pallida.* To explore the extent of this incongruence, the *P*-element data were examined on the framework of the species branching pattern presented in figure 1. This tree requires 1,482 steps for the *P*-element data, 236 steps more than the most-parsimonious tree of figure 3*A*. Even if the analysis is constrained to require only that the *Drosophila P*-element sequences be monophyletic, the most-parsimonious tree is one with 47 more steps (1,246 vs. 1,293) than the tree in figure 3A. The *P*-element data therefore strongly prefer a tree incongruent with the species phylogeny. This incongruence cannot be due to a misrooting of the *P*-element tree by using *L. cuprina;* there is no alternative rooting that can bring the *P*-element tree into congruence with the species tree.

As seen in table 2, the extent of some of the *P*-element sequence divergences is quite large, especially for those involving *L. cuprina.* However, the alignment of these nucleotides, as based on amino acid translations, is reasonably straightforward. Although gaps were not used in the analyses, the distribution of gaps in the sequence alignments provides additional support for the phylogenetic conclusions. For example, the *P*-element sequence from *D. guanche* and both sequences from *D. subobscura* share an extra 12 nt in exon 1 (encoding four amino acids) that are found in none of the other

Table 2
Nucleotide Sequence Differentiation among *P* Elements from *Lucilia, Scaptomyza,* and *Drosophila* Species

	L. cuprina	*S. pallida* 18	*S. pallida* 2	*D. bifasciata*	*D. subobscura* A1	*D. subobscura* G2	*D. guanche*	*D. melanogaster*	*D. nebulosa*	*D. willistoni*
L. cuprina		0.44	0.44	0.44	0.46	0.46	0.46	0.43	0.45	0.43
S. pallida 18	0.66		0.02	0.07	0.29	0.29	0.29	0.21	0.24	0.21
S. pallida 2	0.67	0.02		0.06	0.30	0.30	0.29	0.21	0.24	0.21
D. bifasciata	0.66	0.07	0.06		0.28	0.28	0.28	0.20	0.23	0.20
D. subobscura A1	0.72	0.38	0.38	0.35		0.01	0.03	0.29	0.32	0.29
D. subobscura G2	0.72	0.38	0.38	0.35	0.01		0.03	0.29	0.32	0.29
D. guanche	0.72	0.37	0.37	0.35	0.03	0.03		0.29	0.32	0.29
D. melanogaster	0.65	0.25	0.25	0.23	0.38	0.38	0.38		0.06	0.00
D. nebulosa	0.68	0.30	0.30	0.28	0.42	0.42	0.42	0.07		0.06
D. willistoni	0.65	0.25	0.25	0.23	0.38	0.38	0.38	0.00	0.07	

NOTE.—Numbers above the diagonal are the proportion of differences between taxa, of 1,336 nucleotide postions confined to exons 1 and 2; numbers below the diagonal are genetic distances calculated using Kimura's (1980) two-parameter method.

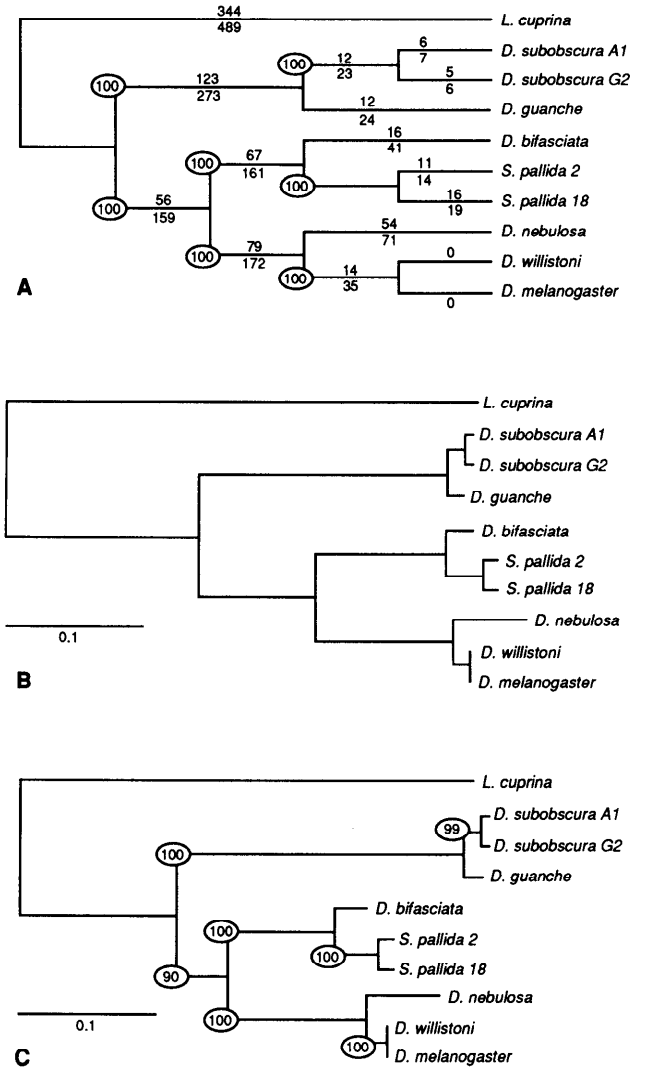

FIG. 3.—Phylogenetic analysis of *P*-element nucleotide sequences. The comparisons were limited to 1,336 positions (excluding gaps) constituting the alignment of exons 1 and 2 of the *P*-element sequences. In all cases, the *P* element of *Lucilia cuprina* was used as an outgroup. *A*, Cladogram generated by parsimony analysis using the branch-and-bound algorithm of PAUP (Swofford 1990). All characters were unordered, and gaps were treated as missing data. The minimum and maximum numbers of nucleotide substitutions on each branch, allowed by parsimonious reconstruction of changes, are shown. This tree is the single most-parsimonious reconstruction and requires 1,246 steps. The consistency index is 0.89, and the retention index is 0.90. The numbers in the ovals indicate the percent of 500 bootstrap replications that contain that clade. *B*, Tree generated by maximum-likelihood (Felsenstein 1981) analysis as implemented by PHYLIP (Felsenstein 1991). The Kimura two-parameter method for genetic distance was used with a transition:transversion ratio of 2.0. Empirical base frequencies were used for this tree, with the global search option as specified in the program DNAML. The branch lengths are expressed in terms of expected nucleotide substitutions per site. *C*, Tree generated by the neighbor-joining method (Saitou and Nei 1987) as implemented by the program NJBOOT2 of K. Tamura. The distance matrix was generated by Kimura's two-parameter method. The branch lengths are given in terms of nucleotide substitutions per site. Bootstrap values are percentages of 500 replications.

P-element molecules. Both of the *D. subobscura* molecules and that from *D. guanche* also exclusively share a valine codon (GTT/G) in exon 2. Such occurrences further illustrate the distinction between the *P* elements of *D. guanche* and *D. subobscura* and those of the other species in the genus, including their sister obscura-group species, *D. bifasciata*. Almost all of the other gaps necessary to align the sequences correctly are associated with the highly divergent *L. cuprina* molecule and, less often, with the nonfunctional *P* element from *D. nebulosa*. When the phylogenetic analyses are repeated using the alignment generated by CLUSTAL V, there are no differences from figure 3 in the branching patterns (or bootstrap values) of any of the taxa.

The affiliation of the *P*-element sequences from *S. pallida* and *D. bifasciata* clearly contradicts the established phylogenetic relationship of these two species, as pointed out elsewhere (Hagemann et al. 1992). In order to determine whether this is because of a peculiarity of the *S. pallida* sequences, the *P* element of a closely related species, *S. elmoi*, was examined. Degenerate oligonucleotide primers to a conserved region spanning exons 1 and 2 (see fig. 2) were designed for use in PCR. The nucleotide sequence of this 838-bp fragment (777 bp when the intron is excluded) from *S. elmoi* is presented in figure 4. It is apparent that the *P*-element sequence sampled from this species does not significantly differ from the published *S. pallida* sequences. Parsimony analysis limited to the region defined by the PCR fragment and including now the sequence from *S. elmoi* yields a branching pattern (not shown) identical with that of figure 3*A*. Again, there is a single most-parsimonious tree, here one of 616 steps, when the branch-and-bound algorithm is used and *L. cuprina* is an outgroup. The next shortest tree (618 steps) and that within 4 steps of the shortest differ only in the branching among the three *Scaptomyza* sequences. The branching patterns of maximum-likelihood and neighbor-joining trees are also in complete agreement with those in figure 3*B* and *C*: the three *Scaptomyza* sequences form a clade that clusters with the *P*-element sequence from *D. bifasciata*.

Discussion

The results of *P*-element phylogenetic comparisons presented here suggest that the evolution of these mobile sequences may be complex. A phylogenetic tree based on *P* elements that is congruent with a tree of their host organisms would suggest that the elements have had a long association with the species lineages and would provide no reason for believing that transmission has been other than vertical. Indeed, the *P*-element and species trees do agree on some of the smaller clades. Specifically, the *P* elements from *Drosophila subobscura* and

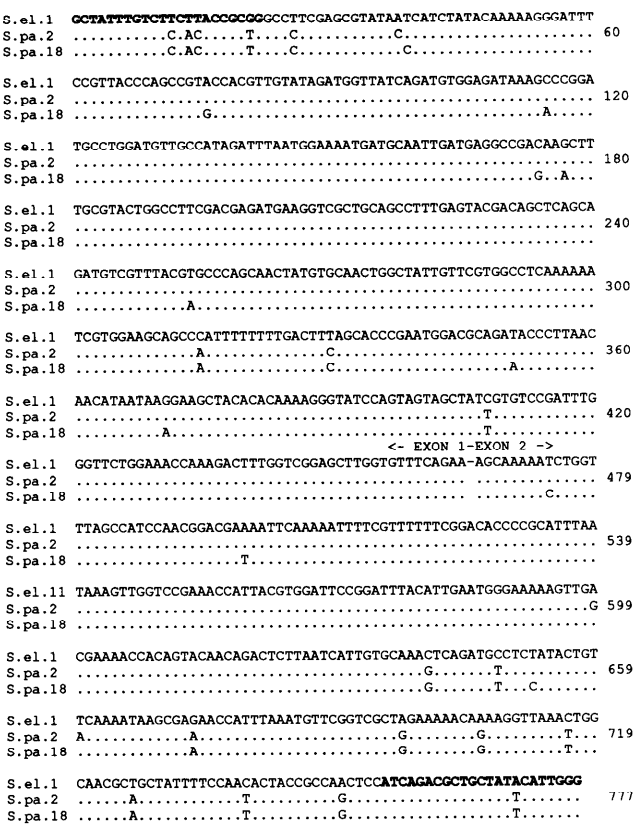

FIG. 4.—Nucleotide sequence comparison of a portion of the *P* element from *Scaptomyza elmoi* (S.el.1) and of those from *S. pallida* (S.pa.2 and S.pa.18). This region from *S. elmoi* was amplified from genomic DNA by using PCR and was sequenced. Dots (.) indicate nucleotide identity in the other sequences compared with S.el.1, while nucleotide substitutions at a particular site are indicated. The boundary of exons 1 and 2 is shown. The primer sites are indicated by boldface type.

D. guanche of the obscura species group form a clade, as do those from *D. willistoni* and *D. nebulosa* of the willistoni species group, and as do those from the two *Scaptomyza* species. This is consistent with at least phylogenetically short-term vertical transmission of *P* elements in these lineages.

However, other aspects of the *P*-element tree are incongruent with the species phylogeny. One explanation for this inconsistency is that either one or both of the trees were reconstructed improperly. Although there is some uncertainty as to the divergence times given for the species phylogeny (fig. 1), there is strong support, from studies using many different methodologies (see table 1), for the branching patterns as depicted. With respect to the *P*-element trees, the nucleotide sequence comparisons presented in figure 3 are highly consistent, irrespective of the type of phylogenetic analysis used. In addition, the bootstrap values for both the parsimony and neighbor-joining trees are quite high. The weakest feature of the *P*-element trees is that the *Lucilia cuprina* sequence may be too divergent for use as an outgroup, resulting in an improper rooting of the drosophilid *P*-element tree. However, there is no rooting of the *P*-element tree that could bring it into congruence with the species tree. Thus, the disagreement between the *P*-element and species trees would appear to be a true conflict and not an artifact due to flawed reconstruction.

There are two obvious explanations for the incongruence between the species tree and the *P*-element tree: (i) the *P* elements may have been transferred horizontally from one lineage to another at various times, or (ii) *P* elements may have diverged within a species lineage and then coexisted within descendant lineages as they were divided by speciation events. If we are willing to assume either widespread horizontal transfers or an extended coexistence of different *P*-element copies with appropriate patterns of extinction or sampling, then either horizontal transfer or vertical transmission-coexistence can explain any disagreement between a gene tree and a species tree. To decide which is more likely, it is necessary to consider what each explanation demands.

An explanation invoking horizontal transfer requires both the physical transfer of the *P* element from one fly to another and the successful integration of the element into the genome. It seems likely that the physical transfer would be an extremely rare event, although there is evidence that viruses (Miller and Miller 1982; Friesen and Nissen 1990) or parasitic mites (Houck et al. 1991) could be involved in the horizontal transfer of transposable elements. Although interspecific transfer of *P* elements from *D. melanogaster* to distantly related Diptera is not always efficient (O'Brochta and Handler 1988), transfer to *D. simulans* (Scavarda and Hartl 1984) and *D. hawaiiensis* (Brennan et al. 1984) has been achieved. In both instances, the *D. melanogaster P* element is integrated into the recipient genome after laboratory transfer. Furthermore, *P* elements from both *S. pallida* (Simonelig and Anxolabéhère 1991) and *D. willistoni* (Daniels et al. 1990) are active in *D. melanogaster* after transfer in the laboratory.

A minimum of three horizontal transfer events would be required to explain the discrepancy between the *P*-element and species trees. Four possible alternative scenarios are outlined in figure 5. In each, there is both a transfer between *D. willistoni* and *D. melanogaster* and an independent transfer between the ancestors of *Scaptomyza* and *D. bifasciata.* A third transfer, which occurs deeper in the phylogeny, is of ambiguous position. With the exception of the *D. willistoni–D. melanogaster* transfer, which has apparently occurred relatively recently, it is difficult to unequivocally identify when, and

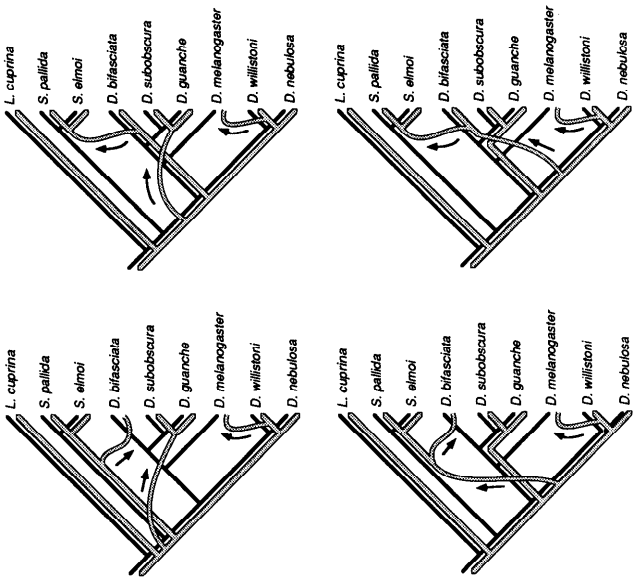

FIG. 5.—Four possible scenarios, involving horizontal transfer, to explain the incongruence between *P*-element phylogeny and species phylogeny. The black trees underneath represent the species phylogeny; and the gray trees above show the hypothesized descent of the *P* elements. In each scenario, there are three independent horizontal transfer events (highlighted by arrows), which are consistent with the observed sequence data.

in what direction, any of the other horizontal transfers occurred. For example, solely on the basis of the comparisons presented here, it is not possible to distinguish whether there was a transfer from *D. bifasciata* to *Scaptomyza* or, as previously proposed (Hagemann et al. 1992), from *Scaptomyza* to *D. bifasciata.*

The alternative explanation supposes that *P* elements diverged within species lineages and that multiple copies were passed down vertically through speciation events. Eventually, only one copy survived (or was sampled) in each species, and, because the sequences compared were, in a sense, paralogous, the *P*-element tree does not match the species tree (Goodman et al. 1979). This explanation (which we will call "vertical transmission-coexistence") requires long-term coexistence of different active *P* elements within the same lineage, as depicted in figure 6. This hypothesis requires the coexistence of at least three *P* elements in the ancestral lineage that connects the split between *Scaptomyza* and *Drosophila* to the base of the genus *Drosophila,* a period of $\geq$16 Myr (Beverley and Wilson 1984). Furthermore, at least two of these coexisting elements must have been functional, because their extant descendants in *D. willistoni, D. melanogaster,* and, apparently, *D. bifasciata* are active. At this point in our investigations, is not clear how likely such extended coexistence is. Although multiple, independently evolving *P* elements have been characterized in *D. nebulosa,* they are inactive (Lansman et al. 1987). If the fate of transposable elements is loss of function and eventual extinction (Kaplan et al. 1985), then long-term coexistence of active elements would seem unlikely.

If it is assumed, however, that long-term coexistence of active *P* elements could occur, then the explanation of vertical transmission-coexistence would require either radical changes in the rates of evolution in different *P*-element ancestors or extremely deep divergence. For example, there is a relatively small amount of divergence of the *P*-element sequences from *S. pallida* and *D. bifasciata* (table 2), and yet, under the vertical transmission-coexistence model, this divergence would have to be as old as the *Scaptomyza/Drosophila* species split (see branches marked "a"; fig. 6). In contrast, the ancestral *P*-element branches designated "b" in figure 6 show much divergence (see fig. 3*B* and *C*). If the "b" branches are indeed much shorter than the "a" branches, as drawn in figure 6, then they must have experienced a much more rapid rate of change in nucleotide sequences, to account for this divergence. On the other hand, if the rates of change were in fact equal, then the "b" branches must extend much deeper in time than depicted in figure 6, implying an extended coexistence of active elements. Radical shifts in evolutionary rates are possible, of course, especially if some elements are nonfunctional. But the *P* elements along the branches under discussion were almost certainly mobile, given that their extant descendants in *S. pallida* and *D. bifasciata* appear to be functional.

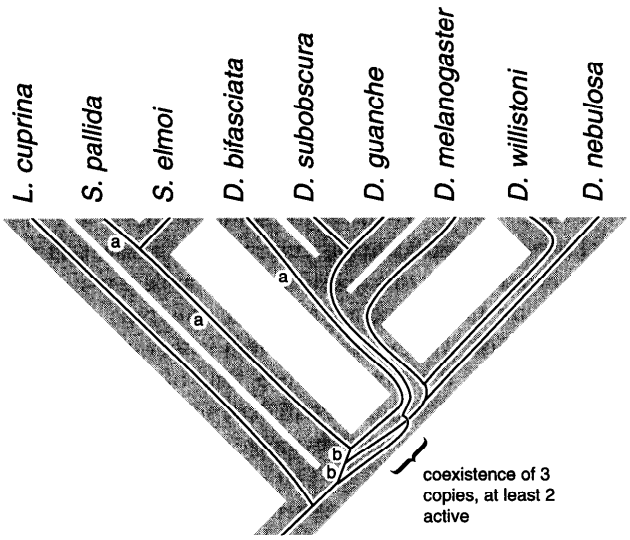

FIG. 6.—A scenario, involving vertical transmission and coexistence, to explain the incongruence between *P*-element phylogeny and species phylogeny. The broad gray outline represents the species tree; and the fine black lines represent the *P*-element tree. Branches separating the *Scaptomyza pallida* and *Drosophila bifasciata P* elements are highlighted by "a"; two other ancestral branches are designated "b."

At this point, an explanation that includes at least some horizontal transfer events seems most tenable. Such an explanation accords well with what is known about the *P* elements carried by a putative recipient of such a transfer, *D. melanogaster.* Within *D. melanogaster,* the distribution of *P* elements is patchy: most of the strains collected from the wild prior to ~40 years ago lack *P* elements, whereas these elements are present in almost all strains collected from the wild after ~1970 (Kidwell 1983). Furthermore, the *P* elements of *D. melanogaster* are remarkably homogeneous (O'Hare and Rubin 1983; Sakoyama 1985). This pattern suggests that *P* elements (or at least the *P* elements now found in *D. melanogaster*) were not present at the time that this species arose during the sophophoran radiation. The most likely candidate for a donor species is *D. willistoni,* whose *P*-element nucleotide sequence differs from that of *D. melanogaster* by only a single substitution (Daniels et al. 1990) and whose geographic range in the New World tropics overlaps that of the cosmopolitan *D. melanogaster.* This geographic overlap includes southern North America, where the earliest *P*-element activity was detected in *D. melanogaster* (Anxolabéhère et al. 1988). We have recently identified a possible vector for such a transfer: the parasitic mite, *Proctolaelaps regalis* (Houck et al. 1991), which shares with these two species the same geographic range and ecology.

The determination of *P*-element sequences from a number of additional species of *Drosophila,* which is currently underway in our laboratory, should provide an opportunity to resolve this issue and may allow us to assess the strength of the horizontal transfer hypothesis. Horizontal transfer of transposable elements may, in fact, be a more common phenomenon than has previously been believed. A number of claims for transfer of different types of transposable elements involving animals, plants, and fungi has recently been reviewed (Kidwell 1992). Whether horizontal transfer has any evolutionary significance is currently not known. However, the answer to this question would appear to depend strongly on the frequency at which this type of event occurs and on whether the problem is viewed from the perspective of the host or from that of the elements themselves. For the host, it will depend on whether mobile elements confer an advantage to their host or are merely genomic parasites. The long-term evolutionary fate of mobile-element families appears to be eventual loss of transposition function or excision and eventual divergence into anonymity (Kaplan et al. 1985). The unusual properties of these elements, which allow them to jump to new locations within a single genome, would appear to give them a distinct advantage in moving between genomes, should an appropriate opportunity be presented. Therefore, from the perspective of the elements, horizontal transfer could be critical for avoiding inevitable extinction in a particular lineage and for ensuring their long-term evolutionary survival.

Acknowledgments

This work was supported by National Science Foundation grant DEB-9119349 and by a postdoctoral fellowship, to J.B.C., from the N.S.F.-sponsored Research Training Group in the Analysis of Biological Diversification, University of Arizona. We thank an anonymous reviewer, Michael Cummings, Junhyong Kim, and members of the Phylogenetics Discussion Group for their comments and criticisms.

LITERATURE CITED

ANDERSON, C. L., E. A. CAREW, and J. R. POWELL. 1993. Evolution of the *Adh* locus in the *Drosophila willistoni* group: the loss of an intron, and shift in codon usage. Mol. Biol. Evol. **10**:605–618.

ANXOLABÉHÈRE, D., M. G. KIDWELL, and G. PERIQUET. 1988. 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.

BARRIO, E., A. LATORRE, A. MOYA, and F. J. AYALA. 1992. Phylogenetic reconstruction of the *Drosophila obscura* group, on the basis of mitochondrial DNA. Mol. Biol. Evol. **9**:621–635.

BEVERLEY, S. M., and A. C. WILSON. 1984. Molecular evolution in *Drosophila* and the higher Diptera. II. A time scale for fly evolution. J. Mol. Evol. **21**:1–13.

BOCK, I. R. 1980. Current status of the *Drosophila melanogaster* species group (Diptera). Syst. Entomol. **5**:341–356.

BRENNAN, M. D., R. G. ROWAN, and W. J. DICKINSON. 1984. Introduction of a functional P element into the germ-line of *Drosophila hawaiiensis.* Cell **38**:147–151.

BUZZATI-TRAVERSO, A. A., and R. E. SCOSSIROLLI. 1955. The "obscura" group of the genus *Drosophila.* Adv. Genet. **7**: 47–92.

CACCONE, A., J. M. GLEASON, and J. R. POWELL. 1992. Complementary DNA-DNA hybridization in *Drosophila.* J. Mol. Evol. **43**:130–140.

CAIROU, M. L., D. LACHAISE, L. TSACAS, J. SOURDIS, C. KRIMBAS, and M. ASHBURNER. 1988. New African species in the *Drosophila obscura* species group: genetic variation, differentiation and evolution. Heredity **61**:73–84.

CARSON, H. L., D. E. HARDY, H. T. SPIETH, and W. W. STONE. 1970. The evolutionary biology of Hawaiian Drosophilidae. Pp. 437–543 *in* M. K. HECHT and W. C. STEERE, eds. Essays in evolution and genetics in honor of Theodosius Dobzhansky. Appelton-Century-Crofts, New York.

DANIELS, S. B., K. R. PETERSON, L. D. STRAUSBAUGH, M. G. KIDWELL, and A. C. CHOVNICK. 1990. Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. Genetics **124**:339–355.

DANIELS, S. B., and L. D. STRAUSBAUGH. 1986. The distribution of *P*-element sequences in *Drosophila:* the *willistoni* and *saltans* species groups. J. Mol. Evol. **23**:138–148.

DESALLE, R. 1992. The phylogenetic relationships of flies in the family Drosophilidae deduced from mtDNA sequences. Mol. Phylogenet. Evol. **1**:31–40.

DOBZHANSKY, T., and J. R. POWELL. 1975. The *willistoni* group of sibling species of *Drosophila.* Pp. 589–622 *in* R. C. KING, ed. Handbook of genetics. Vol. **3:** Invertebrates of genetic interest. Plenum, New York.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17**:368–376.

———. 1991. PHYLIP: Phylogenetic inference package, version 3.4. University of Washington, Seattle.

FRIESEN, P. D., and M. S. NISSEN. 1990. Gene organization and transcription of TED, a lepidopteran retrotransposon integrated within the baculovirus genome. Mol. Cell Biol. **10**:3067–3077.

GOOD, A. G., G. MEISTER, H. BROCK, T. GRIGLIATTI, and D. HICKEY. 1989. Rapid spread of transposable elements in experimental populations of *Drosophila melanogaster.* Genetics **122**:387–396.

GOODMAN, M., J. CZELUSNIAK, G. W. MOORE, A. E. ROMERO-HERRERA, and G. MATSUDA. 1979. Fitting the gene lineage into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences. Syst. Zool. **28**:132–163.

GONZÁLEZ, A. M., M. HERNÁNDEZ, A. VOLZ, J. PESTANO, J. M. LARRUGA, D. SPERLICH, and V. M. CABRERA. 1990. Mitochondrial DNA evolution in the *obscura* species subgroup of *Drosophila.* J. Mol. Evol. **31**:122–131.

GRIMALDI, D. 1987. Amber fossil Drosophilidae (Diptera) with particular reference to the Hispaniolan taxa. Am. Museum Novitates **2880**:12–23.

———. 1990. A phylogenetic, revised classification of genera in the Drosophilidae (Diptera). Bull. 197. American Museum of Natural History, New York.

HAGEMANN, S., W. J. MILLER, and W. PINSKER. 1992. Identification of a complete P-element in the genome of *Drosophila bifasciata.* Nucleic Acids Res. **20**:409–413.

HIGGINS, D. G., A. J. BLEASBY, and R. FUCHS. 1992. CLUSTAL V: improved software for multiple sequence alignment. Comput. Appl. Biosci. **8**:189–191.

HOUCK, M. A., J. B. CLARK, K. R. PETERSON, and M. G. KIDWELL. 1991. Possible horizontal transfer of *Drosophila* genes by the mite *Proctolaelaps regalis.* Science **253**:1125–1129.

KAPLAN, N., T. DARDEN, and C. LANGLEY. 1985. Evolution and extinction of transposable elements in Mendelian populations. Genetics **10**:459–480.

KIDWELL, M. G. 1983. Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **80**:1655–1659.

———. 1992. Horizontal transfer. Curr. Opin. Genet. Dev. **2**:868–873.

KIDWELL, M. G., J. F. KIDWELL, and J. A. SVED. 1977. Hybrid dysgenesis in *Drosophila melanogaster:* a syndrome of aberrant traits including mutation, sterility, and male recombination. Genetics **36**:813–833.

KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitution through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

LAKOVAARA, S., and A. SAURA. 1982. Evolution and speciation in the *Drosophila obscura* group. Pp. 1–59 *in* M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. The genetics and biology of *Drosophila.* Vol. **3b.** Academic Press, New York.

LANSMAN, R. A., R. O. O'SHADE, T. A. GIRIGIATTI, and H. W. BROCK. 1987. Evolution of *P* transposable elements: sequences of *Drosophila nebulosa P* elements. Proc. Natl. Acad. Sci. USA **84**:6491–6495.

LOUKAS, M., C. DELIDAKIS, and F. C. KAFATOS. 1986. Genomic blot hybridization as a tool of phylogenetic analysis: evolutionary divergence in the genus *Drosophila.* J. Mol. Evol. **24**:174–188.

LOUKAS, M., C. B. KRIMBAS, and Y. VERGINI. 1984. Evolution of the *obscura* group *Drosophila* species. II. Phylogeny of ten species based on electrophoretic data. Heredity **53**:483–493.

MCALPINE, J. F. 1989. Phylogeny and classification of the Muscomorpha. Pp. 1397–1518 *in* J. F. MCALPINE, ed. Manual of Nearctic Diptera. Monograph 32. Agriculture Canada, Ottawa.

MADDISON, W. P., and D. R. MADDISON. 1992. MacClade: analysis of phylogeny and character evolution. Version 3.01. Sinauer, Sunderland, Mass.

MILLER, D. W., and L. K. MILLER. 1982. A virus mutant with an insertion of a *copia*-like transposable element. Nature **299**:562–564.

MILLER, W. J., S. HAGEMANN, E. REITER, and W. PINSKER. 1992. *P*-element homologous sequences are tandemly repeated in the genome of *Drosophila guanche.* Proc. Natl. Acad. Sci. USA **89**:4018–4022.

MISRA, S., and D. C. RIO. 1990. Cytotype control of *Drosophila P* element transposition: the 66 kd protein is a repressor of transcription activity. Cell **62**:269–284.

O'BROCHTA, D. A., and A. M. HANDLER. 1988. Mobility of *P* elements in drosophilids and nondrosophilids. Proc. Natl. Acad. Sci. USA **85**:6052–6056.

O'HARE, K., and G. M. RUBIN. 1983. Structures of *P* transposable elements and their insertion and excision in the *Drosophila melanogaster* genome. Cell **34**:25–35.

PARICIO, N., M. PEREZ-ALONSO, M. J. MARTÍNEZ-SEBASTIÁN, and R. DE FRUTOS. 1992. *P* sequences of *Drosophila subobscura* lack exon 3 and may encode a 66 kd repressor-like protein. Nucleic Acids Res. **19**:6713–6718.

PÉLANDAKIS, M., D. G. HIGGINS, and M. SOLIGNAC. 1991. Molecular phylogeny of the subgenus *Sophophora* of *Drosophila* derived from large subunit of ribosomal RNA sequences. Genetica **84**:87–94.

PERKINS, H. D., and A. J. HOWELLS. 1992. Genomic sequences with homology to the *P* element of *Drosophila melanogaster* occur in the blowfly *Lucilia cuprina.* Proc. Natl. Acad. Sci. USA **89**:10753–10757.

RIO, D. C., F. A. LASKI, and G. M. RUBIN. 1986. Identification and immunochemical analysis of biologically active *Drosophila P* element transposase. Cell **44**:21–32.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SAKOYAMA, Y., T. TODO, S. ISHIWA-CHIGUSA, T. HONJO, and S. KANDO. 1985. Structures of defective *P* transposable elements prevalent in natural Q and Q-derived strains of *D. melanogaster.* Proc. Natl. Acad. Sci. USA **82**:6236–6239.

SCAVARDA, N. J., and D. L. HARTL. 1984. Interspecific DNA transformation in *Drosophila.* Proc. Natl. Acad. Sci. USA **81**:7515–7519.

SIMONELIG, M., and A. ANXOLABÉHÈRE. 1991. A *P* element of *Scaptomyza pallida* is active in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **88**:6102–6106.

STURTEVANT, A. H. 1942. The classification of the genus *Drosophila* with description of nine new species. Univ. Tex. Publ. **4213**:5–51.

SWOFFORD, D. 1990. PAUP: phylogenetic analysis using parsimony, version 3.0s. Illinois Natural History Survey, Champaign.

THROCKMORTON, L. H. 1975. The phylogeny, ecology and geography of Drosophila. Pp. 421–469 *in* R. C. KING, ed. Handbook of genetics. Vol. **3**: Invertebrates of genetic interest. Plenum, New York.

ZWIEBEL, L. J., V. H. COHN, D. R. WRIGHT, and G. P. MOORE. 1982. Evolution of single-copy DNA and the ADH gene in seven drosophilids. J. Mol. Evol. **19**:62–71.

MASATOSHI NEI, reviewing editor

Received March 22, 1993

Accepted September 15, 1993