Molecular Evolution of *P* Transposable Elements in the Genus *Drosophila*. III. The *melanogaster* Species Group

Jonathan B. Clark, Pyong C. Kim, and Margaret G. Kidwell

Department of Ecology and Evolutionary Biology, University of Arizona

Phylogenetic relationships were determined for 76 partial *P*-element sequences from 14 species of the *melanogaster* species group within the *Drosophila* subgenus *Sophophora*. These results are examined in the context of the phylogeny of the species from which the sequences were isolated. Sequences from the *P*-element family fall into distinct subfamilies, or clades, which are often characteristic for particular species subgroups. When examined locally among closely related species, the evolution of *P* elements is characterized by vertical transmission, whereby the *P*-element phylogeny traces the species phylogeny. On a broader scale, however, the *P*-element phylogeny is not congruent with the species phylogeny. One feature of *P*-element evolution in the *melanogaster* group is the presence of more than one *P*-element subfamily, differing by as much as 36%, in the genomes of some species. Thus, *P* elements from several individual species are not monophyletic, and a likely explanation for the incongruence between *P*-element and species phylogenies is provided by the comparison of paralogous sequences. In certain instances, horizontal transfer seems to be a valid alternative explanation for lack of congruence between species and *P*-element phylogenies. The canonical *P*-element subfamily, which represents the active, autonomous transposable element, is restricted to *D. melanogaster*. Thus, its origin clearly lies outside of the *melanogaster* species group, consistent with the earlier conclusion of recent horizontal transfer.

Introduction

P elements comprise a family of mobile genetic sequences that were first described in *Drosophila melanogaster*, where their mobility is responsible for the phenomenon of hybrid dysgenesis (Kidwell, Kidwell, and Sved 1977). Since their molecular characterization (Bingham, Kidwell, and Rubin 1982; Rubin, Kidwell, and Bingham 1982), *P* elements have become invaluable as vectors for germ line transformation in *D. melanogaster* (Spradling and Rubin 1982). The manifestation of hybrid dysgenesis and the efficiency of transformation are both a result of the high frequency of transposition exhibited by *P* elements in susceptible genetic backgrounds. A detailed review of the genetics and molecular biology of *P* elements in *Drosophila* can be found in Engels (1996).

The identification of *P* elements in *D. melanogaster* notwithstanding, *P* elements have not been detected in all species of the genus (Daniels et al. 1990). They are fairly common in the subgenus *Sophophora*, which includes the *melanogaster*, *obscura*, *saltans*, and *willistoni* species groups. However, *P* elements are conspicuously absent in the *melanogaster* subgroup to which *D. melanogaster* belongs (Brookfield, Montgomery, and Langley 1984). Furthermore, not all strains of *D. melanogaster* possess detectable *P* elements (Anxolabéhère, Kidwell, and Periquet 1988). The temporal and geographic patterns of this absence led to the recent invasion hypothesis, which maintains that *P* elements have only recently been introduced into the *D. melanogaster* genome (Kidwell 1983). This was confirmed by subsequent studies which identified the source of the canonical *P* element as horizontal transfer from *D. willistoni* (Clark, Maddison, and Kidwell 1994; Daniels et al. 1990).

In an attempt to understand the evolution of *P* elements in more detail and to assess the extent of horizontal transfer, we initiated a comprehensive phylogenetic survey of *P* elements in the subgenus *Sophophora*. Results from the *saltans*, *willistoni* (Clark et al. 1995), and *obscura* (unpublished data) species groups suggest that the evolution of *P* is dominated by vertical transfer, with occasional instances of horizontal transfer. However, even when obvious cases of horizontal transfer are excluded, there is not strict congruence between *P*-element and species phylogenies. This report addresses *P*-element evolution in the *melanogaster* species group, an Old World lineage of *Sophophora* and the largest and most diverse within the subgenus. A total of 76 partial *P*-element sequences were obtained from 14 species in this group, including additional sequences from *D. melanogaster*. On a broad scale, *P*-element and species phylogenies within the *melanogaster* species group are not congruent. Furthermore, *P* elements from as many as three different subfamilies, differing by up to 36%, are found in the genomes of some species. Multiple subfamilies are also seen in the other species groups (Clark et al. 1995; unpublished data), indicating that this is a common feature of *P*-element evolution in *Sophophora*.

Materials and Methods

DNA Amplification and Sequencing

Most of the flies used in this analysis were obtained from the National Drosophila Species Resource Center (Bowling Green State University, Bowling Green, Ohio). *Drosophila rufa* and *D. lacteicornis* were obtained from Dr. M. Watada, Ehime University, Japan. *Drosophila melanogaster* Harwich is a standard laboratory P strain. The identities of the species used in this study are given in table 1. The following *P*-element sequences were obtained from the literature: *D. melano-*

Key words: *Drosophila*, molecular evolution, transposable element, mobile DNA, horizontal transfer.

Address for correspondence and reprints: Margaret G. Kidwell, The University of Arizona, Department of Ecology and Evolutionary Biology, 310 BioSciences West, Tucson, Arizona 85721. E-mail: kidwell@azstarnet.com.

Mol. Biol. Evol. 15(6):746–755. 1998

Table 1
List of *Drosophila* Species Surveyed and Their Geographical Distributions and Sources

Subgroup	Species	Distribution[a]	Source	*P* Element[b]	
				Southern	PCR
ananassae	*D. ananassae*	Circumtropical	BG 14024-0371.0	−	−
	D. bipectinata	East Palearctic; oriental; Australasian; Afrotropical	BG 14024-0381.0	F	+
	D. malerkotliana	Oriental; Afrotropical; Neotropical	BG 14024-0391.0	F	+
elegans	*D. elegans*	Oriental; Australasian	BG 14027-0461.0	−	−
eugracilis	*D. eugracilis*	Oriental; Australasian	BG 14026-0451.0	−	−
ficusphila.	*D. ficusphila*	Oriental; east Palearctic	BG 14025-0441.0	+	+
melanogaster	*D. mauritiana*	Afrotropical	BG 14021-0241.0	−	−
	D. melanogaster	Cosmopolitan	Harwich	+	+
	D. simulans	Cosmopolitan	BG 14021-0251.[c]	−	−
	D. yakuba	Afrotropical	BG 14021-0261.0	−	−
montium	*D. auraria*	East Palearctic; oriental	BG 14028-0471.0	+	+
	D. biauraria	East Palearctic	BG 14028-0501.0	+	+
	D. lacteicornis	Oriental (Okinawa)	M. Watada	+	+
	D. nikananu	Afrotropical	BG 14028-0601.0	+	+
	D. quadraria	Oriental (Taiwan)	BG 14028-0651.0	+	+
	D. rufa	Oriental; east Palearctic	M. Watada	+	+
	D. triauraria	East Palearctic; oriental	BG 14028-0691.0	+	+
	D. tsacasi	Afrotropical	BG 14028-0701.0	+	+
suzukii	*D. lucipennis*	Oriental	BG 14023-0331.0	+	+
	D. mimetica	Oriental	BG 14023-0341.0	+	+
	D. pulchrella	East Palearctic; oriental	BG 14023-0351.0	+	−
takahashii	*D. lutescens*	East Palearctic	BG 14022-0271.0	F	−
	D. prostipennis	Oriental	BG 14022-0291.0	F	−
	D. takahashii	East Palearctic; oriental; Australasian	BG 14022-0311.0	−	−

[a] Information on distribution obtained from Lemeunier et al. (1986).

[b] Presence (+) or absence (−) of *P*-element sequences as determined by Southern blotting (Daniels et al. 1990) or PCR (this study). F = faint signal. Sequences were obtained from those species designated PCR (+).

[c] Multiple strains surveyed.

gaster π25.1 (O'Hare and Rubin 1983), *D. guanche* G1 (Miller et al. 1992), *D. nebulosa* N10 (Lansman et al. 1985), *D. subobscura* A1 and G1 (Paricio et al. 1991), *D. willistoni* 13 (Daniels et al. 1990), *Scaptomyza pallida* 2 and 18 (Simonelig and Anxolabéhère 1991), and *Lucilia cuprina* P1 (Perkins and Howells 1992).

Total genomic DNA was isolated from about 100 individuals of each species using standard methods described elsewhere (Daniels and Strausbaugh 1986). Two degenerate oligonucleotide primers were designed to conserved regions within exon 2 of the *P* element (see fig. 1). Primers 2015 (5′-TGGTTTASCCATCCWRCRGAYG-3′) and 2017 (5′-CCWTCMAGGGAWGCATTRTTSAC-3′) amplify a phylogenetically informative region of about 530 bp between *D. melanogaster P*-element positions 1305 and 1780 (O'Hare and Rubin 1983). This region was chosen because phylogenetic results from a small number of complete *P*-element sequences (Clark, Maddison, and Kidwell 1994) are identical to those using this region.

Amplification reactions were usually carried out in 25-μl volumes, with 20 ng template DNA and 0.125 U of *Taq* polymerase (Cetus, Emeryville, Calif.; GIBCO-BRL, Gaithersburg, Md.). The reaction conditions were template denaturation for 1 min at 94°C, primer annealing for 1 min at 50°C, and primer extension for 1 min at 72°C (with 2 s added for each cycle) for a total of 30 cycles. PCR products were purified using Microcon-100 filtration units (Amicon, Beverley, Mass.) and were then ligated into the vector pCR-script (Stratagene, La Jolla,

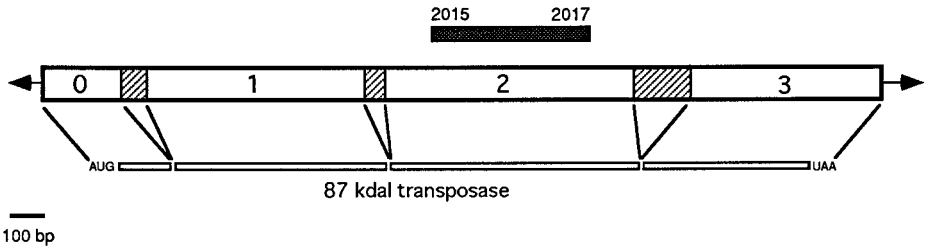

FIG. 1.—Schematic representation of the *P* element showing the locations of primers used for phylogenetic analysis. The canonical *P* element from *D. melanogaster* is 2907 bp in length and is flanked by perfect 31-bp inverted repeats (arrowheads). Exons 0–3 (open boxes) encode the transposase necessary for *P*-element mobility. The relative locations of primers 2015 and 2017, used to amplify the DNA fragment used for phylogenetic analysis, are indicated above exon 2.

Calif.) or pCR 2.1 (Invitrogen, Carlsbad, Calif.). Following transformation into *Escherichia coli,* plasmid DNA from individual clones was isolated and sequenced using Sequenase 2.0 (United States Biochemical, Cleveland, Ohio). For each species, between 4 and 10 individual clones were chosen at random as representative of the *P* elements amplified in a particular reaction. The sequence of 20% of each clone (corresponding to the middle of the PCR fragment) was obtained from both strands. One fourth of the clones were sequenced twice in their entirety, and all sequences were checked by comparing multiple clones from each species. Where discrepancies existed between clones from the same species, the differences were verified by consulting the original autoradiograms. The sequences are available from GenBank, accession numbers AF047088–AF047163.

Phylogenetic Analysis

Nucleotide sequences of the 530-bp region flanked by primers 2015 and 2017 were aligned by hand after consideration of the codon assignments. The alignment was relatively straightforward, with only occasional gaps needed in certain sequences to preserve the reading frame. The database consists of 76 sequences isolated from 14 species plus the 10 reference sequences. To coincide with other studies (Clark et al. 1995), the analysis was confined to 449 total characters, 331 of which were parsimony informative. Parsimony analysis of the aligned data matrix was performed using the heuristic option of PAUP 3.1.1 (Swofford 1993) with the random stepwise addition of sequences (10 replicates) and TBR branch swapping. Ten separate searches were performed, each terminating after the accumulation of 1,000 equally parsimonious trees. In all searches, the most-parsimonious tree was one of 1,253 steps; one of the most-parsimonious trees was arbitrarily chosen for presentation. Because divergence among some sequences exceeds 40%, the analysis was also performed on a data set in which third codon positions were eliminated. This had no effect on the major branching patterns or on the placement of the root of the tree. The sequences were also analyzed using neighbor-joining with the Kimura two-parameter correction (Saitou and Nei 1987). The resulting distance tree was identical to those obtained from parsimony analysis. Bootstrapping was performed using both parsimony (100 replicates) and neighbor-joining (500 replicates).

Results

Distribution of *P* Elements in the *melanogaster* Species Group

Eight of 11 subgroups, corresponding to the species most readily available within the *melanogaster* species group, were examined with PCR primers specific for the *P*-element sequence. As depicted in figure 1, these primers flank a 530-bp region within exon 2 of the *P*-element transposase gene. The species examined and their sources are listed in table 1. These species were chosen because, in an initial Southern blot screen, a canonical *P*-element probe from *D. melanogaster* hybridized to their genomic DNA (Daniels et al. 1990). That study examined a total of 50 species in the *melanogaster* species group, over half of which were from the largest subgroup, *montium.* In addition, a number of species that showed no hybridization were reexamined using PCR to verify that they did indeed lack detectable *P*-element sequences.

In general, the results from Southern blot and PCR analyses are in good agreement. With the exception of *D. pulchrella* of the *suzukii* species group, *P*-element sequences were obtained for all species showing hybridization in the earlier study (denoted by "+" in table 1). Sequences were not obtained from any species that earlier showed no hybridization (denoted by "−" in table 1). Varying results were obtained for four species exhibiting a faint hybridization signal in the Southern screen (denoted by "F" in table 1). *P*-element sequences were not obtained from *D. lutescens* or *D. prostipennis* of the *takahashii* subgroup (a third species sampled from this subgroup, *D. takahashii,* showed no hybridization). However, sequences were obtained from *D. bipectinata* and *D. malerkotliana* of the *ananassae* subgroup.

Consistent with the findings of Brookfield, Montgomery, and Langley (1984) and Daniels et al. (1990) with the exception of *D. melanogaster,* detectable *P* elements were absent in all flies examined from the *melanogaster* subgroup. Detectable *P* elements were also absent in *elegans* and *eugracilis,* two small subgroups represented by a total of four species Lemeunier et al. 1986). In a third small subgroup, *ficusphila, P*-element sequences were obtained from the lone species sampled. *P*-element sequences were amplified from all eight species examined in the *montium* subgroup, by far the largest subgroup in this species group. In all, *P*-element sequences were detected by PCR in 14 of 24 species examined from the *melanogaster* species group (see table 1).

One possible explanation for the failure to amplify *P*-element sequences in some species is sequence degeneration, including deletions at priming sites. Table 2 shows that 31 of 76 (41%) of the sequences in this study have deletions or insertions. For two species, *D. bipectinata* and *D. malerkotliana,* both of the *ananassae* subgroup, all sequences analyzed have deletions. In *D. malerkotliana,* there are three different deletion patterns; in addition, three sequences from *D. malerkotliana* share identical 83-bp insertions in this region. Deletions can be used as markers for particular sequences isolated from a genome. Two sequences which share a deletion and are also identical in sequence almost certainly represent the same genomic *P* element amplified with PCR and then cloned more than once and sequenced.

Although the information on the distribution of indels among these *P*-element sequences is useful, the length of the PCR product surveyed represents only about 15% of the complete *P*-element sequence. There may be additional indels in portions of the sequence that were not sampled here. It seems reasonable to assume that sequences with indels that are not multiples of three do not correspond to functional transposable elements. However, in some species, *P*-element sequences with large deletions are transcribed, although they are not capable of

Table 2
Insertions and Deletions Among *P*-Element Sequences from the *melanogaster* Species Group

Subgroup	Element (indel)	Clade
ananassae . .	*D. bipectinata* 5 (−12)	*E*
	D. bipectinata 7 (−12)	*E*
	D. bipectinata 9 (−12)	*E*
	D. bipectinata 10 (−8)	*E*
	D. bipectinata 13 (−12)	*E*
	D. malerkotliana 1 (−5) (−54)	*Q*
	D. malerkotliana 1N (−39)	*Q*
	D. malerkotliana 4 (−5) (−54)	*Q*
	D. malerkotliana 4N (+83) (−35) (−3) (−3)	*G*
	D. malerkotliana 5N (+83) (−35) (−3) (−3)	*G*
	D. malerkotliana 6 (−5) (−54)	*Q*
	D. malerkotliana 7 (−5) (−54)	*Q*
	D. malerkotliana 8 (−5) (−54)	*Q*
	D. malerkotliana 10 (−5) (−54)	*Q*
	D. malerkotliana 11N (+83) (−35) (−3) (−3)	*G*
montium. . . .	*D. auraria* 1 (−96)	*O*
	D. auraria 2 (−96)	*O*
	D. auraria 4 (−96)	*O*
	D. auraria 6 (−96)	*O*
	D. biauraria 4 (−3) (−3)	*O*
	D. biauraria 9 (−3) (−3)	*O*
	D. biauraria 15 (−3) (−3)	*O*
	D. quadraria 5 (−219)	*L*
	D. lacteicornis 11 (−36)	*L*
	D. lacteicornis 15 (−13)	*P*
	D. rufa 1 (−36)	*L*
	D. rufa 6 (−3) (−3) (−20) (−20)	*P*
	D. rufa 9 (−36)	*L*
	D. rufa 10 (−3) (−3) (−20) (−20)	*P*
	D. rufa 16 (−3) (−3) (−20) (−20)	*P*
	D. rufa 17 (−3) (−3) (−20) (−20)	*P*

NOTE.—Size and nature of insertion (+) or deletion (−) are given in parentheses. Noncontiguous indels are denoted with separate parentheses. Letters correspond to clades designated in figure 2.

transposition (Paricio et al. 1991; Nouaud and Anxolabéhère 1997). It should be stressed that the results of the phylogenetic analysis are presented here without regard to potential functionality of the sequences.

P-Element Phylogeny

Phylogenetic relationships among the 76 *P*-element sequences from the *melanogaster* species group were examined using parsimony analysis. Included in the analysis were 10 *P*-element sequences obtained from the literature, which provide a frame of reference for comparison to other studies of *P*-element phylogeny (Clark et al. 1995). A tree representative of 1,000 most-parsimonious reconstructions is shown in figure 2, with major clades designated *A–Q* and the reference sequences boxed. Clade *B* includes three reference sequences, two from *S. pallida* and one from *D. bifasciata* of the *obscura* species group. These are discussed in more detail elsewhere (Clark, Maddison, and Kidwell 1994). Clade *H* comprises reference sequences representative of the *obscura* species group, which, together with the *melanogaster* species group, represents the Old World lineage of *Sophophora.*

With the exception of the *P* elements from *D. melanogaster (*clade *A*), sequences from the *melanogaster* species group are monophyletic (clade *C*). However, *P* elements fall into major subfamilies, discussed in order starting at the top of the tree. Clade *M* includes two distinct lineages. The maximum sequence divergence between these two lineages (clades *N* and *Q*) is 31%, indicating that they are in fact quite distinct in spite of their adjacent positions in clade *M*. Clade *N* comprises sequences from four species of the *auraria* complex of the *montium* subgroup. Here, there is clear distinction between sequences from two closely related species, *D. auraria* and *D. biauraria* (clade *O*), and those from *D. lacteicornis* and *D. rufa* (clade *P*). Clade *Q* includes six identical sequences from *D. malerkotliana* of the *ananassae* subgroup. A seventh sequence in clade *Q,* and the most distinct, *D. malerkotliana* 1N, differs from the other six sequences by 5.3% and lacks 54-bp and 5-bp deletions; however, it possesses a unique 39-bp deletion (see table 2). Thus, this sequence shares a common ancestor with the other six sequences, but has accumulated substitutions and a distinct deletion since its divergence from the other sequences.

Clade *J* comprises sequences from six species in the *montium* subgroup. There are two lineages within this subfamily, consistent with geographical distribution. The first, clade *L,* includes sequences from four closely related species of the *auraria* complex, which are distributed in the oriental and east Palearctic regions. The most divergent sequence in clade *L* is *D. quadraria* 5, which has a unique 219-bp deletion (see table 2). Interestingly, *D. lacteicornis* 11 and *D. rufa* 1 and 9 share the same 36-bp deletion (see table 2), although there is a clear distinction between sequences from the two species (3.7%). This suggests that the deleted element was inherited from a common ancestor and then subsequently diverged in the separate species. The second lineage, clade *K,* includes sequences from two afrotropical species of the *montium* subgroup, *D. nikananu* and *D. tsacasi.* Within clade *K,* there is a clear distinction between sequences of *D. nikananu* (*nikananu* complex) and *D. tsacasi* (*bakoue* complex), consistent with the species phylogeny. The 12% divergence among sequences from these two complexes represents the differentiation of the sequences since the evolutionary split of the two African complexes.

Recent work by Nouaud and Anxolabéhère (1997) provides a detailed understanding of the sequences from *D. tsacasi.* They characterized a genomic *P*-element clone that is identical in sequence to those obtained from *D. tsacasi* in the present (PCR) survey (clade *K*). The genomic clone is a truncated *P* element that lacks the fourth exon and, because both a complete transposase and terminal inverted repeats are necessary for *P*-element mobility, this truncated sequence is not a functional transposable element. However, some of the important domains of the transposase protein are preserved in the truncated sequence, and, indeed, this sequence is transcribed in *D. tsacasi* (Nouaud and Anxolabéhère 1997). One explanation for its expression is that the truncated sequence has been recruited by *D. tsacasi* and may have evolved a new (and as yet unknown) function unrelated to transposition. The truncated sequence may also function as a repressor of *P*-element transposition,

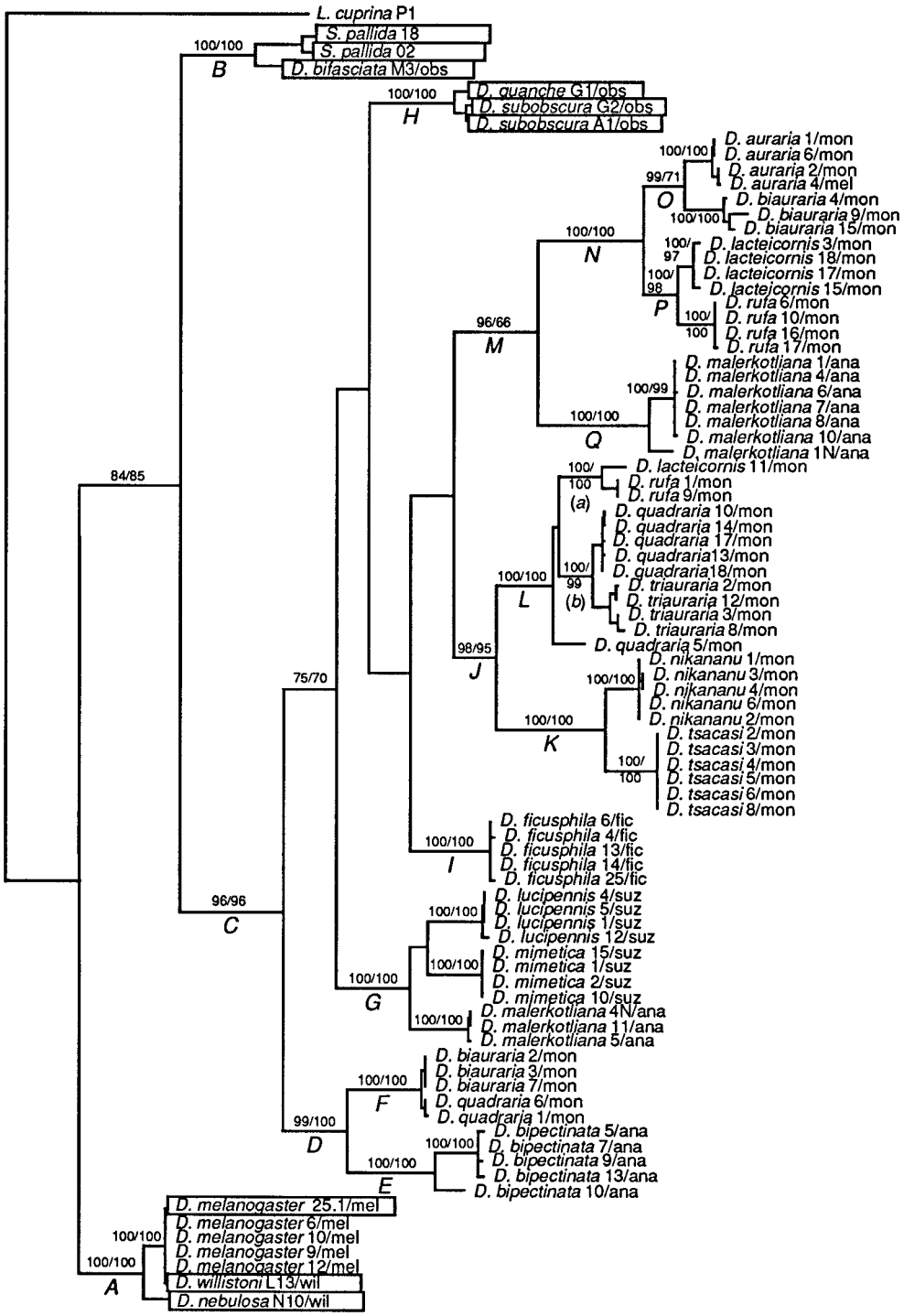

FIG. 2.—Phylogenetic analysis of *P*-element nucleotide sequences from the *melanogaster* species group. Comparisons were limited to 449 bp between primers 2015 and 2017, and the sequence from *L. cuprina* was used as an outgroup. This cladogram was generated by parsimony analysis as implemented by PAUP 3.1.1 (Swofford 1993) using the heuristic search algorithm with TBR branch swapping and random stepwise addition of taxa. This is an arbitrarily chosen representative of 1,000 equally parsimonious trees, each requiring 1,253 steps. The consistency index is 0.551, and the retention index is 0.911. Numbers above the branches are bootstrap percentages; those before the slashes are values for parsimony analysis (100 replicates); and those after the slashes are values for neighbor-joining (500 replicates). Only values of at least 50% are shown. Letters refer to clades that are discussed in the text. The *a* and *b* labels in clade *L* are also used in the discussion. Species names are given in italics followed by a numbered clone designation. Ten reference sequences are boxed, and species group designations are given after the slashes (obs, *obscura* species group; wil, *willistoni* species group). For species from the *melanogaster* species group, subgroup affiliations are given after the slashes (ana, *ananassae*; fic, *ficusphila*; mel, *melanogaster*; mon, *montium*; suz, *suzukii*).

perhaps serving to mitigate the deleterious effects of transposition of full-length active elements that exist elsewhere in the genome. The truncated sequence is present in eight other members of the *montium* subgroup; however, none of these other species was included in the survey described here.

Clade *I*, another subfamily of *P* elements from the *melanogaster* species group, comprises five nearly identical sequences from *D. ficusphila* of the *ficusphila* subgroup. This species is the only geographically widespread member of this small subgroup, and the distinction of its *P* elements suggests that it is indeed a separate subgroup within the *melanogaster* species group. Clade *G* includes *P*-element sequences from two species of the *suzukii* subgroup, *D. lucipennis* and *D. mimetica*, and from *D. malekotliana* of the *ananassae* subgroup. Sequences from the two species of the *suzukii* subgroup differ from each other by an average of 11%, while sequences from *D. malerkotliana* differ from those of the *suzukii* subgroup by an average of 14%. This implies that the *suzukii* and *ananassae* subgroups are closely related; however, alternative explanations, such as horizontal transfer, are considered in the *Discussion.*

Clade *D* comprises sequences that were isolated from the *ananassae* and *montium* subgroups. *P* elements from the *auraria* complex of the *montium* subgroup are again represented here (clade *F*). Clade *E* is another monotypic clade, comprising closely related sequences from *D. bipectinata* of the *ananassae* subgroup. It is noteworthy that these sequences are distinct from those of *D. malerkotliana*, also of the *ananassae* subgroup (clades *G* and *Q*). The most divergent sequence, *D. bipectinata* 10, has a unique 8-bp deletion and does not share a 12-bp deletion with the other sequences in clade *E* (see table 2).

Clade *A* includes the final subfamily of *P* elements from the *melanogaster* species group. The phylogeny of these sequences, referred to as canonical *P* elements, is unusual in several respects. First, sequences in clade *A* clearly have a distinct evolutionary origin from the other sequences from the *melanogaster* group, which fall into clade *C*. As mentioned above, *D. melanogaster* is the only member of the *melanogaster* subgroup with detectable *P* elements, and the only member of the *melanogaster* species group that possesses the canonical element. Canonical *P* elements are common, however, among species in the New World *saltans* and *willistoni* species groups (Clark et al. 1995). The sequences of the canonical elements from *D. melanogaster* are identical over this 500-bp region to those isolated from *D. willistoni* of the *willistoni* species group. This confirms the results of an earlier study and is consistent with the conclusion that the canonical *P* element was transferred horizontally from *D. willistoni* to *D. melanogaster* (Daniels et al. 1990; Kidwell 1994).

Although good evidence exists for the short-term vertical transmission of *P* elements in the *melanogaster* species group, the *P*-element phylogeny is not congruent with that of the species themselves. For example, as mentioned above, sequences from the *ananassae* subgroup are not monophyletic, falling into three distinct subfamilies (clades *E*, *G*, and *Q*). Another example is the three distinct subfamilies of *P* elements from the *auraria* complex of the *montium* subgroup (clades *F*, *L*, and *O*). For single-copy nuclear genes, sequences from the *montium* and *ananassae* subgroups are expected to be monophyletic. The genomes of at least five species within the *melanogaster* species group possess multiple *P*-element subfamilies. These are listed in table 3, along with the maximum percentages of divergence among the *P*-element subfamilies. The presence of multiple subfamilies of *P* elements in some species indicates that in some cases, the sequences compared are not orthologous. Thus, comparison of paralogous sequences is one explanation for the lack of congruence between *P*-element phylogeny and species phylogeny.

Table 3
Maximum Divergence Among *P*-Element Sequences from the Same *Drosophila* Species

Species	% Divergence	Clades
D. biauraria	30.5	*F, O*
D. lacteicornis	33.6	*L, P*
D. malerkotliana	35.8	*G, Q*
D. quadraria	32.0	*F, L*
D. rufa	23.8	*L, P*

NOTE.—Maximum percentages of divergence (uncorrected) are shown for only those species with multiple *P*-element subfamilies. Clade designations correspond to letters in figure 2.

Discussion

Phylogeny of the *melanogaster* Species Group

It is believed that the *Drosophila* subgenus *Sophophora* originated in the Old World tropics, probably between 40 and 55 MYA (Throckmorton 1975). The two extant Old World species groups, *melanogaster* and *obscura*, are believed to have diverged approximately 30–45 MYA, and the origin of the *melanogaster* species group was probably in southeast Asia (Throckmorton 1975; Bock 1980; Ashburner, Bodmer, and Lemeunier 1984). From here, secondary radiations occurred into adjacent regions, consistent with the current distribution throughout the eastern Palearctic, oriental, Australasian, and Afrotropical regions.

Relationships among the eight extant principle subgroups are uncertain. However, evidence from a number of studies supports three major lineages, one represented by the *ananassae* subgroup, a second by the *montium* subgroup, and a third comprising the *elegans*, *eugracilis*, *ficusphila*, *melanogaster*, *suzukii*, and *takahashii* subgroups (summarized in Lemeunier et al. 1986). This view is supported by the limited number of molecular studies that have included species from the *melanogaster* group (Tsakas and Tscas 1984; Pélandakis, Higgins, and Solignac 1991; Pélandakis and Solignac 1993). A general phylogeny, synthesized from a number of morphological, biogeographical, and molecular studies, is presented in figure 3. This species phylogeny is intended to guide and inform discussion of the *P*-element phylogeny that follows.

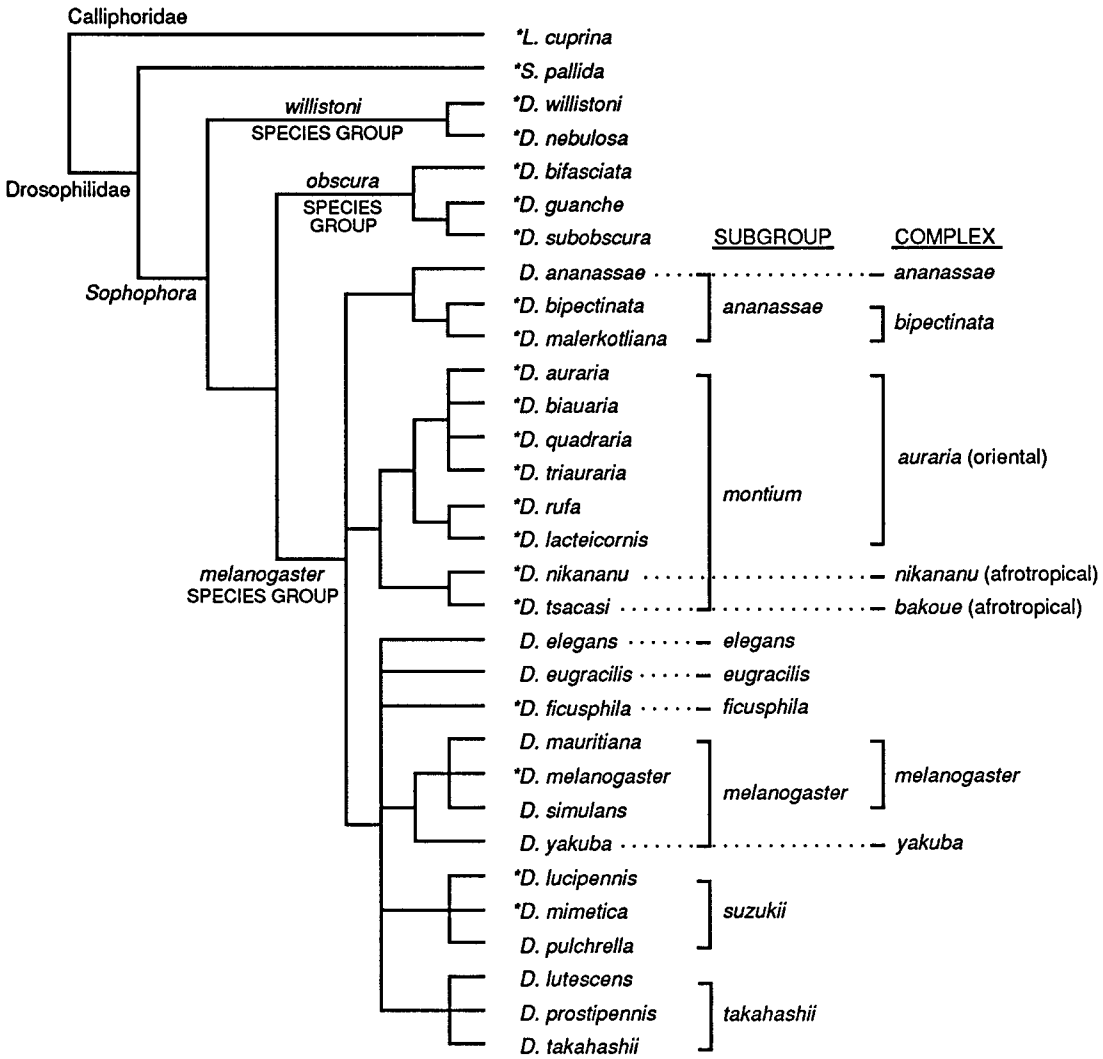

FIG. 3.—Summary of phylogenetic relationships among species in the *melanogaster* species group. Only those species used in this study are shown; other species, both inside and outside of the genus *Drosophila*, are also included. For the *melanogaster* group, subgroup and complex information is shown. Species identified with asterisks are those from which *P*-element sequences were obtained (see fig. 2). This tree is based on a number of morphological (Bock and Wheeler 1972; Bock 1980), biogeographical (Throckmorton 1975; Lemeunier et al. 1986) and molecular (Ashburner, Bodmer, and Lemeunier 1984; Beverly and Wilson 1984; Tsakas and Tscas 1984; Kim, Aotsuka, and Kitagawa 1993; Pélandakis and Solignac 1993; Russo, Takezaki, and Nei 1995) studies. Not all species in the tree were included in all of the studies.

P-Element Evolution in the *melanogaster* Species Group

On a local scale, the *P*-element phylogeny is consistent with that of the flies themselves. Even though multiple sequences were obtained from each species, sequences isolated from one species are distinct from those of other species. Clade *J*, which includes only sequences from the *montium* subgroup, provides a striking illustration of short-term vertical transmission. There is a gradient of sequence differentiation that parallels species relationships. In clade *L*, there is relatively little differentiation among sequences isolated from *D. rufa* and *D. lacteicornis* (*L*(*a*)), and among sequences isolated from *D. quadraria* and *D. triauraria* (*L*(*b*)). There is considerably more differentiation when sequences from *L*(*a*) are compared to those from *L*(*b*). The largest degree of differentiation is seen between species with Afrotropical (clade *K*, *D. nikananu* and *D. tsacasi*) and oriental (clade *L*, *auraria* complex) distributions. The two Afrotropical species are clearly differentiated from each other, which is expected for these two species belonging to different complexes (*D. nikananu*, *nikananu* complex; *D. tsacasi*, *bakoue* complex).

On a broader scale, *P*-element phylogeny and species phylogeny are not congruent. A single origin for *P* elements from a particular subgroup is expected, but the genomes of some species support more than one *P*-element subfamily. This is most apparent in figure 2 for the *montium* subgroup (clades *F*, *J*, and *N*). The divergence among these different subfamilies is as much as 32%, suggesting that they do indeed have distinct evolutionary histories. This finding of multiple *P*-element subfamilies within a species is consistent with *P*-element evolution in the *saltans*, *willistoni*, and *obscura* species groups (Clark et al. 1995; unpublished data). The origins of these different subfamilies within a species are un-

certain. They could represent different *P*-element invasions of the genome at different times, or ancient polymorphisms that have been retained in a particular lineage. *P* elements from different subfamilies are not detected in all species (e.g., *D. ficusphila,* clade *I*; *D. bipectinata,* clade *E*; *D. tsacasi* and *D. nikananu,* clade *K*). Assuming that different subfamilies were at one time present in the ancestor of these species, this could be due to the stochastic loss of a particular subfamily or the retention of only a single member of an ancestral polymorphism. Alternatively, multiple subfamilies may indeed exist in these species but were simply not detected in this PCR-based survey, either because of limited sampling or because the primers used were not successful in amplifying all divergent sequence variants that were present.

Among the more puzzling aspects of this *P*-element phylogeny are the sequences from *D. bipectinata* (clade *E*) and *D. malerkotliana* (clades *G* and *Q*), both of the *bipectinata* complex of the *ananassae* subgroup (see fig. 3). These two species are very similar morphologically and can be hybridized in the laboratory to produce fertile females and sterile males (Bock 1978). This implies that these two species diverged from one another only relatively recently, perhaps within the past few million years, yet their corresponding *P* elements are clearly distinct both in sequence and in the presence of characteristic deletions that are not shared between them (see table 2). It is interesting that some sequences from *D. malerkotliana* are affiliated with sequences from the *montium* subgroup (clade *N*), while sequences from *D. bipectinata* are affiliated with a different subfamily of *montium* sequences (clade *F*). One explanation for this pattern is that these two species have retained in their genomes different ancestral *melanogaster* group *P* elements. However, this seems unlikely, since hybridization studies indicate that these two species have only recently become reproductively isolated (Bock 1978). *P*-element evolution in the *ananassae* subgroup is further complicated by a second subfamily of sequences from *D. malerkotliana,* which are affiliated with sequences from the *suzukii* subgroup (clade *G*).

Examination of the degree of divergence between sequences from the *ananassae* subgroup and related sequences from other subgroups can be used to identify other explanations for this unusual distribution of *P* elements in *D. bipectinata* and *D. malerkotliana.* Within clade *D,* the average divergence among sequences from *D. bipectinata* (clade *E*) and those from the *montium* subgroup (clade *F*) is 19.7%. Within clade *M,* the average divergence between sequences from *D. malerkotliana* (clade *Q*) and those from various species in the *montium* subgroup (clade *N*) is 29.3%. This compares to maximum divergence of 16.6% among the 15 sequences isolated from any of the four species of the *montium* subgroup (clade *N*). This pattern of relatively modest differentiation among sequences from the same subgroup and greater differentiation among sequences from different subgroups is what would be expected for vertical transmission.

The situation in the third *ananassae P*-element subfamily is different. Within clade *G,* the average divergence between sequences from *D. malerkotliana* of the *ananassae* subgroup and either *D. lucipennis* or *D. mimetica* of the *suzukii* subgroup is 13.7%. However, the divergence of sequences between the two *suzukii* species is 11.2%. Thus, the divergence between *P* elements from the same subgroup is similar to the divergence of *P*-element sequences from different subgroups. One explanation for this pattern is that the *ananassae* and *suzukii* subgroups are in fact closely related. Although phylogeny among the various subgroups within the *melanogaster* species group is by no means resolved, these two subgroups appear to be only distantly related (see Lemeunier et al. 1986). Thus, the phylogeny of the flies provides no support for this explanation. Another explanation is that the *P* elements from *D. malerkotliana* originated in the *suzukii* subgroup and were transferred horizontally to *D. malerkotliana* at some time in the past. This raises the possibility that horizontal transfer is responsible for some aspects of the unusual distribution of *P* elements in the *ananassae* subgroup. There is considerable geographic overlap in India and southeast Asia between *D. malerkotliana* and many of the species in the *suzukii* subgroup (Lemeunier et al. 1986), providing at least an opportunity for horizontal transfer. However, based on the information we have now, it is not possible to identify the specific donor species for such transfer.

As many as 16 species examined from the *melanogaster* species group lack detectable *P*-element sequences (see also Daniels et al. 1990). There are three explanations for this observation. First, it is possible that these species never carried *P* elements. This is not likely because *P* is present throughout all four principle species groups of *Sophophora,* suggesting that it has been in this subgenus since its origin (Daniels et al. 1990). Second, the *P*-element sequences from these species may have simply diverged, in sequence or structure, beyond recognition by the methods employed. Third, *P* elements may have been lost from these genomes. A more detailed genomic characterization will be necessary to distinguish between the latter two alternatives for the *melanogaster* species group. Within the *obscura* species group, there is evidence for the second explanation in *D. pseudoobscura* (unpublished data) and *D. bifasciata* (Hagemann, Miller, and Pinsker 1990), which possess internally deleted *P*-element sequences. Although detectable in Southern blots, they cannot be amplified with the primers used here. Strong evidence for the third explanation comes from *D. insularis* of the *willistoni* species group, which possesses recognizable *P* elements but lacks the canonical *P*-element subfamily characteristic of all other members of this group (Clark et al. 1995). Theoretical studies have shown that stochastic loss from a genome may be the ultimate fate of transposable elements (Kaplan, Darden, and Langley 1985). Although the exact mechanism(s) by which such loss occurs is not known, it most probably involves random genetic drift (Kaplan, Darden, and Langley 1985; Brookfield 1995; Lohe et al. 1995).

Horizontal Transfer and the Origin of *P* Elements in *D. melanogaster*

As seen in figure 2, the canonical *P* element (clade *A*) is, with the exception of *D. melanogaster*, absent from the *melanogaster* species group. Thus, its origin must lie outside of this species group. The alternative explanation, that canonical *P* elements were lost in all species except for *D. melanogaster*, is untenable. Not only are canonical *P* elements absent from all members of the *melanogaster* group examined, but they are also absent from all members of the *obscura* species group (unpublished data). In all, 66 species from the Old World lineage of *Sophophora* have been examined and found to lack the canonical *P* element (Daniels et al. 1990; this study). The results reported here are consistent with the earlier conclusion that the canonical *P* element was transferred horizontally from *D. willistoni* of the New World *willistoni* species group to the cosmopolitan *D. melanogaster* (Daniels et al. 1990; Clark, Maddison, and Kidwell 1994). It is also possible that horizontal transfer may explain some of the other unusual findings of *P* element phylogeny in the *melanogaster* species group that were described above. However, the explanation of horizontal transfer requires that certain assumptions be met (Houck et al. 1991), such as geographical and ecological overlap between the donor and recipient species involved. Until these conditions can be satisfied for particular cases where *P*-element and species phylogenies are incongruent, alternative explanations, such as *P*-element loss and the presence of paralogous sequences, cannot be ruled out.

The present findings of predominantly vertical transmission, occasional instances of horizontal transfer, and the presence of multiple subfamilies are similar to those obtained previously from the *saltans, willistoni,* and *obscura* species groups of *Sophophora.* Although the frequency of horizontal transfer appears to be relatively high among the *saltans* and *willistoni* species groups (Clark and Kidwell 1997), sampling and ascertainment problems preclude any meaningful comparison of frequencies among different taxa. These results can also be compared with those for the *mariner* element, which, like *P,* is a class II transposable element (Finnegan 1989). *Mariner* elements also belong to distinct subfamilies, and the phylogeny of *mariner* sequences is characterized by both vertical transmission and horizontal transfer (Robertson and MacLeod 1993). In general, class II elements, which transpose by means of a DNA–DNA mechanism, appear to be more prone to horizontal transfer than class I elements that use an RNA intermediate (Kidwell 1993). This may be related to their mode of transposition or their relatively small size. It is possible that for class II elements, ecological conditions and geographical opportunities may influence this frequency. It is becoming apparent that horizontal transfer has played a role in the evolution of both *mariner* and *P* and may be responsible for the persistence of these transposable elements in natural populations (Lohe et al. 1995; Clark and Kidwell 1997). The present findings of predominantly vertical transfer with occasional instances of horizontal transfer are similar to those for the *willistoni, saltans,* and *obscura* subgroups. Although the frequency of horizontal transfer appears to be relatively high among the *willistoni* and *saltans* species groups, sampling and ascertainment problems preclude any meaningful comparison of frequencies.

Acknowledgments

This work was supported by NSF grant DEB-9119349 to M. G. Kidwell. We are grateful to Dr. M. Watada for help with identification of some of the species, for the samples of *D. lacteicornis* and *D. rufa,* and for comments on the manuscript. We thank Dr. J. S. Yoon and his staff at the National Drosophila Species Resource Center, Bowling Green State University, for providing most of the flies used in this study.

LITERATURE CITED

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.

ASHBURNER, M., M. BODMER, and F. LEMEUNIER. 1984. On the evolutionary relationships of *Drosophila melanogaster*. Dev. Genet. **4**:295–312.

BEVERLY, 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.

BINGHAM, P. M., M. G. KIDWELL, and G. M. RUBIN. 1982. The molecular basis of P-M hybrid dysgenesis: the role of the P element, a P strain-specific transposon family. Cell **29**: 995–1004.

BOCK, I. R. 1978. The *bipectinata* complex: a study in interspecific hybridization in the genus *Drosophila* (Dipt.; Drosophilidae). Aust. J. Biol. Sci. **31**:197–208.

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

BOCK, I. R., and M. R. WHEELER. 1972. The *Drosophila melanogaster* species group. Univ. Tex. Publ. **7213**:1–102.

BROOKFIELD, J. F. Y. 1995. Transposable elements as selfish DNA. Pp. 130–153 *in* D. J. SHERRATT, ed. Mobile genetic elements. IRL Press, Oxford.

BROOKFIELD, J. F. Y., E. MONTGOMERY, and C. LANGLEY. 1984. Apparent absence of transposable elements related to the *P* elements of *D. melanogaster* in other species of *Drosophila.* Nature **310**:330–332.

CALVI, B. R., T. J. HONG, S. D. FINDLEY, and W. M. GELBART. 1991. Evidence for a common evolutionary origin of inverted repeat transposons in *Drosophila* and plants: *hobo, Activator* and *Tam3.* Cell **66**:465–471.

CLARK, J. B., T. K. ALTHEIDE, M. J. SCHLOSSER, and M. G. KIDWELL. 1995. Molecular evolution of *P* transposable elements in the genus *Drosophila.* I. The *saltans* and *willistoni* species groups. Mol. Biol. Evol. **12**:902–913.

CLARK, J. B., and M. G. KIDWELL. 1997. A phylogenetic perspective on *P* transposable element evolution in *Drosophila.* Proc. Natl. Acad. Sci. USA **94**:11428–11433.

CLARK, J. B., W. P. MADDISON, and M. G. KIDWELL. 1994. Phylogenetic analysis supports horizontal transfer of P transposable elements. Mol. Biol. Evol. **11**:40–50.

DANIELS, S. B., K. R. PETERSON, L. D. STRAUSBAUGH, M. G. KIDWELL, and A. 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.

ENGELS, W. R. 1996. *P* elements in *Drosophila.* Pp. 103–123 *in* H. SAEDLER and A. GIERL, eds. Transposable elements. Springer-Verlag, Berlin.

FINNEGAN, D. J. 1989. Eukaryotic transposable elements and genome evolution. Trends Genet. **5**:103–107.

HAGEMANN, S., W. J. MILLER, and W. PINSKER. 1990. *P*-related sequences in *Drosophila bifasciata*: a molecular clue to the understanding of *P*-element evolution in the genus *Drosophila.* J. Mol. Evol. **31**:478–484.

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 **109**:459–480.

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

———. 1993. Lateral transfer in natural populations of eukaryotes. Annu. Rev. Genet. **27**:235–256.

———. 1994. The evolutionary history of the P family of transposable elements. J. Hered. **85**:339–346.

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.

KIM, B. K., T. AOTSUKA, and O. KITAGAWA. 1993. Evolutionary genetics of the *Drosophila montium* subgroup. II. Mitochondrial DNA variation. Zool. Sci. **10**:991–996.

LANSMAN, R. A., S. N. STACEY, T. A. GRIGLIATTI, and H. W. BROCK. 1985. Sequences homologous to the P mobile element of *Drosophila melanogaster* are widely distributed in the subgenus Sophophora. Nature **318**:561–563.

LEMEUNIER, F., J. R. DAVID, L. TSACAS, and M. ASHBURNER. 1986. The *melanogaster* species group. Pp. 147–256 *in* M. ASHBURNER, H. L. CARSON, and J. N. THOMPSON, eds. The genetics and biology of Drosophila. Vol. 3e. Academic Press, London.

LOHE, A. R., E. N. MORIYAMA, D. A. LIDHOLM, and D. L. HARTL. 1995. Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. Mol. Biol. Evol. **12**:62–72.

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. **89**:4018–4022.

NOUAUD, D., and D. ANXOLABÉHÈRE. 1997. *P* element domestication: a stationary truncated *P* element may encode a 66 kD repressor-like protein in the *Drosophila montium* species subgroup. Mol. Biol. Evol. **14**:1132–1144.

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

PARICIO, N. M., M. PEREZ-ALONSO, M. J. MARTÍNEZ-SEBASTIÁN, and R. DE FRUTOS. 1991. *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.

PÉLANDAKIS, M., and M. SOLIGNAC. 1993. Molecular phylogeny of *Drosophila* based on ribosomal RNA sequences. J. Mol. Evol. **37**:525–543.

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.

ROBERTSON, H. M., and D. J. LAMPE. 1995. Distribution of transposable elements in arthropods. Annu. Rev. Entomol. **40**:333–357.

ROBERTSON, H. M., and E. G. MACLEOD. 1993. Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. Insect Mol. Biol. **2**:125–139.

RUBIN, G. M., M. G. KIDWELL, and P. M. BINGHAM. 1982. The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. Cell **29**:987–994.

RUSSO, C. A., N. TAKEZAKI, and M. NEI. 1995. Molecular phylogeny and divergence times of drosophilid species. Mol. Biol. Evol. **12**:391–404.

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

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

———. 1994. P-elements are old components of the *Scaptomyza pallida* genome. J. Mol. Evol. **38**:232–240.

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

SWOFFORD, D. 1993. PAUP: phylogenetic analysis using parsimony. Version 3.1.1. Smithsonian Institution, Washington, D.C.

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

TSAKAS, S. C., and L. TSCAS. 1984. A phenetic tree of eighteen species of the *melanogaster* group of *Drosophila* using alozyme data as compared with classification based on other criteria. Genetica **64**:139–144.

THOMAS H. EICKBUSH, reviewing editor

Accepted March 4, 1998