

## Gypsy Homologous Sequences in *Drosophila subobscura* (*gypsyDS*)

T.M. Alberola and R. de Frutos (1993)

Departament de Genètica, Facultat de Ciències Biològiques, Universitat de València, Spain

**Summary.** Characterization of sequences homologous to the *Drosophila melanogaster gypsy* transposable element was carried out in *Drosophila subobscura* (*gypsyDS*). They were found to be widely distributed among natural populations of this species. From Southern blot and in situ analyses, these sequences appear to be mobile in this species. *GypsyDS* sequences are located in both euchromatic and heterochromatic regions. A complete *gypsyDS* sequence was isolated from a *D. subobscura* genomic library, and a 1.3-kb fragment which aligns with the ORF2 of the *D. melanogaster gypsy* element was sequenced. Comparisons of this sequence in three species (*D. subobscura*, *D. melanogaster*, and *D. virilis*) indicate that there is greater similarity between the *D. subobscura*–*D. virilis* sequences than between *D. subobscura* and *D. melanogaster*. Molecular divergence of *gypsy* sequences between *D. virilis* and *D. subobscura* is estimated at 16 MY, whereas the most likely divergence time of these two species is more than 60 MY. These data strongly suggest that *gypsy* sequences have been horizontally transferred between these species.

**Key words:** Transposable elements — *Drosophila* — *Gypsy* — Horizontal transfer — In situ hybridization — Molecular evolution

### Introduction

Transposable elements (TEs) are ubiquitous constituents of living organisms. They seem to be par-

ticularly important in evolution as a source of variation. Among eukaryotes, the best-characterized TEs are those of *Drosophila melanogaster*. To date, 30 different families have been described in this species (Finnegan and Fawcett 1986). All of these families are formed by active members which are able to transpose. In *D. melanogaster*, a considerable proportion of the morphologically spontaneous mutations are caused by the insertion of mobile genetic elements. (For reviews see Bingham and Zachar 1989; McDonald 1989.) Transposition may also induce major genetic instability, for example, hybrid dysgenesis (Kidwell et al. 1977; for current reviews, see Engels 1989; Finnegan 1989).

Extensive surveys have been carried out on the distribution of *D. melanogaster* TEs in *Drosophila* species in an effort to study their dispersal and evolution (Martin et al. 1983; Brookfield et al. 1984; Lansman et al. 1985; Daniels and Strausbaugh 1986; Stacey et al. 1986; Silber et al. 1989; Daniels et al. 1990a,b; de Frutos et al. 1992). *F*-homologous sequences have been detected in all species tested (Stacey et al. 1986). *Gypsy*-, *412*-, and *cop*-homologous sequences are widely distributed in the *Drosophila* genus, with discontinuities (Martin et al. 1983; Brookfield et al. 1984; Stacey et al. 1986; de Frutos et al. 1992). *P*-homologous sequences are widely distributed within the *Sophophora* subgenus, although they are absent in species most closely related to *Drosophila melanogaster* (Stacey et al. 1986; Daniels et al. 1990a). *Hobo*-, *297*-, *I*-, and *FB*-homologous sequences are almost completely restricted to the *melanogaster* group (Martin et al. 1983; Brookfield et al. 1984; Silber et al. 1989; Daniels et al. 1990b). Although the extent of distribution varies depending on the element, these data indicate that sequences homologous to some TEs of

*D. melanogaster* are present in other *Drosophila* species. However, as the majority of data is based on Southern blot analysis for the presence or absence of each homologous sequence, it is open to question whether they are active elements or just inactive residues. The set of TEs in each species is most probably composed of inactive relict junk sequences plus some functional active elements. Since functional element activity changes depending on the strain or natural population analyzed (Csink and McDonald 1990), it is necessary to determine to what extent the set of TEs is functional in a species or strain in order to determine whether TEs can be considered an evolutionary force at present. It is with this objective in mind that we have initiated a molecular analysis of *gypsy*-homologous sequences in *D. subobscura* as an initial step in the examination of actively transposing sequences in *Drosophila* species other than *D. melanogaster*. As mentioned earlier, sequences homologous to the retrotransposon *gypsy* of *D. melanogaster* are widely distributed among *Drosophila* species (Martin et al. 1983; Stacey et al. 1986; de Frutos et al. 1992). Differences were found, in terms of sequence similarity among species, using Southern analysis at different stringency levels. Only some species of the *virilis* and *obscura* groups showed positive hybridization signals under high-stringency conditions. It is not clear whether *gypsy* sequences are almost always active in *Drosophila* species, or if these widespread sequences are only the relics of functional ones. Unfortunately there is little information on this matter. Recently, a *gypsy* sequence from *D. virilis* has been characterized which seems to be transpositionally active (Mizrokhi and Mazo 1991). The authors suggest that horizontal transfer could be involved in the process of spreading and evolution of *gypsy*.

In this paper we report the molecular characterization of *gypsy* sequences of *D. subobscura* (*gypsyDS*), which seem able to transpose in this species. On comparing *D. virilis* and *D. subobscura* *gypsy* sequences, they were found to be closer to each other in terms of nucleotide sequence similarity than they were to the *D. melanogaster gypsy*. Our results suggest that horizontal transfer has been involved in the evolution of *gypsy* in these species.

## Materials and Methods

### *Drosophila subobscura* Strains

The strains of *D. subobscura* used in this study were obtained from different natural populations over the distribution area of this species. They were as follows:

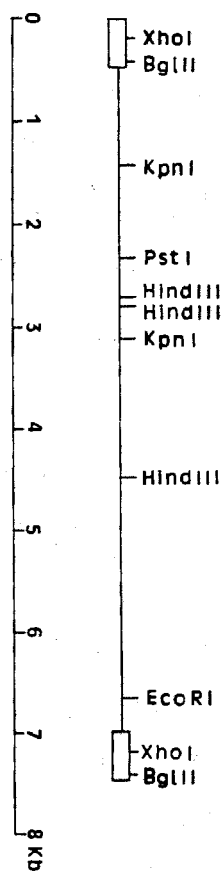


Fig. 1. Restriction map of the *D. melanogaster gypsy* element (Marlor et al. 1986)

**European strains:** H271 (Finland), SN, L, and UP (Sweden), G (Denmark), FT (Scotland), TÛ (Germany), TA (Holland), DIE, RO, and BU (Switzerland), CY (France), FP (Italy), and PM, COL, and R (Spain).

**African strains:** MAR (Morocco), BI (Tunisia), MA (Madeira Islands), RA (Canary Islands), and AZ (Azores Islands).

**North American strains:** DA, AL, EU, CJ (USA), and PC (Canada).

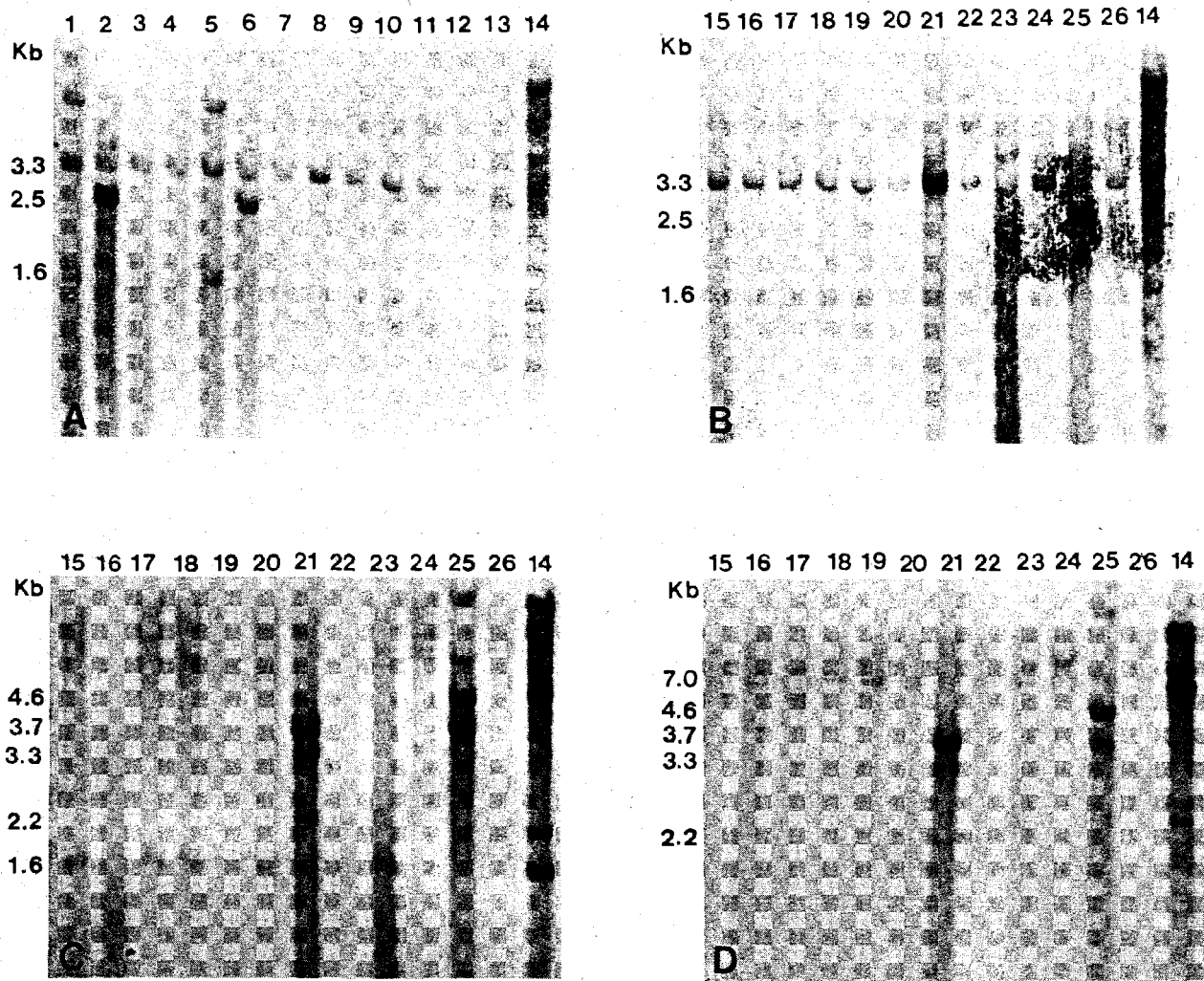
**South American strains:** BA (Argentina) and SL (Chile).

### Probes

A plasmid containing a 7.0-kb *XhoI* fragment from the complete *D. melanogaster gypsy* element was used as a probe in Southern blot hybridizations and in the screening of a *D. subobscura* genomic library. The *XhoI* restriction sites are located in the LTRs (Fig. 1), and this restriction fragment includes most of the *gypsy* element. A 7.0-kb *XhoI* fragment from the DsE1 clone, isolated from the *D. subobscura* genomic library, was used as a probe for the in situ hybridization.

### DNA Preparation, Restriction, and Blot Analysis

These methods were performed using standard procedures. Genomic DNAs were quantified by spectrofluorimetry; 7 µg of DNA per strain was digested and subsequently electrophoresed. Blot hybridizations were made in 5× SSC, 0.02% SDS at 65° C and washed in 1× SSC, 0.1% SDS at 45° C. Probes were labeled



**Fig. 2.** Southern blots of genomic DNAs of different *D. subobscura* strains probed with the *Xho*I fragment from the *D. melanogaster* gypsy element. A European strains. B–D American and African strains. DNA samples were digested with: *Hind*III (A and B), *Kpn*I (C) and *Bgl*II (D). Strains: Lane 1, H271; 2, SN;

3, G; 4, L; 5, TÙ; 6, DIE1; 7, DIE2; 8, FT; 9, CY; 10, TA; 11, FP; 12, PM; 13, COL; 14, OrR of *D. melanogaster*; 15, DA; 16, AL; 17, BA; 18, EU1; 19, EU2; 20, CJ; 21, PC; 22, SL; 23, MA; 24, RA; 25, MAR; 26, BI. Nomenclature of the strains is described in Materials and Methods

with digoxigenin-dUTP. Color detection was performed according to the manufacturer's protocols (Boehringer Mannheim).

DNA sequences were analyzed using DNASTAR and CLUSTAL programs.

### In Situ Hybridization

In situ were carried out following the method described in Terol et al. (1991). Probes were labeled with  $^3\text{H}$ -dCTP.

### Results

#### GypsyDS Sequences in Natural Populations of *D. subobscura*

### Cloning and DNA Sequencing

Gypsy homologous sequences were isolated from a library constructed with genomic DNA from *D. subobscura* H271 strain. 28 positive clones were identified. From these, clone DsE1 was selected for further analyses since it contained a full-length element. Different restriction fragments were subcloned into pUC18 and pUC19 and subsequently sequenced. Plasmid DNA sequencing was performed by the Sanger dideoxy chain termination method (Sanger et al. 1977; Tabor and Richardson 1987).

We have analyzed the presence of sequences homologous to the *D. melanogaster* gypsy element in natural populations over the distribution area of *D. subobscura*. Genomic DNAs from each of the different strains were digested with *Xho*I, *Bgl*II, *Bam*HI, *Hind*III, or *Kpn*I and probed with the almost complete *D. melanogaster* gypsy element. A selection of the blots is shown in Fig. 2. The important points to emphasize from these analyses are as follows:

Table 1. Distribution of *gypsyDS* sequences on polytene chromosomes of *D. subobscura*<sup>a</sup>

Region: Strain	Chromosome												
	A				J		U				E		O
	1A	2A	3A	7B	17B	35AB	37A	38C	39D	50C	54A	55A	84A
PC1	+	-	++	++	++	++	-	-	++	-	-	-	-
PC2	-	-	++	++	++	++	-	-	+	-	-	-	-
TU1	-	++	++	++	-	++	-	++	++	-	-	-	-
TU2	-	-	++	+	-	+	+	+	+	-	++	-	-
DIE1	-	-	+	++	-	++	++	++	++	-	-	++	-
DIE2	-	-	+	+	-	+	+	+	+	-	-	+	-
SN1	-	++	++	++	++	++	-	+	+	++	-	-	-
SN2	-	++	++	++	++	++	-	++	++	++	-	-	-
MAR1	-	-	++	++	-	++	-	++	++	-	-	-	++
MAR2	-	-	++	+	-	++	-	++	++	-	-	-	++
H271	+	-	+	++	-	++	-	+	+	-	-	-	-

<sup>a</sup> ++ strong hybridization signal, + weak hybridization signal, - no hybridization

First, *gypsyDS* sequences are present in all strains of *D. subobscura* analyzed.

Second, the hybridization patterns vary among strains, indicating the possible presence of actively transposing elements.

Third, complete *gypsyDS* sequences were detected in all strains analyzed as 7.0-kb *Xho*I fragments (data not shown).

Fourth, the internal molecular structure of the element tends to be conserved among strains; for example, all of the analyzed strains contained 3.3- and 1.6-kb *Hind*III and 1.6-kb *Kpn*I internal fragments (Fig. 2A-C); 1.6-kb *Hind*III and *Kpn*I fragments are also present in the *D. melanogaster* *gypsy* element.

Finally, five of the 26 strains analyzed (lane 2 from Sweden, lane 5 from Germany, lane 6 from Switzerland, lane 21 from Canada, and lane 25 from Morocco) show striking hybridization patterns with strong signal intensity and different banding patterns. This suggests that *gypsyDS* sequences in these strains could be polymorphic for some restriction sites and be present at a high copy number.

#### Chromosome Localization and Copy Number of *GypsyDS* Sequences

In order to determine the number and chromosomal distribution of *gypsyDS*, we conducted an in situ analysis of polytene chromosomes in larvae from SN, TÛ, DIE, PC, MAR (two individuals per strain), and H271 strains. The results, which are summarized in Table 1, indicate that *gypsyDS* sequences hybridize to both heterochromatic and eu-

chromatic regions. The number of positive euchromatic hybridization sites varies from four to seven per strain. Positively hybridizing sections 1A-2A (Fig. 3B), 17B (Fig. 3A,E), and 37A and 54A (Fig. 3C) are localized in the centromeric ends of A, J, U, and E chromosomes, respectively, which indicates that *gypsyDS* in these regions is located in, or near, the heterochromatin. The *D. melanogaster* *gypsy* element hybridizes extensively to the chromocenter and to only a few sites on the chromosome arms (Modolell et al. 1983; Bajev et al. 1984). Of the 13 chromosomal sites of hybridization, only 3A, 7B, 35AB, and 39D are shared by all the strains, and the rest of them vary among strains. Remarkably, the 50C site of the E chromosome (Fig. 3D) and the 84A site of the O chromosome are restricted, respectively, to SN and MAR strains. The sites tend to be the same for the two individuals of each strain, except in TÛ, in which three sites were altered.

These results suggest that *gypsyDS* sequences are able to transpose in *D. subobscura*. However, most intrastrain differences are found in heterochromatic regions. Heterochromatin in polytene chromosomes is underreplicated and could give negative hybridization even if *gypsyDS* elements were present.

Remarkably, careful analysis of hybridization signals shows that signals are almost always composed of several lines or regions of labeling. While the 17B region from strain PC appears to be composed of two labeled lines (Fig. 3E), 17B from strain SN appears as a single strong line (Fig. 3A). This could be due to the fact that *gypsyDS* sequences are so close on the chromosomes that they can only be

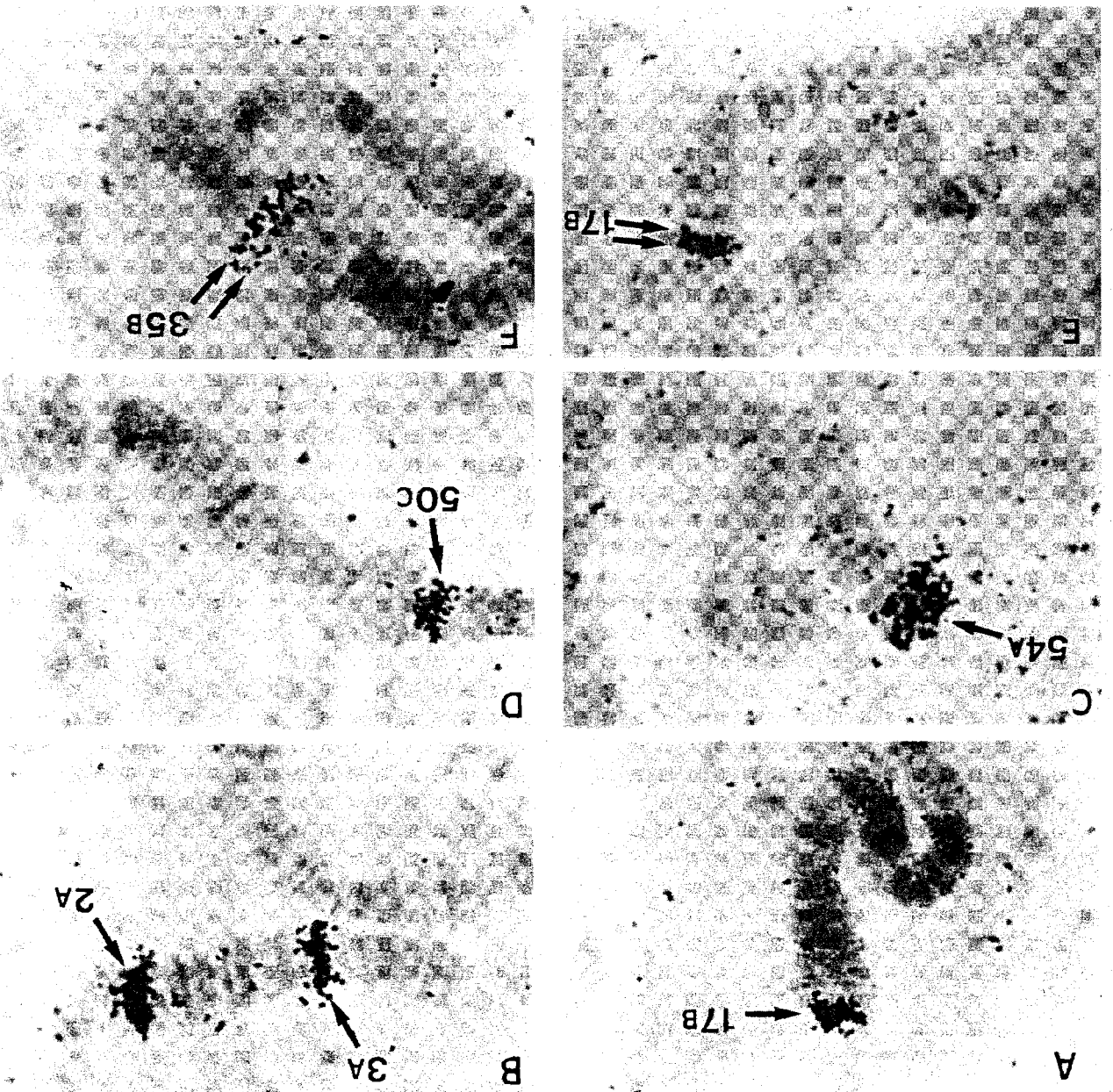
*A. D. subobscura* genomic library from H271 strain was screened with the *gypsy* *Xho*I fragment of *D.*

### Cloning a Full-Sized GypsyDS

discerned as separate lines when chromosomes are stretched (Fig. 3E) or puffed (Fig. 3F loci 35B). An alternative explanation for the intensity differences in Southern blot hybridizations between strains could be the variable copy number of *gypsyDS* sequences per labeled chromosome locus.

*melanogaster*, which encompasses almost the complete element. From the 28 positive clones, at least two of them, including clone DSE1, contained apparently full-length copies. A 1.3-kb fragment from clone DSE1 was sequenced and was found to align with positions 3580–4885 of the *D. melanogaster gypsy* sequence (Fig. 4): 77.1% similarity was found between these nucleic acid sequences with no deletions or insertions. This region showed 77.0% similarity between *D. melanogaster* and *D. virilis* but, surprisingly, 91.6% similarity between clone DSE1 and *D. virilis* (Fig. 4), suggesting that genetic dis-

Fig. 3. In situ hybridization on the polytene chromosomes of *D. subobscura* strains probed with *gypsyDS* sequence from the DSE1 clone. A, B, D, and F SNI strain, C TQ2 strain, E PC strain



mel GACTACCCCTGAACCAGACTGCGCTTACAAAGTTAGCTGCTTCCTTGGTTAGCCAGCTAC 3640	mel CTTAAGCAGGTGGTAAACCCTGACCTCGTGAACGCTATTCACTGCGACCTGCCACTCTG 4480
sub --A-T-----G-A-T--C---GT-----A-----T-G-G-----	sub --C--AC-A--T--G--TT--GA-----A--T--A--TT-----T--C-----GT--
vir --A-T-----A-GAAT--C---GC-----T-G-G-----A----- 3403	vir --C-----A--T--G--TT--CA-----T--T--A--A--T-----T--T-----GT-- 4243
mel TACAGACTCTTCATCAAGACTTTGCTGCCATAGCCCGCCGATACCCGATATCCTAAAA 3700	mel GCAAGCTTCCAACAGCACCTCAITGGCCACTTTCCAGCCCAATTTGGTCACTGTAAG 4540
sub --G--T--T--T--T-----C--T--T--T-----AJ-G-C--A--T-----	sub -----T-----G--TTA--T--C--T--T--G-----G-----T--G--T--C--A
vir --T--G--T--T--T--G-----C--T--T--G--T-----AT-G-C--A--T----- 3463	vir -----T-----G-----TTCT-T--C--T--T-----CT-G--T--C--A 4303
mel GCGGAAAAAGCTTCGGTGAAGCAACACATGCTTAAAAAAATTCCTGTTGACTTTAATGAA 3760	mel AATGTCCTGTTAGACATAACCGACAAAAACCAACAGATCGAAATCCTCACTGCCGACAC 4600
sub -----T-----G--C-A--A--G--G--C-----A--C--C-----	sub --T--A--CAG--TG-----G-----C--T--A-----A-----A-----
vir -----A--T--T--CG--G--A--A--G-----C-----A----- 3523	vir --C--T--CA--T--T--G-----G-----G--T--A-----A----- 4363
mel ACTCAACGCAACGCGTTCCAAAGACTCCCAACATACTAGCATCCGAGGATGTCATACTC 3820	mel AACCGCGTCAACAGCCGACAAACAAACATTAACCAAGTCCCTCGGATTATTACTTT 4660
sub CT-----G-G-A--TCC-T--G--T--TT-G-----A--A-----G--A	sub --T--A--C-----G--T--G-----T-----G-----A--C--TA-C--C--T--
vir CT-----G-T--A--T--CC-T--G-GT--TT-G-----A--A--C-----GT-A 3583	vir --T--A--G-----A--T--C-----C-----A-----C-----T----- 4423
mel AAATACCCCGACTTTAAAAAGCCTTTGACCTTACTACAGATGCTTGGCAAGTGGTATC 3880	mel CCCAAAAAGGCACTTTGACTAAAGAGTAGTAGCTAATTCAGGGCTGCACCCCAAGCA 4720
sub --TC--T--A--T-----C--G--A--A--C--T--A--C--G-----C-----A--T	sub -----C-----A-----C-----G--G--G--C--C-----TA-----C
vir --TC--T--AA--CC-G--A--G--C--TT-A--C--T-----C-----A--T 3643	vir -----T-----C-----C--G--C-----G--T--A--C--CC-----TA----- 4483
mel GGTGCAGTCTATCCAGGAGGGCAGCCCAATCACCATGATATCGCGTACCCTTAAACAG 3940	mel AAGTATCACAGCCACCCGAAAAAGCAAGAGCTCGGGAAAAACCCCATACCCAGCTATACA 4780
sub -----G--T--C-----A-GCAAT--A--C-----T--T--CG-TT-A-----A	sub -----T-----G--A-----A-----T-----T-----T-----T
vir --C--G--T--T--T--A-GCAAT-----A--A-----T--T--CG--T--A----- 3703	vir --A--C--T-----G--A--A-----A-----T-----A--T-----T 4543
mel CCCGACGAGAAGTACGCCCAAAACGAAAGGAAATTCGCGCCATGATGCGCCCTAGCT 4000	mel GGTGAGATGCTGCATATTGACATATTCCTCAACCGACAGGAAGCTATTCTGACCTGATT 4840
sub G-----A--T-----T-----C-----A--G-----C-----G-----TT--C	sub -----A-----T-----T-----T-----G--G--A-----A--A--T-----C--C--A
vir G-----A--T-----A-----C-----A-----C-----G-----TT--C 3763	vir --C--A--T-----T-----G--A-----A--AG--T--T-----C--CG--A 4603
mel AAGTTCAGAACTTCTGTATGGCTCTAGGGAGATTAATAATTTACCGACCATCAACCC 4080	mel GACAAATTTCTAAATATGCAATAGTGCAACAGTCTGCTGTAGA 4885
sub --GAC-----T-----C-----A-----C--T--C--A-----G--G	sub -----C-----T-----C--C-----CA--AT-----
vir --GA-----T-----C--T-----A-----C--T--C--G-----C--G--G 3823	vir -----C--A--T-----G--C--C-----CA--AT--A--C----- 4648
mel CTCACCTTGGCTCTTGGCCAGAGAACCAAGATGCCAAGATAAAGAGGTGCAAAATCTTAC 4120	
sub --G--T-----G--A--A-----A--T-----T--A--A-----	
vir -----G-----GT-G--A-----T-----T--A--A----- 3883	
mel ATAGACCAGCATAATGCCAAGCTTTTCTACAAACCTGGCAAGAAAAATTCGTGGCAGAC 4180	
sub --T--T-----C-----AA-G--T--T-----A-----C--G-----G--T	
vir --T--T-----C-----AA-G--T--T-----G--C--G-----G--T 3943	
mel GCGCTCTAGCCGAAATCTGAATCCCTTACAAAACCAACCCCAATCAGAGCTCGGACC 4240	
sub -----T--A-----A--CA--T-----TC-----G--T--C--T--GT--G-----A--	
vir -----T--A-----A--CA--T-----TC--CG--G--T--CG-----GT--G-----A-- 4003	
mel ATTACAGTCACTCTCCCTGACCTACACGGTGCAGACAACAGCAAAACCGTTAAATGTC 4300	
sub G-----C--A-----A--CT-----A--G-----	
vir G-----C--A-----A-----T-----T--G----- 4063	
mel TTCAGGAACCATGATCTGCGAGGCGACGCTTTTCGGCTCAAAACGAAACCTGGTCTC 4360	
sub --T--T-----TG-C-----A--T--G-----A--AG-G-----T--T--G	
vir --T--T-----TG-C--A--A--T--G-----A--G-----T--T--G 4123	
mel TTTCGAAGCAATCTGCCACTTAATCACTTACTGATAAAAGTGGCTATTAAAAACA 4420	
sub -----C-----A-----C--C-----T--AC-----C-----CACAA--TC--TG--T--	
vir -----C-----A--A-----C-----T--AC-----C--G--C--CAA--TC--C-----C 4183	

Fig. 4. Comparison of the 1.3-kb gypsy nucleic acid sequences among *D. melanogaster*, *D. subobscura*, and *D. virilis*

tances are not correlated with phylogenetic distances between these species. (See Discussion.)

Sequence comparisons were also used in an attempt to determine whether gypsyDS elements are functionally active, by assuming that selection would act against replacement substitutions in functional elements. The 1.3-kb sequenced region aligns with the terminal part of the reverse transcriptase domain and the initial part of the endonuclease domain in ORF2 of the *D. melanogaster* gypsy element (arrowed in Fig. 5). Most of the amino acid changes between DsE1, *D. melanogaster*, and *D. virilis* are conservative (indicated by asterisks in Fig. 5). Furthermore, the values of  $D_S$  (number of silent substitutions per effectively synonymous site) and  $D_R$  (number of replacement substitutions per

nonsilent site) for *D. melanogaster*-*D. subobscura* are  $D_S = 0.6980$  and  $D_R = 0.0828$ , and for *D. subobscura*-*D. virilis*  $D_S = 0.2481$  and  $D_R = 0.0349$  (Table 2). In both cases the values of  $D_S$  are much higher than  $D_R$ , which indicates that gypsy sequences seem to have been subjected to purifying selection against replacement substitutions.

## Discussion

GypsyDS Sequences Seem to be Transpositionally Active in *D. subobscura*

Sequences homologous to TEs of *D. melanogaster* are not well represented in the genome of *D. sub-*

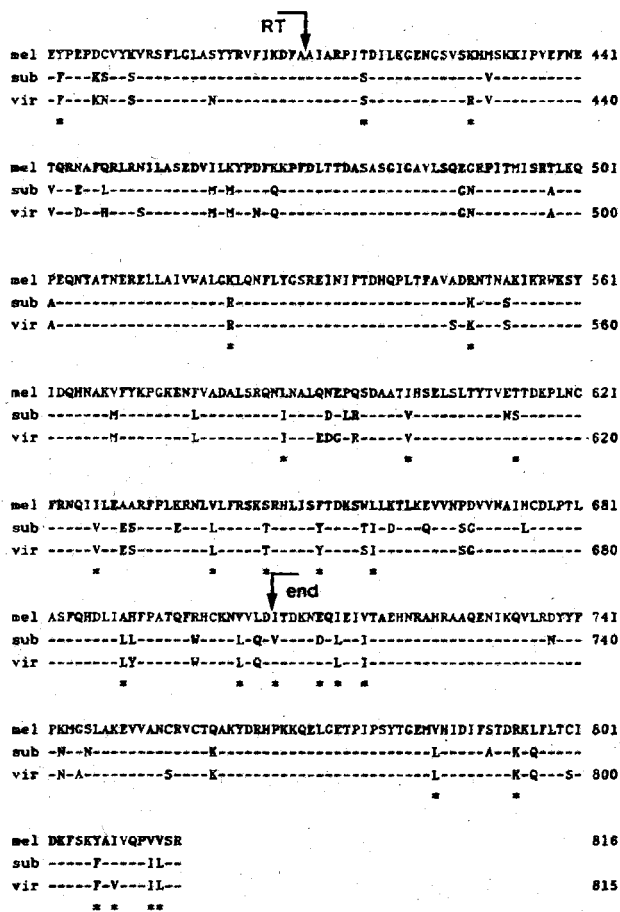


Fig. 5. Alignment of the putative protein sequences from *D. melanogaster*, *D. subobscura*, and *D. virilis*. Conservative changes are indicated by an asterisk. The following changes were considered conservative: L = V = I = A / S = T / K = R / E = D / N = Q / Y = F. The terminal part of the reverse transcriptase domain and the initial part of the endonuclease domain are indicated by arrows

*obscura*. Southern blot analysis has shown that 8 out of 14 *D. melanogaster* elements tested were completely absent from *D. subobscura* (*copia*, 297, *mdg3*, 3S18, *G*, *jockey*, *hobo*, and *FB3*); of the other 6 elements, *mdg1*, *P*, and *I* showed weak hybridization signals, and only *gypsy*, 412, and *F* showed hybridization at high stringency (de Frutos et al. 1992). We do not know whether these 6 families of elements (*mdg1*, *P*, *I*, *gypsy*, 412, and *F*) include true active mobile elements in the *D. subobscura* genome which could have a major influence on variation and evolution or are instead composed only of inactive relics.

From the 6 families of sequences homologous to *D. melanogaster* TEs described in *D. subobscura*, we have information only on *P* sequences and *gypsyDS*. *P* sequences in *D. subobscura* lack exon 3 and thus cannot transpose (Paricio et al. 1991). In contrast, several observations suggest that *gypsyDS* sequences are active in *D. subobscura*:

First, the results of Southern analysis, namely,

heterogeneous banding patterns among different strains, are those expected for true mobile elements.

Second, the distribution of occupied sites on polytene chromosomes varies in the six strains analyzed.

Third, comparison between *D. melanogaster* and *D. subobscura* shows that the number of silent substitutions per effectively synonymous site ( $D_S = 0.6980$ ) is significantly higher than the number of replacement substitutions per nonsilent site ( $D_R = 0.0828$ ). Similar results were obtained when comparing *D. virilis*–*D. subobscura* sequences ( $D_S = 0.2481$  and  $D_R = 0.0349$ ), suggesting that the evolution of *gypsy* sequences is subjected to functional constraints.

### Gypsy Sequences Are Subjected to Horizontal Transfer

The recent-invasion hypothesis proposed by Kidwell (1979, 1983) assumes that *P* elements were transferred horizontally between *Drosophila* species. Subsequent experimental evidence strongly supports the existence of horizontal transfer of *P* elements among *Drosophila* species (Anxolabéhère et al. 1988; Daniels and Strausbaugh 1986; Daniels et al. 1990a). Horizontal transfer also seems to be involved in the spreading and evolution of *hobo*, *I*, and *jockey* retrotransposons (Abad et al. 1989; Mizrokhii et al. 1990; Pascual et al. 1991). However, the extent to which horizontal transfer of these elements and other TEs occurs in *Drosophila* remains unknown (Kidwell 1992).

Retroelements are a widely extended group of genetic elements in living organisms. Comparison of retrosequences indicates interesting evolutionary relationships between them; for example, the *gypsy* retrotransposon group clusters closer to caulimoviruses than to the *copia* element, and *copia* clusters with *Ty1* from yeast (Doolittle et al. 1989). Horizontal transfer has been invoked to explain these relationships. We have little information about evolutionary relationships of retroelements among *Drosophila* species. With respect to the *gypsy* element, an extensive survey by blot hybridization has been carried out in order to detect the distribution of *gypsy*-homologous sequences among the *Drosophila* genus (Stacey et al. 1986). They are widely distributed among both the *Sophophora* and *Drosophila* subgenus, and are also present in *D. busckii*, belonging to the *Dorsilopha* subgenus. It is possible therefore that ancestral *gypsy* sequences were present before early radiations. However, no correlation has been found between phylogenetic distance and conservation of *gypsy*-homologous sequences relative to *D. melanogaster*. The most

Table 2. Comparative analysis of gypsy sequences<sup>a</sup>

Sites	<i>D. subobscura</i>				<i>D. subobscura</i> vs. <i>D. virilis</i>		
	Total no	No of changes	Percent divergence (D)	Corrected percent divergence (k)	No of changes	Percent divergence (D)	Corrected percent divergence (k)
Silent sites	302	211	D <sub>S</sub> = 0.6980	k <sub>S</sub> = 2.0103	75	D <sub>S</sub> = 0.2481	k <sub>S</sub> = 0.3013
Replacement sites	1003	83	D <sub>R</sub> = 0.0828	k <sub>R</sub> = 0.0877	35	D <sub>R</sub> = 0.0349	k <sub>R</sub> = 0.0357
Total site	1305	298	D <sub>T</sub> = 0.2284	k <sub>T</sub> = 0.2724	110	D <sub>T</sub> = 0.0843	k <sub>T</sub> = 0.0894

<sup>a</sup> The number of silent sites was calculated according to Hartl and Clark (1988). Corrected percent divergence was estimated as  $k = -\frac{3}{4} (\ln 1 - \frac{1}{2} D)$

conserved sequences were observed in species from *Drosophila* subgenus. A similar degree of conservation has only been found in species of the *Sophophora* subgenus belonging to the *obscura* subgroup. Subsequently, a survey for gypsy-homologous sequences among the *obscura*-group species indicated that, although they are present in all 15 species analyzed, conserved sequences are restricted to the *obscura* and *affinis* subgroups. Sequence comparisons of a 1.3-kb gypsy ORF2 region of *D. melanogaster*, *D. virilis*, and *D. subobscura* reinforces the degree of conservation inferred by blot analysis. The percent identity of the 1.3-kb nucleotide sequence is 77.2% between *D. subobscura* and *D. melanogaster*, 91.5% between *D. subobscura* and *D. virilis*, and 77.0% between *D. melanogaster* and *D. virilis*. There is a remarkable similarity between these values and those from conserved conventional genes among species of *obscura-melanogaster-virilis* groups (Blackman and Meselson 1986; Wilde and Akam 1987; Seeger and Kaufman 1990), particularly since retrotransposons might be expected to evolve more rapidly than other sequences because of the use of error-prone reverse transcriptase in their replication. Most significantly, the degree of similarity in gypsy sequences is not correlated with the phylogenetic distances between species. Assuming that the rate of evolution of gypsy sequences has been similar in the three species, the molecular divergence time between *D. subobscura* and *D. virilis* sequences is estimated as 16 MY. This time is much lower than the most likely divergence time of more than 60 MY for the divergence between the *virilis* and *melanogaster-obscura* groups (Beverly and Wilson 1983; Throckmorton 1975). These results strongly suggest that gypsy sequences have been transferred horizontally between these species.

The geographical distribution of these species is also consistent with this hypothesis because they overlap extensively. While *D. melanogaster* and *D. virilis* are cosmopolitan species, the distribution of *D. subobscura* is restricted to Europe and North

Africa. At present, this last species is colonizing North and South America (Brncic et al. 1981; Beckenbach and Prevosti 1986).

Examples of horizontal transfer between species are becoming increasingly numerous. A phenomenon that at first seemed unlikely is now accepted as a probable mechanism of sequence spreading between individuals (Kidwell 1992). Moreover, if a lot of horizontal transfer is going on, the different in situ patterns observed in gypsyDs could also be explained by independent and repeated horizontal transfer events that went on many times, in many places. This latter explanation seems unlikely due to the low probability of the occurrence of many horizontal transfer events. The fact that gypsyDs could be mobile in *D. subobscura* is a simpler explanation for the in situ hybridization patterns and our present research aims to confirm it.

**Acknowledgments.** This work was supported by grants from the Spanish CICYT (No. PB87-0952) and FPI programmes. The nucleotide sequence was determined at Servicio de Secuenciación de la Universitat de València. Sequence analyses were done at Servicio de Bioinformática de la Universitat de València. We would like to thank R. González-Duarte and G. Marfany for the *D. subobscura* genomic library, Dr. J.L. Ménsua for helping in the evolutionary aspects of the manuscript, N. Paricio for her help in sequencing, and J. Ferré and J.F. Barraclough for improving the English of the manuscript.

## References

- Abad P, Vaury C, Pelisson A, Chaboissier MC, Busseau I, Bucheton A (1989) A long interspersed repetitive element—the *I* factor of *Drosophila teissieri*—is able to transpose in different *Drosophila* species. Proc Natl Acad Sci USA 86: 8887–8891
- Anxolabéhère D, Kidwell MG, Periquet G (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
- Bajev AA, Lyubomirskaya NV, Dzhumagaliev EB, Ananiev EV, Amiantova IG, Ilyin YV (1984) Structural organization of transposable element *mdg4* from *Drosophila melanogaster* and a nucleotide sequence of its long terminal repeats. Nucl Acid Res 12:3707–3723



- Beckenbach AT, Prevosti A (1986) Colonization of North America by the European species, *Drosophila subobscura* and *Drosophila ambigua*. *Am Midland Naturalist* 115:10-18
- Bingham PM, Zachar Z (1989) Retrotransposons and the *FB* transposon from *Drosophila melanogaster*. In: Berg DE, Howe MM (eds) *Mobile DNA*. American Society for Microbiology, Washington, DC, pp 485-502
- Blackman RK, Meselson M (1986) Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila hsp82* gene. *J Mol Biol* 188:499-515
- Brcic D, Prevosti A, Budnik M, Ocaña J (1981) Colonization of *Drosophila subobscura* in Chile I. First population and cytogenetic studies. *Genetica* 56:3-9
- Brookfield JFY, Montgomery E, Langley CH (1984) Apparent absence of transposable elements related to the *P* elements of *Drosophila melanogaster* in other species of *Drosophila*. *Nature* 310:330-332
- Csink AK, McDonald JF (1990) *Copia* expression is variable among natural populations of *Drosophila*. *Genetics* 126:375-385
- Daniels SB, Strausbaugh LD (1986) The distribution of *P* element sequences in *Drosophila*: the *willistoni* and *saltans* species groups. *J Mol Evol* 23:138-148
- Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A (1990a) Evidence for horizontal transmission of the *P* transposable element between *Drosophila* species. *Genetics* 124:339-355
- Daniels SB, Chovnick A, Boussy IA (1990b) Distribution of *hobo* transposable elements in the genus *Drosophila*. *Mol Biol Evol* 7:589-606
- Doolittle RF, Feng DF, Johnson MS, McClure MA (1989) Origins and evolutionary relationships of retroviruses. *Quart Rev Biol* 64:1-30
- Engels WR (1989) *P* elements in *Drosophila melanogaster*. In: Berg DE, Howe MM (eds) *Mobile DNA*. American Society for Microbiology, Washington, DC, pp 437-484
- Finnegan D, Fawcett D (1986) Transposable elements in *Drosophila melanogaster*. In: Maclean N (ed) *Oxford surveys on eukaryotic genes*, vol 3, pp 1-62
- Finnegan DJ (1989) The *I* factor and I-R hybrid dysgenesis in *Drosophila melanogaster*. In: Berg DE, Howe MM (eds) *Mobile DNA*. American Society for Microbiology, Washington, DC, pp 503-517
- de Frutos R, Peterson KR, Kidwell MG (1992) Distribution of *Drosophila melanogaster* transposable element sequences in species of the *obscura* group. *Chromosoma* 101:293-300
- Hartl DL, Clark AG (1988) *Principles of population genetics*. Sinauer, Sunderland, MA, pp 682
- Kidwell MG, Kidwell JF, Sved JA (1977) Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics* 86:813-833
- Kidwell MG (1979) Hybrid dysgenesis in *Drosophila melanogaster*: the relationship between the P-M and I-R interaction systems. *Genet Res* 33:105-117
- Kidwell MG (1983) Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 80:1655-1659
- Kidwell MG (1992) Horizontal transfer of *P* elements and other short inverted repeat transposons. *Genetica* (in press)
- Lansman RA, Stacey SN, Grigliatti TA, Brock HW (1985) Sequences homologous to the *P* mobile element of *Drosophila melanogaster* are widely distributed in the subgenus *Sophophora*. *Nature* 318:561-563
- McDonald JF (1989) The potential evolutionary significance of retroviral-like transposable elements in peripheral populations. In: Fontdevila A (ed) *Evolutionary biology of transient unstable populations*. Springer-Verlag, New York, pp 190-205
- Martin G, Wiernasz D, Schedl P (1983) Evolution of *Drosophila* repetitive-dispersed DNA. *J Mol Evol* 19:203-213
- Marlor RL, Parkhurst SM, Corces VG (1986) The *Drosophila melanogaster gypsy* transposable element encodes putative gene products homologous to retroviral proteins. *Mol Cell Biol* 6:1129-1134
- Mizrokhi LJ, Mazo AM (1990) Evidence for horizontal transmission of the mobile element *jockey* between distant *Drosophila* species. *Proc Natl Acad Sci USA* 87:9216-9220
- Mizrokhi LJ, Mazo AM (1991) Cloning and analysis of the mobile element *gypsy* and *Drosophila virilis*. *Nucl Acid Res* 19:913-916
- Modolell J, Bender W, Meselson M (1983) *Drosophila melanogaster* mutations suppressible by the suppressor of *Hairy-wing* are insertions of a 7.3-kilobase mobile element. *Proc Natl Acad Sci USA* 80:1678-1682
- Paricio N, Perez-Alonso M, Martinez-Sebastian MJ, de Frutos R (1991) *P* sequences of *Drosophila subobscura* lack exon 3 and may encode a 66 kd repressor-like protein. *Nucl Acid Res* 19:6713-6718
- Pascual L, Periquet G (1991) Distribution of *hobo* transposable elements in natural populations of *Drosophila melanogaster*. *Mol Biol Evol* 8:282-296
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci* 74:5463-5467
- Silber J, Bazin C, Lemeunier F, Aulard S, Volovitch M (1989) Distribution and conservation of the Foldback transposable element in *Drosophila*. *J Mol Evol* 28:220-224
- Seeger MA, Kaufman TC (1990) Molecular analysis of the *bicoid* gene from *D. pseudoobscura*: identification of conserved domains within coding and noncoding regions of the *bicoid* mRNA. *EMBO J* 9:2977-2987
- Stacey SN, Lansman RA, Brock HW, Grigliatti TA (1986) Distribution and conservation of mobile elements in the genus *Drosophila*. *Mol Biol Evol* 3:522-534
- Tabor S, Richardson CC (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc Natl Acad Sci USA* 84:4767-4771
- Terol J, Perez-Alonso M, de Frutos R (1991) In situ localization of the *Antennapedia* gene on the chromosomes of nine *Drosophila* species of the *obscura* group. *Hereditas* 114:131-139
- Throckmorton LH (1975) The phylogeny, ecology and geography of *Drosophila*. In: King RC (ed) *Handbook of genetics*, vol 3. Plenum Press, New York, pp 421-469
- Wilde CD, Akam M (1987) Conserved sequence elements in the 5' region of the *Ultrabithorax* transcription unit. *EMBO J* 6:1393-1401