Genet. Res., Camb. (2000), **75**, *pp.* 13–23. *With* 4 *figures.* Printed in the United Kingdom © 2000 Cambridge University Press

A scenario for the *hobo* transposable element invasion, deduced from the structure of natural populations of *Drosophila melanogaster* using tandem TPE repeats

ERIC BONNIVARD[1], CLAUDE BAZIN[2], BEATRICE DENIS[2] AND DOMINIQUE HIGUET[1,*]

[1] *Laboratoire Dynamique du Génome et Evolution, Institut J. Monod, 2 place Jussieu, T42-32, 4ème étage, 75251 Paris Cedex 05, France*
[2] *Laboratoire Population, Génétique et Evolution, CNRS, 13 avenue de la Terrasse, 91178 Gif sur Yvette, France*

(*Received 7 December 1998 and in revised form 19 March 1999*)

Summary

Temporal surveys of *hobo* transposable elements in natural populations reveal a historical pattern suggesting a recent world-wide invasion of *D. melanogaster* by these transposons, perhaps following a recent horizontal transfer. To clarify the dynamics of *hobo* elements in natural populations, and thus to provide further data for our understanding of the *hobo* invasion, TPE tandem repeats, observed in the polymorphic S region of the element, were used as molecular markers. The number of TPE repeats was studied in 101 current populations from around the world, and in 63 strains collected in the past. This revealed a geographical distribution which seems to have been stable since the beginning of the 1960s. This distribution is compatible with a number of hypotheses for the dynamics of *hobo* elements. We propose a scenario based on an invasion in two stages: first, a complete invasion by elements with three TPE repeats, followed by the beginning of a new invasion involving *hobo* elements with five or seven repeats.

1. Introduction

First described in *Drosophila melanogaster* (McGinnis *et al.*, 1983), the *hobo* elements are one of the three transposable element families (*I*, *P* and *hobo*) known to cause hybrid dysgenesis in *D. melanogaster* (for a review see Finnegan, 1989*a*; Bucheton *et al.*, 1992; Engels, 1989; Periquet *et al.*, 1994). Structurally they belong to the *hAT* superfamily (Calvi *et al.*, 1991), which is a member of the class II elements (Finnegan, 1989*b*; Capy *et al.*, 1997), transposing via a DNA intermediate and having short inverted terminal repeats (ITR) of 12 bp.

The distribution of *hobo* elements in *D. melanogaster* has been studied chiefly through restriction surveys of genomic DNA, which detect the presence of the 2·6 kb *Xho*I fragment characteristic of supposedly complete *hobo* elements. Streck *et al.* (1986) described lines which contained the 2·6 kb fragment which were called 'H' (for '*hobo*'), whereas other lines, called 'E' (for 'empty'), lacked it. All lines (H or E) also show bands of greater than 3 kb (which may be as heavy as 20 kb). These high-molecular-weight bands could correspond to vestigial sequences localized in heterochromatin (Lim, 1988; Daniels *et al.*, 1990).

More extensive surveys of *hobo* elements in natural populations have revealed a historical pattern, with E lines most prevalent in North America before 1950 and in Europe before 1960, and with H lines found thereafter (Periquet *et al.*, 1989*a*, *b*, 1990; Boussy & Daniels, 1991; Pascual & Periquet, 1991). On the basis of the correlation between collection date and the presence of the 2·6 kb *Xho*I fragment in strains, Periquet *et al.* (1989*a*) suggested a recent world-wide invasion of *D. melanogaster* by the *hobo* element, similar to the scenario proposed for *P* elements (Anxolabéhère *et al.*, 1988). They suggested that the introduction of *hobo* could have occurred before about 1950 in America and that the complete *hobo* element has a history of recent horizontal transfer, *D. simulans* being the likely donor species (Periquet *et al.*, 1990; Pascual & Periquet, 1991). Boussy & Daniels (1991) corroborated the historical pattern and agreed with the possibility of a recent introduction and spread of *hobo* sequences in *D. melanogaster*. However, as they also found E strains in *D. simulans*, they raised doubts about the direction of transfer and,

* Corresponding author.
e-mail: higuet@ccr.jussieu.fr

https://doi.org/10.1017/S001667239900395X Published online by Cambridge University Press

indeed, about the validity of any simple explanation. Simmons (1992) found evidence from DNA sequencing of horizontal transfer within the *D. melanogaster* complex, but concluded that no available data supported an unambiguous interpretation about the time or direction of horizontal transfer events among these species.

On the basis of the *hobo* system of hybrid dysgenesis, a new scenario for the invasion of *D. melanogaster* by *hobo* elements was proposed according to the geographical distribution of *hobo* activity in 13 world populations (Bonnivard *et al.*, 1997). This scenario takes into account the particular properties of the African population, which possessed all kinds of activity (GD sterility, *hobo* mobilization using two reporter genes $h(w^+)$ and hvg^{a1}). Two hypotheses were suggested considering Africa either as the start of the invasion or as its end point. However, as the geographical distribution of *hobo* activity does not reveal a clear differentiation in world populations, the dynamics of *hobo* elements in natural populations remains unclear.

The *hobo* element contains, in the ORF 1, a polymorphic S region (Fig. 1) consisting of tandem repeats of a 9 bp 'actccagaa' sequence (Streck *et al.*, 1986; Calvi *et al.*, 1991; Bazin & Higuet, 1996). These sequences are called TPE repeats as they encode a 'Threonine (T)–Proline (P)–Glutamic acid (E)' motif in the protein. Ten repeats were revealed in the first sequenced *hobo* element, $hobo_{108}$ (Streck *et al.*, 1986), and three were sequenced in the autonomous *hobo* reference element *Hfl*I (Blackman *et al.*, 1989; Calvi *et al.*, 1991). Different types of *hobo* elements thus exist with regard to the number of TPE repeats. New types of element with five, six or eight TPE repeats have been described in laboratory strains (Bazin & Higuet, 1996). Moreover, it appears that different types of *hobo* elements can be found in the same strain (e.g. elements with three, five or ten repeats in the AL strain; Bazin & Higuet, 1996). In a previous study we have shown that there is also a variability between *hobo* elements in current natural populations of *D. melanogaster* (Bonnivard *et al.*, 1997), which allows these populations to be characterized. Based on only 13 populations, this study revealed that most populations are monomorphic (presenting only one type of element, with three TPE repeats) whereas some are polymorphic (presenting different types of elements).

TPE repeats are useful molecular markers, which provide new information on the history of invasion of *hobo* elements in *D. melanogaster*. The number of TPE repeats were surveyed in 101 current populations from around the world, as well as in 63 laboratory strains. At present, a particular geographical distribution can be observed, which seems to have been stable since the beginning of the 1960s. This suggests a new scenario for the dynamics of *hobo* elements, based on the existence of two different invasion stages. First, there was a successful and total invasion by elements with three TPE repeats, followed by the start of a new invasion of *hobo* elements with five or seven repeats.

2. Experimental procedure

(i) *Strains*

The *Drosophila melanogaster* strains used were kept at 25 °C under standard laboratory conditions on a cornmeal–sugar–yeast–agar medium.

(*a*) *Current natural populations*. One hundred and one natural populations, derived from diverse localities around the world, were investigated for their TPE motif characteristics (the location of most of these populations was described in Bonnivard & Higuet, 1999). In all cases these populations were represented by strains derived from a large sample (> 10 individuals) collected since 1992. Most of the strains were maintained as isofemale lines, in order to maintain part of the initial polymorphism. At least two generations before molecular analysis, mass cultures were established from these lines. A mass culture is a proportional mixture of isofemale lines.

(*b*) *Old wild-type strains*. Sixty-three strains collected between 1952 and 1987 were studied from the author's collection or from *Drosophila* stock centres (Table 1). As with current populations, samples derived from the same locality at the same time were mixed to make up a mass culture. However, the majority of strains are established from a few or an unknown number of females.

(*c*) *Control strains*. **CyHBL1** (Calvi & Gelbart, 1994) and **vg^{al*}** (Bazin & Higuet, 1996) strains possess only one type of *hobo* element with three and five TPE repeats respectively. The Cy/Sp white strain **341** comes from the laboratory of Professor J. A. Lepesant. Three PCR products can be obtained from this strain, the shortest corresponding to *hobo* elements with six TPE repeats (Bazin, unpublished data).

(ii) *Determining the TPE status of strains*

(*a*) *PCR amplification and sequencing*. For each strain, DNA was extracted from 50 females of the mass culture, using the method described by Junakovic *et al.* (1984). PCR amplification was performed using internal primers *h11* and *h6* of the *Hfl*1 element (Fig. 1), described in Bazin & Higuet (1996). About 50 ng of DNA were used in 25 μl of total reaction volume

https://doi.org/10.1017/S001667239900395X Published online by Cambridge University Press

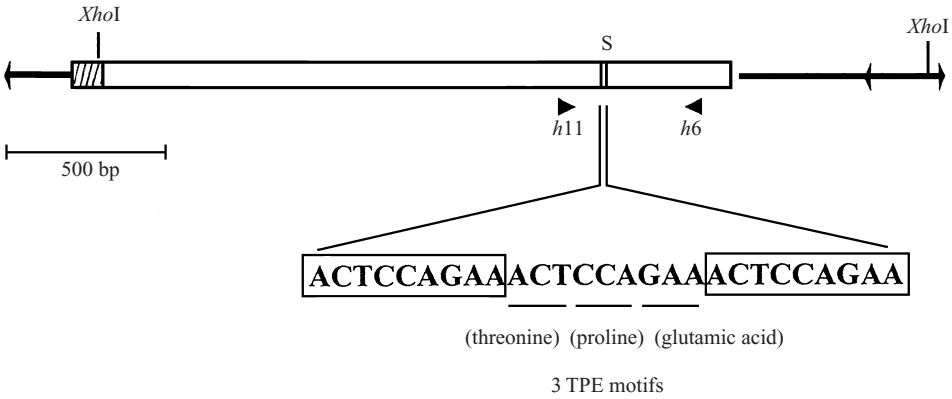

Fig. 1. Map of the *hobo* transposable reference element *Hfl1* showing *Xho*I restriction sites, localization of primers *h*11 (1756–1774) and *h*6 (2168–2148) and major structural features. This element possesses three TPE repeats in the S region. Each inverted terminal repeat is represented by a black arrowhead and the ORF0 and ORF1 regions are shown by the hatched and white horizontal bars respectively.

containing 1 unit of *Taq* polymerase (Perkin-Elmer). Amplifications were performed on a MJ Research Minicycler for 30 cycles. Cycling conditions were 95 °C for 45 s, 65 °C for 2 min and 72 °C for 2 min. For some strains, PCR products were cloned with the pGEM-T Easy Vector system (Promega) and sequenced with the T7 polymerase according to the manufacturer's recommendations.

(*b*) *Electrophoresis*. PCR products were electrophoretically separated by size in a 2·5% agarose gel. The number of TPE repeats was determined by comparing the length of the PCR product with a sequenced control fragment from the control strains (see above). In some cases, the numbers of TPE repeats was confirmed by sequences obtained from cloned PCR products.

(iii) *Southern blotting*

(*a*) *Southern blotting and hybridization*. Southern blots were performed using the same DNA extraction as for PCR amplification. About 3 μg of genomic DNA from each strain was digested with 15 units of the restriction enzyme *Xho*I, according to the manufacturer's instructions. Southern blots were performed using standard techniques (Sambrook *et al*., 1989). Filters were prehybridized at 65 °C for 2 h with a hybridization mixture (6 × SSC, 5 × FPG, 0·5% SDS), to which salmon testis DNA was added to 100 μg/ml after denaturation at 100 °C for 10 min. Hybridization was carried out overnight at 65 °C in a similar mixture, which contained 80 μg/ml salmon testis DNA and ^{32}P-labelled probe (see below). After two post-hybridization washes of 10 min in 2 × SSC, 0·1% SDS at 65 °C, the filters were wrapped in plastic film and exposed to X-ray film.

(*b*) *Probes*. Two different probes were used separately. The first was a synthetic oligonucleotide corresponding to five TPE repeats and labelled by terminal deoxynucleotidil transferase using [^{32}P]dCTP and following the protocol recommended by the manufacturer (Amersham). Southern blots were subsequently rehybridized with the ^{32}P-labelled p*Hfl*1 plasmid as probe (Blackman *et al*., 1989).

3. Results

(i) *Nature of* hobo *elements carrying TPE repeats*

To determine which kinds of *hobo* elements are potentially revealed by the *h*11–*h*6 PCR amplification, 18 strains were investigated by Southern blotting (Fig. 2). These strains were chosen according to both the location and the TPE status of the populations they derived from. Two internal *Xho*I sites (Fig. 1) allow one to distinguish the full-size element (2·6 kb fragment) and deleted elements (fragments less than 2·6 kb), including the deleted element *Th* (1·1 kb fragment; Periquet *et al*., 1989*a*). This element, which occurs in all current wild-type strains examined (Periquet *et al*., 1989*b*; Boussy & Daniels 1991; Bonnivard, unpublished data), has a deletion of 1·5 kb in the central part of the sequence including the S region (Periquet *et al*., 1990).

Two different probes were used separately to distinguish *hobo* elements carrying TPE repeats from those deleted in the S region. Using the oligonucleotide probe corresponding to five TPE repeats it was possible to detect only *hobo* elements carrying TPE repeats (Fig. 2.I). As expected, in each lane full-size *hobo* elements harbour TPE repeats. However, several deleted elements with TPE repeats are also found in each strain. Different types of such elements were present in some strains (e.g. Cuba, lane o; Spain, lane

https://doi.org/10.1017/S001667239900395X Published online by Cambridge University Press

Table 1. *Description, location and TPE status of strains derived from natural populations sampled between 1954 and 1986*

Strain (year of capture)	Latitude[a]	Longitude[a]	Source[b]	No. of initial females[c]	No. of TPE repeats
America					
USA					
Wisconsin – Maderia (196–)			B	4	3
Ohio – Mt Sterling (1966)			B	2	3
Ohio – Painesville (1966)	41·43	−81·15	B	3	3,5
New York – Commack (1961)			B	12	3,5
New York – Monroe County (1961)	41·20	−74·12	B	8	3
New York – Ceres (1966)	42·01	−78·16	B	2	3
New York – Ellenville (1966)	41·43	−74·23	B	4	3,5,7
New York – Varna (1977)	42·27	−76·27	B	12	3
Georgia – Red Mountain (1966)			B	3	3
North Carolina – Oxford (1966)	36·22	−78·37	B	2	3,5
South Carolina – Manning (1966)	33·42	−80·12	B	3	3
California – Riverside (1954)	33·59	−117·22	B	4	3
California – Riverside (1963)	33·59	−117·22	B	12	3
Colombia – Bogota (1962)	4·38	−74·05	B	4	3
Argentina – Buenos Aires (1983)	−34·40	−58·30	L	10	3,5
Africa					
Zimbabwe – Kariba Dam (1963)	−16·31	28·50	B	12	3
Europe					
France					
Kerbiniou (1960)	48·24	−4·29	L	> 20	3,5,7
Le Châtelet (1970)	46·39	2·17	L	> 20	5,9
Marseillan (1965)	43·22	3·32	L	> 20	5
Marseillan (1978)	43·22	3·32	L	> 20	3,5,7
Marseillan (1983)	43·22	3·32	L	> 20	3,5,7
Tautavel (1967)	43·20	3·30	L	> 20	3,5,7
Tautavel (1985)	43·20	3·30	L	> 20	3,5,7
Mèze (1971)	43·25	3·43	L	> 20	3,5,7
Mèze (1987)	43·25	3·43	L	> 20	3,5,7
Venteuil (1984)	49·00	3·60	L	> 30	3,5,7
St Geneviève (1985)	48·38	2·20	L	> 20	3,5,7
Portugal – Madeira (1965)	39·57	−8·02	B	16	3
Sweden					
Stromsvreten (1969)			U	un.	3
Strängnäs (1974)	59·22	17·02	U	un.	3
Finland					
Turku (1996)	60·27	22·15	U	un.	3
Nokia (1971)	61·29	23·31	U	un.	3
Hämeenlinna (1972)	61·00	24·25	U	un.	3
Naantali (1973)	60·28	22·05	U	un.	3
Austria – Vienna (1966)	48·13	16·20	U	un.	3
Czechoslovakia					
Suchumi (1964)			U	un.	3
Moravski Pisek (1972)	49·18	14·1	U	un.	3
Zabcice (1977)			U	un.	3
Former Yugoslavia					
Slankamen (1982)			U	un.	3
Golubac (1986)	46·11	16·03	L	> 10	3
Greece – Athens (1965)	38·00	23·44	B	4	3
Armenia – Ashtarak (1977)	40·19	44·23	U	un.	3,5
Former USSR					
Gurzuf (1961)	44·33	34·15	U	un.	3
Uman (1970)	48·45	30·10	L	> 10	3
Uman (1983)	48·45	30·10	L	> 10	3
Birsk (1974)	55·28	55·31	U	un.	3,5,7
Kurdamir (1977)			U	un.	3
Tashkent (1977)	41·16	69·13	U	un.	3

https://doi.org/10.1017/S001667239900395X Published online by Cambridge University Press

Table 1 (*cont.*)

Strain (year of capture)	Latitude[a]	Longitude[a]	Source[b]	No. of initial females[c]	No. of TPE repeats
Tashkent (1981)	41·16	69·13	L	un.	3
Anapa (1978)	44·54	37·20	U	un.	3
Anapa (1979)	44·54	37·20	U	un.	3
Gelendzhik (1979)	44·34	38·07	U	un.	3
Alma Ata (1981)	43·19	76·55	L	> 10	3
Israel					
Qiryat – anavim (1981)	31·49	35·07	U	un.	3,7,9
Qiryat – anavim (1983)	31·49	35·07	U	un.	3
Asia					
Japan – Kurume (1976)	33·20	130·29	L	> 20	3
Tasmania – Avondale – 2 (1982)	−24·42	152·10	U	un.	3

[a] Latitude: negative number, SLat; positive number, NLat. Longitude: negative number, WLong; positive number, ELong.
[b] Letters denote the laboratories from which the stocks were obtained: B, Bloomington Stock Center; U, Umea Stock Center; L, our laboratory collection.
[c] Number of females from which the mass culture were established; un., unknown.

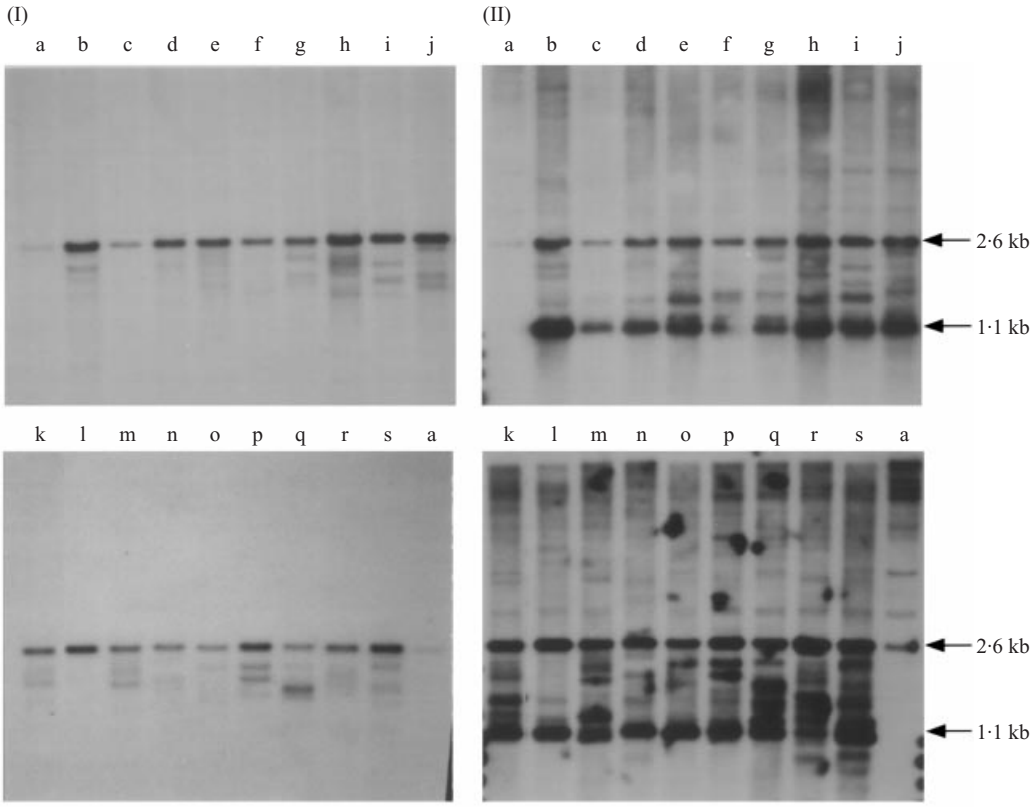

Fig. 2. *Xho*I digests of genomic DNA from a representative set of 18 current strains of *D. melanogaster* and from the laboratory strain CyHBL1, probed with (I) a TPE repeat oligo-probe (oligonucleotide corresponding to five TPE repeats) and (II) a *hobo* element probe (p*Hfl*1). The expected 2·6 kb internal *Xho*I fragment from full-sized *hobo* elements is indicated, as is the 1·1 kb fragment corresponding to the *Th* elements (Periquet *et al.*, 1989). Strain names are as follows: lane (a) CyHBL1; (b) France – Brest; (c) France – Château Thierry; (d) France – Naussac; (e) Germany – Freiburg; (f) Germany – Frankenthal; (g) Germany – Bayreuth (h) Italy – Florence; (i) Italy – Padua; (j) Spain – Madrid; (k) USA – Boston; (l) Costa Rica – Limon; (m) Bolivia; (n) Guyana – Kourou; (o) Cuba – Havana; (p) Guinea Bissau – Bissau; (q) Ivory Coast – Nimba; (r) Kenya; (s) Madagascar.

https://doi.org/10.1017/S001667239900395X Published online by Cambridge University Press

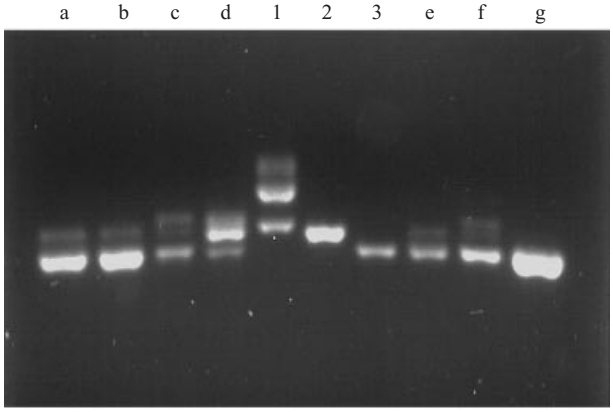

Fig. 3. PCR amplification products using *h11* and *h6* primers (Bazin & Higuet, 1996). Lanes (1)–(3) represent the control strains and lanes (a)–(g) the strains from natural populations. The lanes are as follows: lane (1) 341 (440 bp: 6 'TPE' repeats for the shortest fragment); (2) vg^{al}* (431 bp: 5 TPE repeats); (3) CyHBL1 (413 bp: 3 TPE repeats); (a) Madagascar, class [3*,5*]; (b) Kenya, [3*,5]; (c) Spain – Barcelona, [3,5,7]; (d) France – Naussac, [3*,5*,7]; (e) Guadeloupe, [3,5]; (f) Colombia – Carthagena [3*,5*,7]; (g) USA – Boston [3]. *TPE repeats whose number was controlled by sequence analysis.

j), and in some cases they even made up over half the elements (Ivory Coast, lane q). In the control strain CyHBL1 (lane a), which only contains a full-size element, no band less than 2·6 kb was revealed, confirming the specificity of this probe as a marker of the S region.

These results were compared with those obtained with supplementary p*Hfl*1 hybridization (Fig. 2.II) of the same blots. All the bands previously observed also hybridized with p*HFl*1. For each strain, the p*Hfl*1 probe revealed at most only four new bands. They correspond to the *Th* element observed in all strains that was never revealed with the oligonucleotide probe, or to other various additional fragments which could be detected in some strains (e.g. France, lane b; USA, lane k; Costa Rica, lane l). Some of these fragments, which have an approximate size of 1·5 kb (e.g. Germany, lane e; Italy, lane i; USA, lane k), could correspond to the *Oh* element described by Periquet *et al.* (1990). This element is, like the *Th* element, deleted in the central part of the sequence including the S region. With regard to the vestigial sequences of high molecular weight, hybridization with the TPE probe is undetectable or extremely weak in comparison with that obtained with p*Hfl*1.

(ii) *TPE repeats in current natural populations*

The S region of *hobo* elements was investigated using primers (*h*11, *h*6) that surround it. The numbers of TPE repeats were estimated in accordance with the length of the PCR products compared with the sequenced control fragments (Fig. 3). In several populations randomly chosen from around the world, the number of TPE repeats of some elements was confirmed by sequencing.

The PCR amplification results show variability between strains, the different migration profiles obtained being shown in Fig. 3. Some strains are monomorphic (Boston, lane g), containing only elements with three TPE repeats. The other strains are polymorphic, as they harbour elements with five or seven TPE repeats in addition. Within polymorphic strains, differences were observed in the intensity of the PCR products. These within-strain differences may express differences in the ratio of the number of elements of each type (Brunet *et al.*, 1996). Most of the strains appear to have a majority of elements with three TPE repeats (Fig. 3), to the extent that other types of element are sometimes difficult to detect (e.g. elements with five TPE repeats: Kenya, lane b; or Colombia, lane f). Only some French strains were found to harbour elements with five or seven repeats in equal or higher number than those with three repeats (e.g. Naussac, lane d).

Hobo elements with three TPE repeats are found in all the populations studied around the world (Fig. 4). Indeed, most strains are monomorphic for the three TPE element. However, several polymorphic strains were also observed, which could be grouped according to the number of PCR products detected (two or three) and the types of elements they contain. Four classes of populations around the world can thus be described (Fig. 4). The first class includes the monomorphic strains, represented by circles, which have elements with only three repeats. The two other main classes, which refer to polymorphic populations, are the [3,5] class (i.e. populations that harbour both elements with three TPE repeats and with five TPE repeats), represented by triangles, and the [3,5,7] class, represented by squares. A fourth class is represented by only one population from Greece, which harbours elements with three, five and potentially nine TPE repeats.

The monomorphic strains are found all over the world: North and Central America, North Africa, Central and Eastern Europe, Asia and Oceania. In contrast, the polymorphic populations seem restricted to three regions: Western Europe, South America and Equatorial Africa. Only strains from these regions possess several types of elements (Fig. 4). For these polymorphic populations there is a strong geographical structuring, as geographically close populations show similar kinds of polymorphisms. All the African polymorphic populations belong to the [3,5] class. The situation is more complex in Western Europe, where different classes of populations are found. In and around France, populations have elements with

https://doi.org/10.1017/S001667239900395X Published online by Cambridge University Press

Fig. 4. Geographical distribution of the number of TPE repeats in natural populations of *D. melanogaster*. Populations are distributed among four classes according to both the number and types of element detected. Due to the large number of populations analysed in France, only a proportion of them are represented.

three, five and seven TPE repeats. Around this area populations of the [3,5] class are predominant, and finally more distant populations only have elements with three TPE repeats. Lastly, in South America, two populations from the [3,5] class are observed, and two other polymorphic populations were found in Colombia and Guyana, which show a status similar to those of French populations, i.e. [3,5,7] class.

(iii) *TPE repeats in old strains*

To compare the current distribution of TPE repeats in natural populations with those of the past, old wild-type strains were analysed in the same way; results are shown in Table 1. Of the 63 strains studied, six do not show PCR amplification. Five of these (Brazil – Salvador (1952), Louisiana – New Orleans (1954), South Africa – Capetown (1954), Spain – Barcelona (1954), Peru – Ica (1959)) were collected before 1959, probably prior to the *hobo* invasion in these locations. The sixth, Sweden – Valdemarsvik (1979), was derived from a single female who could have carried no *hobo* sequence. Moreover, new classes of strains can be distinguished among the old strains, including in France a monomorphic strain with five TPE repeats (Marseillan – 1965) and a polymorphic strain which belongs to a [5,9] class (Le Châtelet – 1970). Since 1960, whatever the period considered, variability between strains can be observed, with a majority of monomorphic strains, containing only *hobo* element with three TPE repeats.

Analysis of the number of TPE repeats in old strains reveals a strong and unexpected stability of the geographical distribution: (1) nine of the 11 old French strains studied belong to the [3,5,7] class, as the current French populations; (2) 37 of the 44 old strains collected in North America, Asia, Central or Eastern Europe belong to the [3] class, as the current populations in these areas. This agreement between past and current distributions is not complete, with two types of peculiar observations: (1) old strains that do not belong to the [3,5,7] class in France; (2) elements with five or seven repeats being seen outside their current geographical range. Regarding the appearance or disappearance of some types of elements in the two French strains (Marseillan – 1965 and Le Châtelet – 1970, Table 1), this could be explained by the founding effect at the establishment of these old strains collected in polymorphic areas. This assumes that polymorphic populations have an inter-line variability. Under this hypothesis, elements with nine TPE repeats, probably rare in some French natural populations and thus undetectable through mass culture analysis, could have increased in frequency in the France – Le Châtelet strain. Moreover, in such polymorphic strains the loss of elements types could also be enhanced by the accelerated evolution that is possible when strains are kept in laboratory conditions, through drift and their having more generations per year than wild populations. Regarding the appearance of new TPE variants in four North American strains, two European strains and in Israel (Table 1), two hypotheses can be suggested. The first is that five- or seven-repeat variants were present in the wild populations, although they are now absent in the same geographical locations. However, this hypothesis suppose a heterogeneous distribution in the

https://doi.org/10.1017/S001667239900395X Published online by Cambridge University Press

past which contrasts with the current strong geographical structuring observed and its stability. The second hypothesis is that these new variants in the laboratory lines have arisen either by mutation or by contamination events. Mutations events would imply a high rate of mutation, which would contrast with the stable nature of this microsatellite, and especially with the stability of the geographical distribution. Thus, we suppose that most of the anomalous observations result from contamination events. This is all the more likely because strains from the same locality can belong to different classes depending on the year of capture, even if the two samples were collected close together (e.g. France – Marseillan (1965): [5] class, 1978 and 1983: [3,5,7] class; Israel – Qiryat anavim (1981): [3,7,9] class, 1983: [3] class).

Discussion

To investigate the invasion dynamics of *hobo* elements in natural populations we used a new molecular approach that exploits the existence of a polymorphism with regard to the number of TPE repeats in the S region of the *hobo* element. Thus, 101 current populations and 63 old strains derived from wild flies were analysed by PCR amplification to determine which types of elements they harbour. Such PCR amplifications are only effective with euchromatic elements, as no amplification was observed in E strains, although such strains harbour some vestigial sequences (unpublished data). Natural populations currently show a geographical variation with regard to their numbers of TPE repeats, in which the majority of populations are monomorphic, harbouring only elements with three TPE repeats. These populations are distributed all over the world except in limited areas where polymorphic populations show clear geographical structuring. Such a geographical distribution could account for the dynamics of *hobo* elements.

The TPE repeats can be considered to be microsatellites (Tautz, 1993) according to both their length and the variation in their number. However, some of their particularities need to be stressed. With a size of 9 bp, TPE repeats are longer than most microsatellites studied in *D. melanogaster* (Goldstein & Clark, 1995; Schlötterer *et al.*, 1997). They have only a moderate degree of repetition, with the number of repeats varying between three and ten. The frequencies of the different types of alleles follow an unusual distribution: elements with an even number of TPE repeats seem to be rare. Indeed, elements with four TPE repeats were never noted and elements with six, eight or ten repeats were observed only in a few laboratory lines, and never in the natural populations analysed in this study. On the contrary, elements with an odd number of TPE repeats were frequently found: elements with three repeats are common to all current natural populations; elements with five and seven repeats are found in several populations in South America, Western Europe and/or Africa. These differences in the frequencies of the different variants could result either from the reliability of the mechanisms involved in the mutation of the number of repeats or from a selective advantage conferred on elements carrying an odd number of TPE repeats. Indeed, an obvious characteristic of TPE repeats is that they are localized in the ORF1 of the *hobo* element, thought to encode the transposase (Calvi *et al.*, 1991). Thus they cannot simply be considered as neutral, even if previous studies do not show clear evidence of a relationship between the S region polymorphism and a *hobo* element's activity (Bazin & Higuet, 1996; Bonnivard *et al.*, 1997).

The TPE repeats can be considered as useful molecular markers of *hobo* element dynamics. Indeed, this polymorphic microsatellite appears to be stable, as argued by two observations: (1) the TPE status of the reference strains has not changed for at least 5 years, (2) isofemale lines derived from natural population do not show intra-line variability (unpublished data). Moreover, TPE repeats are original microsatellites as they are localized on a transposable element, present in several copies in the genome. Thus, TPE variants are stable, non-Mendelian loci useful as diagnostic markers to monitor *hobo* element dynamics.

Regarding the geographical distribution of the number of TPE repeats in current natural populations (Fig. 4), several scenarios for the invasion of *D. melanogaster* by *hobo* elements can be proposed. A first hypothesis is that the variability observed between populations might be explained by the simultaneous complete invasions of most natural populations by the different types of elements. During this process, some types of element could have been lost randomly by genetic drift. This would have led to different classes being observed in different areas. However, this type of process should have created different patches of populations, leading to a very heterogeneous geographical distribution compared with the one we observed. A second hypothesis would involve separate invasions by different types of elements, which have come to be found together in the current polymorphic areas. However, this would imply multiple origins for the *hobo* elements and contrasts with the fact that all the monomorphic current populations described and most of the old monomorphic strains belong to class [3]. Finally, the high levels of polymorphism observed in Western Europe could indicate that this region was the origin of the worldwide invasion of *D. melanogaster* by the *hobo* elements. However, this European origin contrasts with the historical pattern described by Pascual & Periquet (1991) and confirmed by

Boussy & Daniels (1991). These authors showed that H strains (which harbour full-size elements) were prevalent first in America after 1950, and only about 10 years later in Western Europe; the latest E strain that they found was collected in 1970 in Italy.

A model of *hobo* element invasion should account for three results: (1) the preponderance of *hobo* elements with three TPE repeats, which characterize all current populations and are in the majority in most of the polymorphic populations, as shown by differences in the intensity of the signals obtained (Fig. 3), (2) the structure of current natural populations and (3) the presence of *hobo* elements first in America and then in Europe. According to these data, we proposed a *hobo* element 'two-stage' invasion, which occurred as follows. First, there was a complete invasion by *hobo* elements with three TPE repeats. One possible starting point for this initial invasion could be Central America as suggested by Periquet *et al.* (1989*a*). However, in the absence of additional temporal data, especially from old African populations, the geographic origin of the invasion remains unknown. This scenario also agrees with the supposed occurrence of horizontal transfer of *hobo* transposable elements within the *melanogaster* species complex (Periquet *et al*., 1990; Boussy & Daniels, 1991; Simmons, 1992). Indeed, the analysis of 10 current natural populations of *D. simulans* has revealed only elements with three TPE repeats (unpublished data), supporting the ancestral status of these elements. Second, the initial invasion by *hobo* elements with three TPE repeats might have been followed by a new invasion stage involving other types of *hobo* elements with five or seven TPE repeats. Such elements could have appeared as a result of mutation of pre-existing elements with three repeats, according to models described for microsatellites or minisatellites (for a review see Debrauwere *et al*., 1997).

Under this hypothesis, the restricted distribution of elements with five and seven TPE repeats contrasts with the complete invasion of natural populations by elements with three TPE repeats. These latest are autonomous (Blackman *et al*., 1989) and they can present different types of activities in the *hobo* system of hybrid dysgenesis (Bazin & Higuet, 1996; Bonnivard *et al*., 1997). Thus, these active elements could have increased in number via duplicative transposition and invaded natural populations of *D. melanogaster*, following fly migrations. Considering the five or seven repeat variants, they invaded populations already harbouring *hobo* elements with three TPE repeats. To explain this situation, we must assume that elements with five or seven TPE repeats are mobilized in *trans* by elements with three TPE repeats and have an advantage in transposition or regulation. However, in this situation elements with five or seven repeats could be widespread in populations, as is observed with the *KP* elements in the PM system (Anxolabéhère *et al*., 1988, 1990; Bonnivard & Higuet, 1999). Another explanation is that elements with five or seven TPE repeats are autonomous and not completely regulated by elements with three TPE repeats. An argument is that elements with five TPE repeats are able to product active transposase (Bazin & Higuet, 1996; unpublished data).

Data from old laboratory strains do not allow us to follow up the invasion of the different types of *hobo* elements. Due to the presence in France of polymorphic populations of the [3,5,7] class as early as the 1960s, it is not possible to observe the potential distribution corresponding to the complete invasion by elements with only three repeats, which could represent the first stage of the invasion. The oldest French population (Kerbiniou – 1960) has the same status as current strains from the same region (Brest – 1997). Elements with five or seven TPE repeats are thus ancient, appearing in a period just following the supposed invasion of European populations (the end of the 1950s; Pascual & Periquet, 1991; Boussy & Daniels, 1991). It thus suggests that the two supposed stages of the invasion were very close in time.

In our two-step invasion model, the geographical origin of the second invasion remains uncertain. The assumption of a high rate of mutation would favour the hypothesis according to which elements with five and seven TPE repeats have arisen numerous times in wild populations, possibly on different continents (South America, Western Europe and Africa). Thus five and seven repeat variants would have multiple origins. On the contrary, if new mutation are rare, as we supposed, the hypothesis of a single origin is more likely. These variants might have appeared in only one region, from where they invaded natural populations of other regions. If Europe is considered to be this single origin, it seems difficult to explain why such an invasion stopped in Central Europe, when it spread through South America or Africa. Moreover, the monomorphic North African populations, which show only elements with three TPE repeats, are found between the European and African polymorphic populations. Such a geographical distribution does not support the hypothesis of a single origin either in Europe or Africa. Thus South America seems to be the best candidate for a single origin of the invasion. This is all the more likely if *hobo* elements are ancient in this region, considered as the possible origin of the invasion by elements with three TPE repeats.

To confirm our hypothesis on a single geographical origin, we are currently analysing *hobo* sequences. Moreover, to determine whether some types of elements are present in current natural populations, even if they cannot be detected in mass cultures, we are studying some populations using isofemale lines. This should allow us to determine the components of

the polymorphism observed in some populations. In particular it could explain a part of the variability observed between results of PCR amplifications for a given strain. This study will also enable us to obtain a better knowledge of the process of invasion. Furthermore, the study of *hobo* dynamics also requires data on the potential activity of the different types of elements differing in the number of TPE repeats.

We are grateful to M. Cobb for reading the manuscript, as well as to anonymous referees for their helpful comments. We wish to thank all those who supplied us with strains derived from natural populations: A. Adoutte, A. Atlan, D. Anxolabéhère, E. Aspect, M. F. Aspect, J. Barrat, V. Benassi, F. Bonhomme, P. Bonnivard, V. Bonnivard, C. Breyton, J. Brookfield, D. Bulcke, P. Capy, M. L. Cariou, M. Cazemajor, J. David, F. Debets, M. Delattre, X. Draye, G. Dubois, C. Dutech, A. Fleuriet, L. Haring, H. Izaabel, L. Kraaijeveld, R. Kuras, D. Lachaise, C. Lagane, M. Lehmann, B. Llorente, T. Mackay, H. Mercot, W. Miller, C. Montchamp, S. Netter, C. Petit, Y. Pierre, D. Poinsot, J. L. Popot, H. Quesneville, M. Raymond, J. Roote, J. Shykoff, J. Sved, Y. H. M. Svoboda, M. Veuille, F. Villate and the anonymous students from Paris 6 University, who spent some of their vacation fly hunting. This work was supported by ACC SV3 no. 9503079.

References

Anxolabéhère, D., Kidwell, M. G. & Périquet, G. (1988). Molecular characteristics of diverse populations are consistent with a recent invasion of *Drosophila melanogaster* by mobile *P* element. *Molecular Biology and Evolution* **5**, 252–269.

Anxolabéhère, D., Hu, K., Nouaud, D. & Périquet, G. (1990). The distribution of the P–M system in *Drosophila melanogaster* strains from the People's Republic of China. *Genetique*, *Selection*, *Evolution* **22**, 175–188.

Bazin, C. & Higuet, D. (1996). Lack of correlation between dysgenic traits in the *hobo* system of hybrid dysgenesis in *Drosophila melanogaster*. *Genetical Research* **67**, 219–226.

Blackman, R. K., Koehler, M. M. D., Grimaila, R. & Gelbart, W. M. (1989). Identification of a fully-functional *hobo* transposable element and its use for germline transformation of *Drosophila*. *EMBO Journal* **8**, 211–217.

Bonnivard, E. & Higuet, D. (1999). Stability of European natural populations of *Drosophila melanogaster* with regard to the P–M system: a buffer zone made up of Q populations. *Journal of Evolutionary Biology* (in press).

Bonnivard, E., Higuet, D. & Bazin, C. (1997). Characterisation of natural populations of *Drosophila melanogaster* with regard to the *hobo* system, a new hypothesis on the invasion. *Genetical Research* **69**, 197–208.

Boussy, I. A. & Daniels, S. B. (1991). *hobo* transposable elements in *Drosophila melanogaster* and *D. simulans*. *Genetical Research* **58**, 27–34.

Brunet, F., Godin, F., Bazin, C., David, J. R. & Capy, P. (1996). The *mariner* transposable element in natural populations of *Drosophila teissieri*. *Journal of Molecular Evolution* **42**, 669–675.

Bucheton, A., Vaury, C., Chaboissier, M.-C., Abad, P., Pélisson, A. & Simonelig, M. (1992). I elements and the *Drosophila* genome. *Genetica* **86**, 175–190.

Calvi, B. R. & Gelbart, W. M. (1994). The basis for germline specificity of the *hobo* transposable element in *Drosophila melanogaster*. *EMBO Journal* **13**, 1636–1644.

Calvi, B. R., Hong, T. J., Findley, S. D. & Gelbart, W. M. (1991). Evidence for a common evolutionary origin of inverted repeat transposon in *Drosophila* and plants: *hobo*, *activator* and *Tam3*. *Cell* **66**, 465–471.

Capy, P., Bazin, C., Higuet, D. & Langin, T. (1997). *Dynamics and Evolution of Transposable Elements*, pp. 36–52. Austin, TX: R. G. Landes Bioscience.

Daniels, S. B., Chovnick, A. & Boussy, I. A. (1990). Distribution of *hobo* transposable elements in the genus *Drosophila*. *Molecular Biology and Evolution* **7**, 589–606.

Debrauwere, H., Gendrel, C. G. & Dutreix, M. (1997). Differences and similarities between various tandem repeat sequences: minisatellites and microsatellites. *Biochimie* **79**, 577–586.

Engels, X. R. (1989). *P* elements in *Drosophila*. In *Mobile DNA* (ed. D. E. Berg & M. M. Howe), pp. 437–484. Washington, DC: American Society of Microbiology.

Finnegan, D. J. (1989*a*). The I factor and I–R hybrid dysgenesis in *Drosophila melanogaster*. In *Mobile DNA* (ed. D. E. Berg & M. M. Howe), pp. 503–518. Washington, DC: American Society of Microbiology.

Finnegan, D. J. (1989*b*). Eukariotic transposable elements and genome evolution. *Trends in Genetics* **5**, 103–107.

Goldstein, D. B. & Clark, A. G. (1995). Microsatellite variation in North American populations of *Drosophila melanogaster*. *Nucleic Acids Research* **23**, 3882–3886.

Junakovic, N., Caneva, R. & Ballario, P. (1984). Genomic distribution of *copia*-like element in laboratory stocks of *Drosophila melanogaster*. *Chromosoma* **90**, 378–382.

Lim, J. K. (1988). Intrachromosomal rearrangements mediated by *hobo* transposons in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the USA* **85**, 9153–9157.

McGinnis, W., Shermoen, A. W. & Beckendorf, S. (1983). A transposable element inserted just 5′ to a *Drosophila* glue protein gene alters gene expression and chromatin structure. *Cell* **34**, 75–84.

Pascual, L. & Periquet, G. (1991). Distribution of *hobo* transposable elements in natural populations of *Drosophila melanogaster*. *Molecular Biology and Evolution* **8**, 282–296.

Periquet, G., Hamelin, M. H., Bigot, Y. & Lepissier, A. (1989*a*). Geographical and historical patterns of distribution of *hobo* elements in *Drosophila melanogaster* populations. *Journal of Evolutionary Biology* **2**, 223–229.

Periquet, G., Hamelin, M. H., Bigot, Y. & Kai Hu (1989*b*). Presence of the deleted element *Th* in Eurasian populations of *Drosophila melanogaster*. *Genetique*, *Selection*, *Evolution* **21**, 107–111.

Periquet, G., Hamelin, M. H., Kalmes, R. & Eeken, J. (1990). *Hobo* elements and their deletion-derivative sequences in *D. melanogaster* and in its sibling species *D. simulans*, *D. mauritiana* and *D. sechellia*. *Genetique*, *Selection*, *Evolution* **22**, 393–402.

Periquet, G., Lemeunier, F., Bigot, Y., Hamelin, M. H., Bazin, C., Ladevèze, V., Eeken, J., Galindo, M. I., Pascual, L. & Boussy, I. (1994). The evolutionary genetics of the *hobo* transposable element in *Drosophila melanogaster* complex. *Genetica* **93**, 79–90.

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

Simmons, G. M. (1992). Horizontal transfer of *hobo* transposable elements within the *Drosophila melanogaster* species complex: evidence from DNA sequencing. *Molecular Biology and Evolution* **9**, 1050–1060.

Schlötterer, C., Vogl, C. & Tautz, D. (1997). Polymorphism and locus-specific effects on polymorphism at micro-

satellite loci in natural *Drosophila melanogaster* populations. *Genetics* **146**, 309–320.

Streck, R. D., MacGaffey, J. E. & Beckendorf, S. K. (1986). The structure of *hobo* element and their site of insertion. *EMBO Journal* **5**, 3615–3623.

Tautz, D. (1993). Notes on the definition and nomenclature of tandemly repetitive DNA sequences. In *DNA Fingerprinting: State of the Science* (ed. S. D. J. Pena, R. Chakraborty, J. T. Epplen & A. J. Jeffreys), pp. 21–28. Basel: Birkhäuser Verlag.

https://doi.org/10.1017/S001667239900395X Published online by Cambridge University Press