DOI: 10.1534/genetics.106.064378

The Evolutionary History of *Drosophila buzzatii*. XXXVI. Molecular Structural Analysis of *Osvaldo* Retrotransposon Insertions in Colonizing Populations Unveils Drift Effects in Founder Events

María Pilar García Guerreiro[1] and Antonio Fontdevila

Grup de Biología Evolutiva, Departament de Genètica i Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Manuscript received August 3, 2006
Accepted for publication October 26, 2006

ABSTRACT

Previous work on transposable element distribution in colonizing populations of *Drosophila buzzatii* revealed a high frequency of occupancy in several chromosomal sites. Two explanatory hypotheses were advanced: the founder hypothesis, by which founder genetic drift was responsible, and the unstable hypothesis that assigns this unusual distribution to bursts of transposition toward some chromosomal sites. Here, we study the molecular structure of three euchromatic *Osvaldo* clones isolated from sites occupied at high (A4 and B9) and low frequency (B4) in colonizing populations, to test these hypotheses. Large insertions, duplications, and indels in the *Osvaldo* coding region and LTR were detected in the A4 clone and a truncated *Osvaldo* with many substitutions was found in the B9 clone. These altered sequences indicate that the two copies of this retroelement are precolonization insertions. Interestingly, the LTR of the A4 clone and the reverse transcriptase region of B9 show identical sequences in all colonizing populations indicating, most probably, that they are identical by descent. Moreover, *Osvaldo* is inserted at the same nucleotide site in all colonizing populations. On the other hand an almost identical LTR sequence, except by 1 base deletion, was found in the B4 clone compared to the canonical active *Osvaldo* element. These results suggest that *Osvaldo* copies in highly occupied sites are, most probably, identical by descent and strongly favor the founder hypothesis. On the other hand, low-insertion-frequency sites could represent recent transposition events. This work emphasizes the importance of molecular population studies to disentangle the effects of genetic drift and transposition in colonization.

RETROTRANSPOSONS are the most abundant and widely distributed class of transposable elements (TEs) in the eukaryote genome (Berg and Howe 1989). Drosophila contains at least 23 families of LTR retrotransposons and retroviruses and 11 families of non-LTR retrotransposons, representing ~5–10% of the Drosophila genome (Bowen and Mcdonald 2001; Kaminker *et al.* 2002; Lerat *et al.* 2003).

A widespread agreement from population studies is that TE occupancy rate per site of insertion is low in Drosophila (Charlesworth and Langley 1989; Biémont 1992). Yet, some cases have been reported of higher than expected insertion rates in colonizing populations (Labrador *et al.* 1998; Vieira *et al.* 1999). Two main mechanisms were advanced to explain this rate increase: first, the activation of transposition following colonization due to either environmental stress or hybrid release in crosses between invaders and resident flies (Evgen'ev *et al.* 1982, 2000; Labrador *et al.* 1998; Vieira *et al.* 1999; Biémont *et al.* 2003) and, second, a drift effect that increases the originally low-frequency occupancy rate (García Guerreiro and Fontdevila 2001). There are few experimental designs that can distinguish clearly between these hypotheses, but those based on theoretical models point toward the founder hypothesis as the most likely (Labrador *et al.* 1998). Yet, these models are valid only for steady-state populations. Since colonizing populations may be in a nonequilibrium state, this casts serious doubts on their applicability. In addition, support for the founder model does not discard absolutely the possibility of colonization release of transposition. Here, we address the problem focusing on the molecular population genetics of TE insertions by studying sequences of high- and low-occupancy sites.

Osvaldo is a LTR retrotransposon described from *Drosophila buzzatii* (Labrador and Fontdevila 1994; Pantazidis *et al.* 1999), a South American species that colonized the Old World a few centuries ago and Australia in more recent times (David and Tsacas 1980; Fontdevila *et al.* 1981; Barker 1982). Previous studies on the natural distribution of this element showed a low-occupancy rate for *Osvaldo* in original populations from Argentina. Conversely, occupancy in colonizing populations from the Iberian Peninsula is higher, showing a bimodal distribution, with a few highly occupied

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. DQ139455–DQ139462 and DQ861339–DQ861345.

[1]*Corresponding author:* Grup de Biología Evolutiva, Departament de Genética i Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. E-mail: mariapilar.garcia.guerreiro@uab.es

sites, but with the majority of sites of low occupancy (Labrador *et al.* 1998; García Guerreiro and Fontdevila 2001). The colonization process of *D. buzzatii* could explain this bimodal distribution pattern by means of a founder event followed by genetic drift. We hypothesized that TE copies in the Old World populations are characterized now by high-occupancy frequencies due to the ensuing genetic drift experienced by founders and those of low occupancy correspond to recent, after colonization, transpositions. This founder hypothesis contrasts with the hypothesis that colonization instability releases TE transposition to specific "hot spots" of high TE occupancy. This unstable hypothesis has been supported mainly by data from natural populations of *D. simulans* (Biémont *et al.* 2003), *D. melanogaster* (Vieira *et al.* 2002), and Hawaiian Drosophila species (Wisotzkey *et al.* 1997) in which a significant increase in TE copy number is detected in colonizing populations. Although possible, it is difficult to envisage a process that affects equally all the TEs studied (all the TEs do not need to be concerned, only some of them are sufficient), but population- or species-dependent factors are most likely. Thus, host factors seem an unlikely, though possible, cause in the case of *D. buzzatii* colonizing populations since the effect involves two very different elements, *Osvaldo* (Labrador *et al.* 1998) and *blanco* (García Guerreiro and Fontdevila 2001).

Fortunately, the two hypotheses generate very different predictions for TE molecular structure and insertion patterns. Under the unstable hypothesis elements in highly occupied sites must be newly transposed, fully active elements, showing no sequence differences from the canonical active sequence, if the colonization is recent. On the other hand, if the founder hypothesis is correct, copies in sites of high occupancy must be older than the colonization event. These high-occupancy-site copies should thus be more diverged than "newer" copies from low-occupancy sites. These latter copies should show canonical active sequences because there has been insufficient time for substitutions after recent colonizations. Moreover, under the founder hypothesis insertion site structure at the nucleotide level should be identical among populations, something unlikely under the unstable hypothesis. To decide between the hypotheses, we sequenced two euchromatic clones of *Osvaldo* from high-occupancy sites and one from a low-occupancy site. Our results are more consistent with the founder hypothesis and indicate the importance of using TE structure and insertion pattern to assess the role of drift in founder events in relation to transposition release caused by the colonization processes.

MATERIALS AND METHODS

Drosophila stocks: *Ca107J9* is an inbred line started from a population collected in Carboneras (Spain) in 1996 (García Guerreiro and Fontdevila 2001) and maintained by brother–sister matings for 15 generations. At the time of the experiment this line exhibited *Osvaldo* insertions at chromosomal sites 2F1d, 2F4a, 2D5a, (E4a) 3F2b, 3B5f, (4C2g), (4E1b), 5A4b, (5B1f), and (5C2a)—parentheses indicate polymorphic sites.

The control line *63 42/7 F81* was originated from *D. buzzatii* stock Bu42 (28/7) collected in Maimara (Argentina) and maintained by brother–sister matings during the initial 63 generations and kept thereafter by mass culturing (Labrador and Fontdevila 1994; Labrador *et al.* 1998). This line is devoid of *Osvaldo* euchromatic insertions.

Adeje, Puebla del Río, Cullera, Colera, Sanlúcar, Portosín, Llano del Moro, Castelldefels, and Túnez are laboratory stocks derived from Spanish natural populations collected at these localities (except Túnez, collected in Tunisia). These stocks were maintained by mass culturing. To determine the exact position of the A4 clone, different isofemale lines from the five first stocks have been established, numbered, as shown in Table 1, and assayed for *Osvaldo* insertions by both *in situ* hybridization and PCR. The stocks of Cullera, Carboneras, Portosín, Llano del Moro, Castelldefels, Túnez, and Sanlúcar were used to sequence the B9 clone because these are the only available stocks (except Colera) having the 5A4b site .

Genomic library: Total genomic DNA was extracted according to Piñol *et al.* (1988) from the *D. buzzatii* inbred Ca107J9 line. DNA was partially digested with *Sau*3A and treated with RNAse and alkaline phosphatase according to Maniatis *et al.* (1982). Then, 0.3 μg of the partially digested genomic DNA was ligated with 1 μg of λ-DASH II arms predigested with *Bam*HI. Recombinant DNA was packaged using Gigapack III XL extracts (Stratagene, La Jolla, CA).

Hybridization and clone selection: The genomic library was screened with a probe (Osv) containing a 6.4-kb *Eco*RI *Osvaldo* fragment (Pantazidis *et al.* 1999) inserted in the Bluescript KS+ plasmid. This fragment comprises the complete canonical element except the LTRs. Probes were labeled with digoxigenin 11-dUTP (Roche, Indianapolis), using a random primer reaction. Positive clones were isolated and used for *in situ* hybridization of polytene chromosomes from the laboratory *63 42/7 F81* stock devoid of *Osvaldo* euchromatic insertions. Because each hybridization signal in this stock is attributed to *Osvaldo* flanking DNA, this allows the mapping of the *Osvaldo* flanking region in each *Osvaldo* insertion.

***In situ* hybridization:** Polytene chromosome squashes from salivary glands of third-instar larvae were prepared as described in Labrador *et al.* (1990) and *in situ* hybridized with Osv (6.4-kb *Eco*RI *Osvaldo* fragment) and Fla (a 1.4-kb fragment of the 3′-flanking region of the A4 *Osvaldo* clone) probes using procedures described by Labrador and Fontdevila (1994). Prehybridization and posthybridization washes were done following the protocol of Schmidt (1992) published by Roche.

PCR amplifications: PCR reactions were carried out in a final volume of 50 μl, including 1× activity buffer (Ecogen), 1.6 mM $MgCl_2$, 0.2 mM of each dNTP (Roche), 0.4 μM primer (Roche), template DNA (~10–20 ng), and 0.04 units of Taq polymerase (Ecotaq from Ecogen)/μl. Amplifications were run in an MJ Research (Watertown, MA) thermocycler programmed as follows: 5 min preliminary denaturation at 94°; 30 cycles of 45 sec at 94° (denaturation), 45 sec at specific PCR annealing temperatures, and 1.5 min at 72° (extension); and a final extension for 10 min at 72°. The samples were then stored at 4°.

PCR products were gel purified with a Geneclean kit (BIO 101, Vista, CA) and either cloned or directly sequenced with PCR primers and new primers designed to have big overlapping segments and diminishing sequencing errors.

Sequencing of *Osvaldo* insertions and flanking sequences: Sequencing was performed in one strand by Sanger's

dideoxynucleotide chain-termination method (Sanger *et al.* 1977), using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA) with an automatic ALF sequencer (Pharmacia Biotech, Piscataway, NJ).

Sequences of clones from the Ca107J9 inbred line were obtained by sequencing different overlapping restriction fragments as in the A4 clone, sequencing directly the phage clone as in the B9 clone, or by sequencing only the band that displays hybridization signal with an *Osvaldo* probe (B4 and P1 clones). Clone sequence comparisons, between different population stocks, were done by amplification of a 1.6-kb PCR fragment. In the case of the A4 clone PCR includes the A4 *Osvaldo* 3′ LTR plus 249 bp of the flanking region. For this, we used a specific primer (orf3) from the A4 ORF3 region (5′-GCTGGTGAG CAGCGCGGCGCTGTT-3′), and another (popnat3) from the 3′-flanking region (5′-ATCAAAAGTTTAGGCGGGTGC-3′). The B9 amplified region includes 577 bp of the reverse transcriptase (RT) region of *Osvaldo* plus 1027 bp of the flanking genomic DNA. Primers used were B9d (5′-GTCCTTTCTGA GGCACGAAT-3′) and B9r (5′-GCACTAAAGACGGTATCC-3′) specific for the RT and flanking DNA regions, respectively.

DNA sequences were analyzed using BLAST programs to establish paired alignments (bl2 seq option) or to search sequence similarities to DNA (BLASTN option). When multiple sequence alignments were required the CLUSTAL W program (Higgins *et al.* 1996) was used. Alignments were edited using the GENEDOC program (Nicholas *et al.* 1997; http://www.psc.edu/biomed/genedoc).

Age of *Osvaldo* 2F4a and 5A4b insertions was estimated as in Kapitonov and Jurka (1996), using the synonymous Drosophila substitution rate of 0.016 substitutions per site per MY (Li 1997). Pairwise comparisons with the canonical active *Osvaldo* were performed and the number of substitutions per site was computed with the Kimura two-parameter model using the MEGA v.2.1 program (Kumar *et al.* 2001).

Determining the position of the A4 *Osvaldo* clone: In contrast to B9 and B4 clones, which showed only one chromosomal position in the empty stock, several positions of the A4 clone were detected. In this case, different inbred lines of each population were checked simultaneously by PCR (amplifying the 1.6-kb band) and by *in situ* hybridization with *Osvaldo* (Osv) and flanking (Fla) probes, to determine the specific site that corresponds to the isolated clone.

RESULTS

Twenty-two positive *Osvaldo* clones isolated from the *Ca107J9* library were used for *in situ* hybridization on polytene chromosomes from the *Osvaldo*-free *63 42/7 F81* stock. Five of them mapped to a euchromatic site. We selected one heterochromatic (P1) and three euchromatic clones: A4 and B9 (located at the high-occupancy sites 2F4a and 5A4b, respectively) and B4 (at the low-occupancy 2F1c site). 2F4a and 5A4b sites are considered high-occupancy sites because a high insertion frequency was observed in colonizer populations (38 and 54% in Carboneras and 18 and 38% in Sanlúcar, respectively) compared to original ones (see Labrador *et al.* 1998; García Guerreiro and Fontdevila 2001). The remaining isolated clones were not analyzed because they showed very weak hybridization signal and we supposed that they contained mostly genomic DNA plus a very small portion of the *Osvaldo* element.

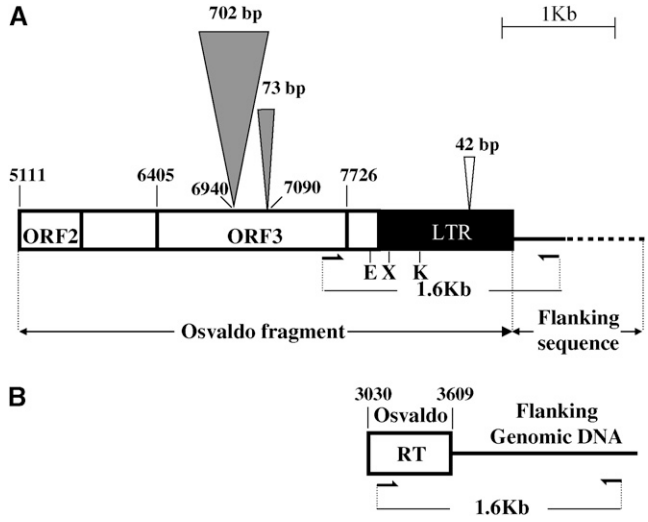

Figure 1.—(A) Structure and restriction map of the *Osvaldo* fragment included in the A4 euchromatic clone (GenBank accession no. DQ139455). Shaded triangles depict insertions and the open triangle depicts a duplication. Abbreviations of restriction enzymes in the map are as follows: E, *Eco*RI; X, *Xba*I; K, *Kpn*I. (B) Structure of an *Osvaldo* fragment included in the B9 euchromatic clone (GenBank accession no. DQ861339). Numbers at the top indicate nucleotide positions in the canonical *Osvaldo* sequence. Arrowheads indicate positions of primers used to amplify the *Osvaldo* fragment plus the flanking region.

Structure of A4 and B9 high-insertion-frequency clones: The 4705-bp A4 clone (Figure 1) comprises an *Osvaldo* 3′-end fragment plus a 394-bp flanking region. Alignment with the canonical active *Osvaldo* sequence (Pantazidis *et al.* 1999) revealed nucleotide substitutions, indels, and two insertions (702 and 73 bp long) in the ORF3 region and one 42-bp imperfect duplication (from which 30 bp corresponds to a perfect duplication) plus four substitutions in the 3′ LTR. The A4 flanking region is a repetitive sequence because it maps by *in situ* hybridization, using a flanking probe, to several dispersed sites (2F1d, 2F4a, 2E4a, 4C2g, and 5B1) in the Ca107J9 line.

The B9 sequenced clone (1.7 kb) shows at the 5′ end 577 bp corresponding to the *Osvaldo* RT region followed by genomic DNA, indicating that this *Osvaldo* insertion is deleted. Alignments to the canonical *Osvaldo* showed seven-nucleotide substitutions and three-nucleotide deletions.

To analyze the two cloned insertions in several colonizing populations we sequenced two 1.6-kb PCR fragments specific to A4 (Figure 2) and B9 clones. The A4 PCR fragment comprises the A4 *Osvaldo* 3′ LTR (Osv) plus a 249-bp fragment (Fla) from the flanking region. The B9 PCR fragment includes 577 bp of the *Osvaldo* RT region plus 1027 bp of the flanking genomic DNA.

We observed that these 1.6-kb sequences are identical in the different independently analyzed populations for A4 and B9 (Figure 3). Interestingly, the Ca107J9 42-bp

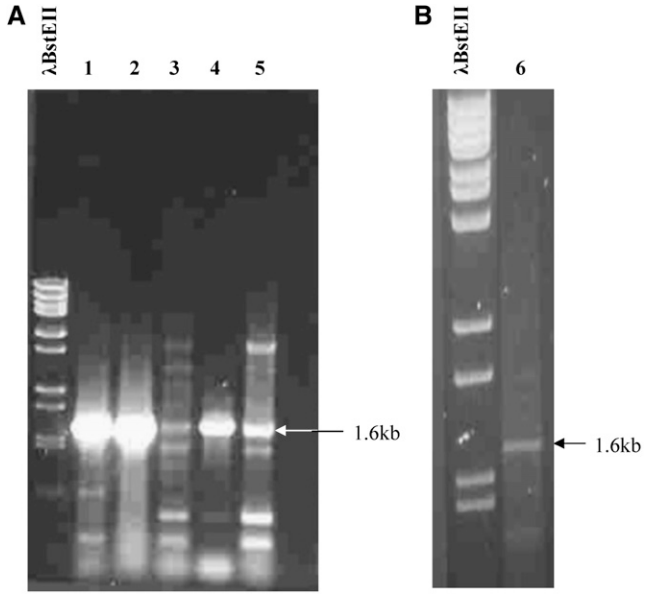

Figure 2.—PCR mass amplifications of an A4 1.6-kb fragment (LTR plus a flanking region) inserted in the 2F4a site in six natural populations (see text for details). The first lane of A and B corresponds to the size-calibrated DNA from *Bst*EII. Lane numbers indicate analyzed populations as follows: 1, Carboneras; 2, Adeje; 3, Colera; 4, Cullera; 5, Puebla del Rio; 6, Sanlúcar. The arrow indicates the 1.6-kb fragment.

duplication and the seven-nucleotide substitutions of A4 and B9, respectively, could be used as markers of colonization, if these copies are rare in original populations, as previous surveys suggest. Thus, we conclude that these A4 and B9 sequences have a common origin derived, most probably, from a founder individual.

The A4 ORF3 region shows several features. First, the 702-bp insertion is flanked by two 10-bp sequences, identical except by one nucleotide [GAGGA(T)GACG], which suggests that the 702-bp insertion is a TE insertion. This insertion shows a high nucleotide similarity (Figure 4) with a sequence, from position 3204 to 4909, present in an *Osvaldo* heterochromatic clone (cDb314) described by Labrador and Fontdevila (1994). Second, alignments of an A4 ORF3 putative protein product with the active *Osvaldo env* protein revealed a high sequence similarity (Figure 5). We believe that this product might correspond to an ancient degenerated *env* protein containing several stop codons.

Estimation of the *Osvaldo* integration time in the 2F4a gives different estimates depending on the region considered, 0.2 MY for LTR, 1.2 MY for the ORF2 region, and 3MY for the *env* region. These differences are due to the different substitution rate in these regions; for example, it is known that the *env* region usually presents the highest substitution rate (Preston and Dougherty 1996). Estimation of *Osvaldo* integration time from the RT region in the 5A4b site is the same as that for the 2F4a *Osvaldo* LTR. All in all, the estimates of divergence indicate that these *Osvaldo* insertions are much older than the colonization event, which occurred not earlier than 300 years ago (Fontdevila 1989, 1991).

Cytological position of the A4 clone: *In situ* hybridizations of the A4 clone on the *Osvaldo*-free stock showed several dispersed hybridization signals, leading to the conclusion that *Osvaldo* is flanked by a repetitive sequence. To detect the exact positions of this clone we combined two methodological approaches: *in situ* hybridization and PCR. First we analyzed the insertion profile of *Osvaldo* A4 with the Osv and Fla probes from this region, using isofemale lines established from each colonizing population by brother–sister matings. Simultaneously, PCR amplifications of each line were performed with A4-specific primers (see materials and methods) to obtain the 1.6-kb fragment previously obtained in mass amplifications. Comparisons between PCR and *in situ* hybridization results allow us to distinguish between A4 and other *Osvaldo* insertions. Results in Table 1 show that amplification is observed only in isofemale lines where *Osvaldo* and its flanking sequence coexist at the 2F4a position. In Colera and Sanlúcar lines, where no 2F4a insertions were observed, in isofemale lines individually analyzed, no amplification was obtained. The fact that some isofemale lines show no 2F4a insertion is probably due to sampling because these lines were established from a single fertilized female, but in mass amplifications (Figure 2) we used DNA from >100 individuals. In fact, mass amplifications give a 1.6-kb band with different intensities depending on the insertion frequency in each stock. Carboneras shows a very strong band intensity and is the population that presents the highest insertion frequency in the 2F4a site (Labrador *et al.* 1998). In the same way Sanlúcar, showing a faint band only after a PCR reamplification process, shows a lower insertion frequency than Carboneras. The insertion frequency of the remaining populations can be deduced from the insertion profile of each group of isofemale lines (Table 1), and populations presenting a weak amplification band correspond to the lines that host very few or no insertions in the 2F4a site. However, when we analyzed data from natural colonizing populations, the 2F4a site was always segregating at high frequency. For example, two independent samples taken in Carboneras (Labrador *et al.* 1998; García Guerreiro and Fontdevila 2001) and separated by 5 years showed very similar 2F4a insertion site frequencies (0.31 and 0.38).

In addition to the 2F4a site, other chromosomal sites (2F1d, 4C2g, and 5A4b) also show cooccurrence of *Osv* and *Fla* in some lines (Table 1). However, we obtain a 1.6-kb band amplification product only when the 2F4a site hybridizes with *Osv* and *Fla* probes simultaneously. We conclude that those sites have a structure different from the A4 clone structure and that they show the existence of either (1) an ancient coinsertion of *Osvaldo* and the flanking repetitive sequence or (2) an *Osvaldo* insertion preference for sites occupied by that repetitive sequence.

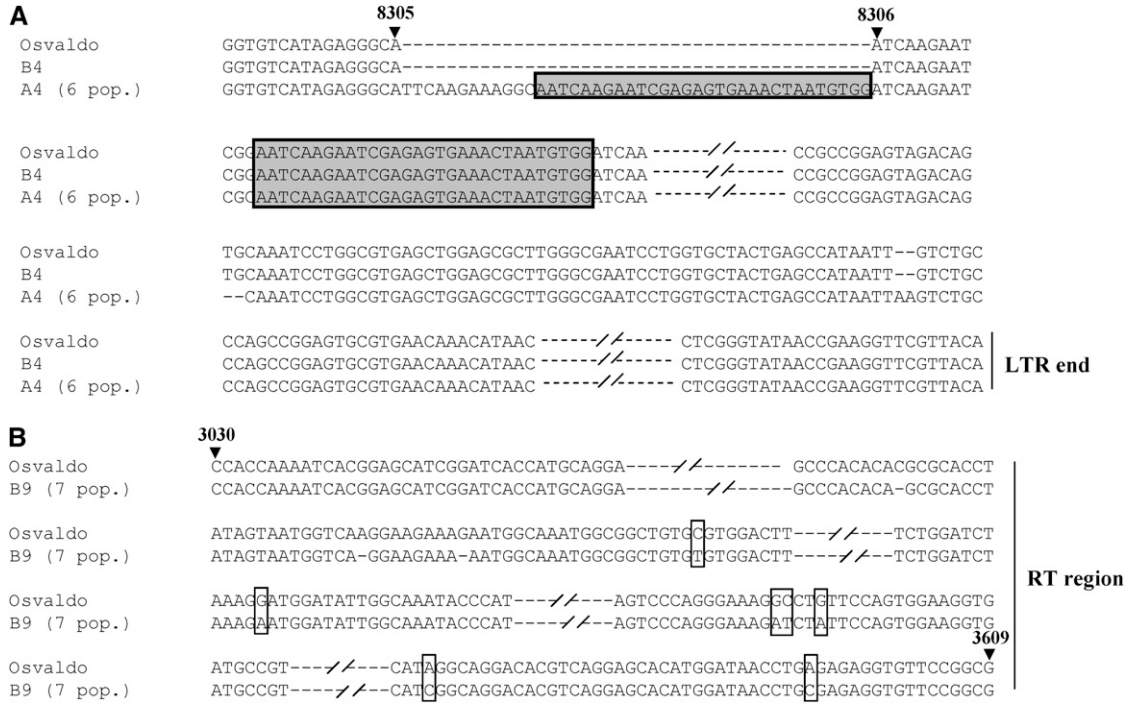

FIGURE 3.—Multiple-sequence alignments of a canonical *Osvaldo* and (A) LTRs of the *Osvaldo* B4 clone (GenBank accession no. DQ139461) and A4 clone from six colonizing populations (GenBank accession nos. DQ139456, DQ139457, DQ139458, DQ139459, DQ139460, and DQ139455 included in the A4 clone). Shaded boxes indicate regions that correspond to a perfect duplication. (B) RT region of the *Osvaldo* B9 clone from seven colonizing populations (GenBank accession nos. DQ861339, DQ861340, DQ861341, DQ861342, DQ861343, DQ861344, and 861345). Open boxes indicate nucleotide substitutions in the B9 clone. Arrowheads and numbers delimit the A4 duplicated region and the beginning and the end of the B9 clone, relative to the *Osvaldo* canonical sequence.

The B4 clone at the 2F1c site and the P1 heterochromatic clone: Two other clones were obtained and compared to the A4 clone. The 13-kb B4 clone contains mainly genomic DNA plus one *Osvaldo* LTR. The sequenced 1100-bp fragment includes a 5′ LTR identical to the active *Osvaldo*, except for a 1-bp deletion (Table 2), plus 56 bp of the *Osvaldo* 5′ region. The flanking genomic DNA hybridizes only at the 2F1c site in the *Osvaldo*-free line, a very low-occupancy site in natural populations. We think that the B4 clone could represent a recent *Osvaldo* insertion because of the high sequence similarity to the canonical active *Osvaldo* element (Figure 3) and the absence of large duplications and deletions.

The P1 clone, including a noncanonical *Osvaldo* element and a flanking region, hybridizes only at centromeric regions of the *Osvaldo*-free line. A 4.6-kb (*Xba*I) product, that shows hybridization signal with an *Osvaldo* probe, includes 1712 bp of flanking DNA plus 2918 bp of *Osvaldo*. P1 *Osvaldo* differs from the canonical *Osvaldo* by the presence of a duplication, an insertion, and a deletion (21, 8, and 322 bp, respectively) in its LTR (Figure 6) and an ORF1–ORF2 region with a 257-bp deletion. The flanking sequence shows no homology with database sequences, but a putative ORF of 522 nucleotides found in this sequence encodes a predicted polypeptide with 54% identity to a *gag-pol* polyprotein precursor of the *Max* retrotransposon of *D. melanogaster*. This retrotransposon exhibits preferential insertion in heterochromatin (MARSANO *et al.* 2004), again indicating the heterochromatic peculiarities of the flanking region of this clone.

DISCUSSION

D. buzzatii colonizing populations show a high insertion frequency of *Osvaldo* and *blanco* TEs in specific chromosomal sites (LABRADOR *et al.* 1998; GARCÍA GUERREIRO and FONTDEVILA 2001), compared to source populations. These results suggest that this TE distribution pattern is associated with colonization either by an increase in the rate of transposition (unstable hypothesis) or by genetic drift (founder hypothesis). To decide between both hypotheses, we sequenced *Osvaldo* clones from heterochromatic (P1) and euchromatic regions corresponding to high- (A4 and B9) and low-occupancy (B4) sites. The P1 heterochromatic clone represents a copy of an inactive *Osvaldo* element that has accumulated many structural changes over time. It is usually thought that heterochromatic copies of retrotransposons represent the end point of an ancient invasion and that heterochromatic regions serve as a sink for ancient, defective copies (SHEVELYOV *et al.* 1989; VAURY *et al.*

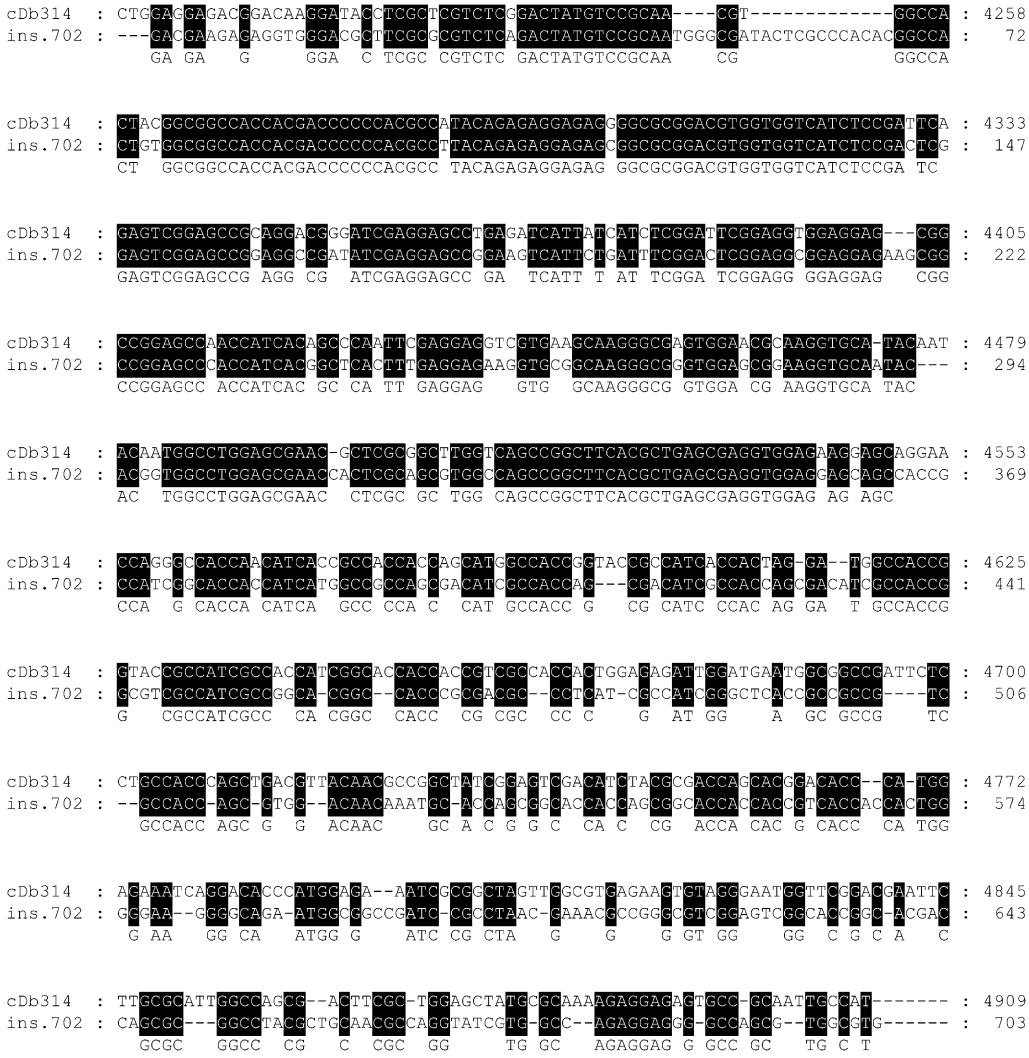

Figure 4.—Alignment of the 702-bp insertion into the *Osvaldo* A4 clone with a putative *env* degenerated sequence from an *Osvaldo* heterochromatic clone (cDb 314) described by Labrador and Fontdevila (1994).

1989; Pimpinelli *et al.* 1995; Dimitri 1997). The clones B4, A4, and B9, are different even though they all are euchromatic. The B4 LTR differs only by one nucleotide deletion from the canonical active *Osvaldo* LTR (Pantazidis *et al.* 1999), whereas the A4 LTR differs by a 42-bp duplication, four substitutions and insertions in the *env* region, from the canonical sequence. Otherwise the B9 clone includes only the *Osvaldo* RT region differing by seven nucleotide substitutions and three deletions from the canonical *Osvaldo* sequence. In view of their accumulated insertions and nucleotide substitutions, A4 and B9 clones host degenerated inactive *Osvaldo* elements and represent precolonization insertions. In contrast, the B4 clone may include a recent insertion (postcolonization) because its LTR sequence is identical to the active *Osvaldo* LTR except for a 1-bp deletion. This deletion is likely due to a reverse transcriptase replication error, as this process has been reported as a frequent event in retroelement replication (Preston and Dougherty 1996). When LTR retrotransposons insert into genomes, they are generally full-sized elements that, over time, change by gradual accumulation of insertions, duplications, substitutions, and small deletions (Petrov 2002). Defective elements often exhibit large changes, such as the deletion of ORFs and LTRs, that impair their transposition (Kaminker *et al.* 2002). In our case we observed, in all analyzed colonizing populations, the same 42-bp duplication in the A4 LTR and seven substitutions in the B9 RT, never found in other *Osvaldo* clones, and an identical *Osvaldo* flanking region inserted in that chromosomal site. It is difficult to explain this result under the unstable hypothesis that predicts similar, but not identical, insertion sites of intact canonical elements (at the nucleotide level). As far as we know, until now no data have been reported documenting an identical insertion site in different, natural colonized populations. However, some cases of similar, but not identical insertion sites have been reported for the *Penelop*e element in *D. virilis* and *D. lummei* (Evgen'ev *et al.* 2000) and for *Zam* and *Idefix* within the white locus of the RevI strain of *D. melanogaster* (Conte *et al.* 2000).

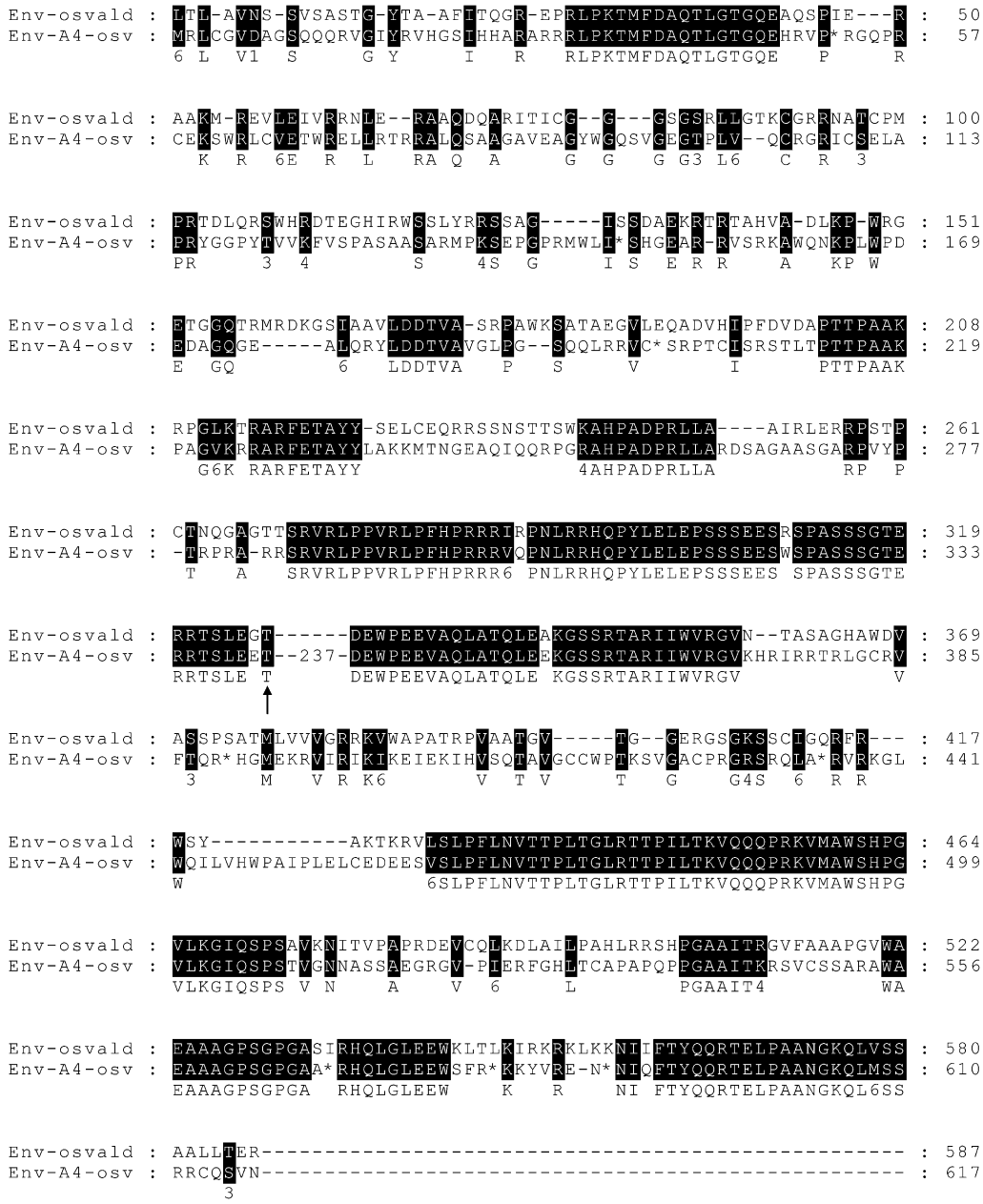

FIGURE 5.—Alignment of an active *Osvaldo env* domain of an active canonical *Osvaldo* with the homologous A4 *env* region. Solid boxes indicate identical or chemically similar amino acid sequences. Asterisks indicate stop codons. After position 341 (arrow) of the A4 *env* domain there are 237 amino acids that cannot be aligned with *Osvaldo*. Numbers below the sequences refer to biochemically similar amino acids as follows: 1, D and N; 3, S and T; 4, K and R; 6, L, I, V, and M.

However, if the unstable hypothesis was correct, insertion events to identical sites should occur independently in many colonizing populations. In our case, identical independent insertion events are unlikely for three reasons. First, new insertions can occur only when active full-length elements are involved in the transposition process. In contrast, we detected insertions, deletions, indels, duplications, and the presence of multiple stop codons in the *env* region of A4 and a deleted *Osvaldo* insertion with seven-nucleotide substitutions in the B9. Second, the age estimation of these insertions, ranging between 0.2 (A4 LTR and B9 RT regions) and 1.5 MY (A4 complete sequence), is much older than the age (300 years) of colonization (FONTDEVILA 1989), reinforcing the conclusion that the occupancy profile is due to genetic drift. It is important to note that even if these insertions are older than the colonization event, they are young in evolutionary time and these age estimates support BOWEN and MCDONALD's (2001) conclusion that Drosophila euchromatic LTR retrotransposons are evolutionarily young. Third, if independent transpositions occur, it is difficult to explain the identical sequences of the 2F4a and 5A4b insertions across populations, unless only one *Osvaldo* copy is active

The fact that the same 2F4a and 5A4b insertions are present in different colonizer populations suggests a common origin and reinforces the founder hypothesis. The absence of these sites in some samples could be due to their loss by sampling drift in some laboratory lines. Moreover, the fact that these two sites were old and were

TABLE 1

Positions of *Osvaldo* (Osv) and its flanking region (Fla) in isofemale lines established from different colonizing populations lines of *D. buzzatii*

Lines:	Adeje-1[a]		Adeje-4[a]		Adeje-7[a]		Adeje-9[a]		Adeje-2[a]	
Probes:	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla
Chrom.										
2	G2, F4a	F1d, F4a, G1c	G2, F5a, F4a	F1d, F4a, G1c	G2a, F4a	F1d, F4a, G1c	E5f, F4a, G2	F1d, F4a	F4a, G2	F1d, F4a
	Puebla del Río-7		Puebla del Río-5		Puebla del Río-4[a]		Puebla del Río-3		Puebla del Río-9	
	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla
Chrom.										
2	—	F1d	—	—	F4a	F4a, F4d	—	F4d	—	F4d
3	—	C5	—	—	F1a	F1d, C1d	—	—	—	—
4	C2g	C2g	C2g	—	—		C2g, F4b	C2g, F4b	C2g	C2g
5	—	—	—	—	A4b	A4b, G1c	—	—	—	—
	Cullera-2		Cullera-4[a]		Cullera-3[a]		Cullera-5[a]		Cullera-7[a]	
	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla
Chrom.										
2	F4a	—	F4a	F4a	F4a	F4a	F4a	F4a	F4a	F4a
3	F1f	E4d	F1f	E4d	F1f	E4d	—	E4d	—	E4d, F1f
4	—	—	—	—	—	—	—	—	—	—
5	A4b	G1d	A4b	G1d	—	G1d, A4b	A4b	G1d	A4b	A4b, G1d
	Colera-1		Colera-7		Colera-5		Colera-4			
	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla		
Chrom.										
2	F1d	F1d	F1d	—	F1d	F1d	F1d	F1d		
3	—	—	—	—	—	—	—	B5		
5	A4b	A4b	A4b	A4b	A4b	A4b	A4b	A4b		
	Sanlucar-3		Sanlucar-1		Sanlucar-2		Sanlucar-7		Sanlucar-10	
	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla	Osv	Fla
Chrom.										
2	E6f	F1d, F4d, F3a	E6f	F4d, F3a	E6f, F1d	F4d, F3a	E6f	F1d, F4d, F3a	E5a	F1d, F4a
3	—	—	—	—	—	—	—	—	F3	—
4	C2g	C2g	C2g	—	C2g	—	C2g	—	C2g	—
5	A4b	A4b	A4b	—	A4b	—	A4b	A4b	A4b	—

Only chromosomes bearing *Osvaldo* insertions are presented. Chrom., chromosomes.

[a] Lines that give a 1.6-kb band (3′ LTR plus flanking region) with specific primers of the A4 clone (2F4a position).

TABLE 2

Characteristics of *Osvaldo* LTRs sequenced in different clones compared to a canonic *Osvaldo* LTR

LTR	Location	bp	Duplications	Insertions	Deletions	Indels ≤3 bases/ substitutions
A4	2F4a	1236	42 bp (455–456)	—	—	4
B4	2F1c	1044	—	—	—	1
P1	Heterochromatin	908	21 bp (69–70)	8 bp (525–526)	322 bp (152)	47

Nucleotide positions are in parentheses.

present in all colonizer populations analyzed in previous works (Labrador *et al.*1998; García Guerreiro and Fontdevila 2001) led us to consider that these two sites are representative of other high-frequency insertions. Different population studies show an accumulation of TEs in regions of low recombination, as inversions (Sniegowsky and Charlesworth 1994) or other nonrecombining regions like the fourth chromosome of *D. melanogaster* (Bartolomé *et al.* 2002; Bartolomé and Maside 2004). However, in our case, previous population analyses (Labrador *et al.* 1998; García Guerreiro and Fontdevila 2001) show no association between high-insertion-frequency sites and regions of low recombination, for example, inversions. Only a case of association was observed and can be explained by a founder effect due to the small effective number of colonizers (Labrador *et al.* 1998), rather than an accumulation mediated by a low-recombination effect.

All evidence reported here favors the founder hypothesis *vs.* the unstable hypothesis but does not negate that colonization may trigger bursts of transposition as suggested by other studies (Vieira and Biémont 1996; Wisotzkey *et al.* 1997; Vieira *et al.* 2002) and that transposition hot spots could exist. The existence of hot spots is difficult to reconcile to the low *Osvaldo* insertion frequency observed in source populations. Even if we suppose an *Osvaldo* transposition rate in original populations lower than that in colonizer ones, we expect a high frequency in putative hot transposition spots. Our study concerns only one clone per band and we cannot discard completely the possibility that new transposition events to new large cytological regions occur after colonization. If this possibility exists the two proposed hypotheses could be correct. This study represents an attempt to combine information for TE sequence structure with chromosomal occupation profiles to disentangle founder effects from transposition bursts in colonization, but it deals only with four clones, one Drosophila species, and one kind of transposable element. Similar analyses, which include additional colonizing species and more TE clones, are under way in our laboratory to test the generality of these results. Yet, these results are the first direct molecular evidence that founder drift effect may be responsible, at least in part, for high-occupancy profiles of TEs in some natural populations.

We acknowledge technical assistance by M. Peiró; some chromosomal slide preparations by B. Chávez; comments and revision of the English text by M. L. Arnold; and helpful discussions and comments by C. Biémont, M. Labrador, X. Maside, and two anonymous reviewers. This research was supported by grants BOS2000-0295-C02 and BOS2003-05904-C02 to A.F. from the Ministerio de Ciencia y Tecnología, Spain.

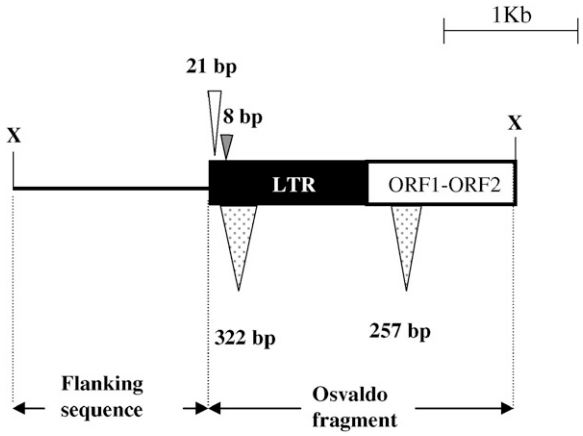

Figure 6.—Structure of the 4.6-kb *Osvaldo* fragment included in the P1 clone (GenBank accession no. DQ139462). Open, shaded, and stippled triangles depict, respectively, duplications, insertions, and deletions. X indicates *Xba*I.

LITERATURE CITED

Barker, J. S. F., 1982 Population genetics of *Opuntia* breeding *Drosophila*, pp. 209–224 in *Ecological Genetics and Evolution*, edited by J. S. F. Barker and W. T. Starmer. Academic Press, New York.

Bartolomé, C., and X. Maside, 2004 The lack of recombination drives the fixation of transposable elements on the fourth chromosome of *Drosophila melanogaster*. Genet. Res. **83:** 91–100.

Bartolomé, C., X. Maside and B. Charlesworth, 2002 On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. Mol. Biol. Evol. **19:** 926–937.

Berg, D. E., and M. M. Howe, 1989 *Mobile DNA*. American Society for Microbiology, Washington, DC.

Biémont, C., 1992 Population genetics of transposable DNA elements. A *Drosophila* point of view. Genetica **86:** 67–84.

Biémont, C., C. Nardon, G. Deceliere, D. Lepetit, C. Loevenbruck *et al.*, 2003 World wide distribution of transposable element copy number in natural populations of *Drosophila*. Evolution **57:** 159–167.

Bowen, N. J., and J. F. Mcdonald, 2001 *Drosophila* euchromatic LTR retrotransposons are much younger than the host species in which they reside. Genome Res. **11:** 1527–1540.

Charlesworth, B., and C. H. Langley, 1989 The population genetics of *Drosophila* transposable elements. Annu. Rev. Genet. **23:** 251–287.

Conte, C., V. Calco, S. Desset, P. Leblanc, B. Dastugue *et al.*, 2000 Impact of multiple insertions of two retroelements, *Zam* and *Idefix* at an euchromatic locus. Genetica **109:** 53–59.

David, J., and L. Tsacas, 1980 Cosmopolitan, subcosmopolitan and widespread species: different strategies within the *Drosophilid* family (*Diptera*). Comptes Rendus de la Société Biogéographique **57:** 11–26.

Dimitri, P., 1997 Constitutive heterochromatin and transposable elements in *Drosophila melanogaster*. Genetica **100:** 85–93.

Evgen'ev, M. B., G. N. Yenikolopov, N. I. Peunova and Y. V. Ilyin, 1982 Transposition of mobile genetic elements in interspecific hybrids of *Drosophila*. Chromosoma **85:** 375–386.

Evgen'ev, M. B., H. Zelentsova, H. Poluectova, G. T. Lyozin, V. Veleikodvorskaja *et al.*, 2000 Mobile elements and chromosomal evolution in the *virilis* group of *Drosophila*. Proc. Natl. Acad. Sci. USA **10:** 11337–11342.

Fontdevila, A., 1989 Founder effects in colonizing populations: the case of *Drosophila buzzatii*, pp. 74–95 in *Evolutionary Biology of Transient Unstable Populations*, edited by A. Fontdevila. Springer, Berlin.

Fontdevila, A., 1991 Colonizing species of *Drosophila*, pp. 249–269 in *Molecular Techniques in Taxonomy*, edited by G. M. Hewitt, A. W. B. Johnston and J. P. W. Young. Springer-Verlag, Berlin/Heidelberg, Germany.

Fontdevila, A., A. Ruíz, G. Alonso and J. Ocaña, 1981 Evolutionary history of *Drosophila buzzatii* I. Natural chromosomal polymorphism in colonized populations of the Old World. Evolution **35:** 148–156.

García Guerreiro, M. P., and A. Fontdevila, 2001 Chromosomal distribution of the transposable elements *Osvaldo* and *blanco* in original and colonizer populations of *Drosophila buzzatii*. Genet. Res. **77:** 227–238.

Higgins, D. G., J. D. Thompson and T. J. Gibson, 1996 Using CLUSTAL for multiple sequence alignments. Methods Enzymol. **266:** 383–402.

Kaminker, J. S., C. M. Bergman, B. Kronmiller, J. Carlson, R. Svirskas *et al.*, 2002 The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. Genome Biol. **3**(12): RESEARCH 0085.

Kapitonov, V., and J. Jurka, 1996 The age of Alus subfamilies. J. Mol. Evol. **42:** 59–65.

Kumar, S., K. Tamura, I. B. Jakobsen and M. Nei, 2001 MEGA-2: Molecular Evolutionary Genetics Analysis software. Bioinformatics **17:** 1244–1245.

Labrador, M., and A. Fontdevila, 1994 High transposition rates of *Osvaldo*, a new *D. buzzatii* retrotransposon. Mol. Gen. Genet. **245:** 661–674.

Labrador, M., H. Naveira and A. Fontdevila, 1990 Genetic mapping of the *Adh* locus in the *repleta* group of *Drosophila* by *in situ* hybridization. J. Hered. **81:** 83–86.

Labrador, M., M. C. Seleme and A. Fontdevila, 1998 The evolutionary history of *D. buzzatii*. XXXIV. The distribution of the retrotransposon *Osvaldo* in original and colonizer populations. Mol. Biol. Evol. **15:** 1532–1547.

Lerat, E., C. Rizzon and C. Biémont, 2003 Sequence divergence of transposable elements in the *D. melanogaster* genome. Genome Res. **13:** 1889–1896.

Li, W., 1997 *Molecular Evolution*. Sinauer, Sunderland, MA.

Maniatis, T. E., E. F. Fritsch and J. Sambrook, 1982 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Marsano, R. M., S. Marconi, R. Moschetti, P. Barsanti, C. Caggese *et al.*, 2004 *Max*, a novel retrotransposon of the BEL-Pao family, is nested within the *Bari 1* cluster at the heterochromatic *h39* region of chromosome 2 in *Drosophila melanogaster*. Mol. Genet. Genomics **270:** 477–484.

Nicholas, K. B., H. B. J. Nicholas and D. W. Deerfield, 1997 Genedoc: analysis and visualization of genetic variation. Embnet News **4:** 14.

Pantazidis, A., M. Labrador and A. Fontdevila, 1999 The retrotransposon *Osvaldo* from *D. buzzatii* displays all structural features of functional retrovirus. Mol. Biol. Evol. **16:** 909–921.

Petrov, D. A., 2002 DNA loss and evolution of genome size in *Drosophila*. Genetica **115:** 81–91.

Pimpinelli, S., M. Berloco, L. Fanti, P. Dimitri, S. Bonaccorsi *et al.*, 1995 Transposable elements are stable structural components of *Drosophila melanogaster* heterochromatin. Proc. Natl. Acad. Sci. USA **92:** 3804–3808.

Piñol, J., O. Francino, A. Fontdevila and O. Cabre, 1988 Rapid isolation of *Drosophila* high molecular weight DNA to obtain genomic libraries. Nucleic Acids Res. **16:** 2736–2737.

Preston, B. D., and J. P. Dougherty, 1996 Mechanisms of retroviral mutation. Trends Microbiol. **4:** 16–21.

Sanger, F., S. Nicklen and A. R. Coulson, 1977 DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

Schmidt, E. R., 1992 A simplified and efficient protocol for non-radioactive in situ hybridisation to polytene chromosomes with a DIG.-labeled DNA probe, pp. 36–38 in *Non Radioactive in Situ Hybridisation Application Manual*. Roche, Indianapolis.

Shevelyov, Y. Y., M. D. Balakireva and V. A. Gvozdev, 1989 Heterochromatic regions in different *Drosophila melanogaster* stocks contain similar arrangements of moderate repeats with inserted *copia-like* elements (*mdg1*). Chromosoma **98:** 117–122.

Sniegowsky, P. D., and B. Charlesworth, 1994 Transposable element numbers in cosmopolitan inversions from a natural population of *Drosophila melanogaster*. Genetics **137:** 815–827.

Vaury, C., A. Bucheton and A. Pélisson, 1989 The heterochromatin sequences flanking the *I* elements are themselves defective transposable elements. Chromosoma **98:** 215–224.

Vieira, C., and C. Biémont, 1996 Geographical variation in insertion site number of retrotransposon *412* in *Drosophila simulans*. J. Mol. Evol. **14:** 185–188.

Vieira, C., D. Lepetit, S. Dumont and C. Biémont, 1999 Wake up of transposable elements following *Drosophila simulans* worldwide colonization. Mol. Biol. Evol. **16:** 1251–1255.

Vieira, C., C. Nardon, C. Arpin, D. Lepetit and C. Biémont, 2002 Evolution of genome size in *Drosophila*. Is the invader's genome being invaded by transposable elements? Mol. Biol. Evol. **19:** 1154–1161.

Wisotzkey, R., I. Felder and A. Hunt, 1997 Biogeographic analysis of the *Uhu* and *LOA* elements in the *Hawaiian Drosophila*. Chromosoma **106:** 465–477.

Communicating editor: T. H. Eickbush