Molecular Characterization of the Breakpoints of an Inversion Fixed Between \textit{Drosophila melanogaster} and \textit{D. subobscura}

Susanna Cirera, Jesús M. Martín-Campos, Carmen Segarra and Montserrat Aguadé

Departament de Genética, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

Manuscript received July 10, 1994
Accepted for publication September 26, 1994

ABSTRACT

The two breakpoints of a chromosomal inversion fixed since the split of \textit{Drosophila melanogaster} and \textit{D. subobscura} lineages have been isolated and sequenced in both species. The regions spanning the breakpoints initially were identified by the presence of two signals after interspecific \textit{in situ} hybridization on polytene chromosomes. Interspecific comparison of the sequenced regions allowed us to delineate the location of the breakpoints. Close to one of these breakpoints a new transcription unit (\textit{ben92}) has been identified in both species. The inversion fixed between \textit{D. melanogaster} and \textit{D. subobscura} does not seem to have broken any transcription unit. Neither complete nor defective transposable elements were found in the regions encompassing the breakpoints. Short thymine-rich sequences (30–50 bp long) have been found bordering the breakpoint regions. Although alternating Pur-Pyr sequences were detected, these putative target sites for topoisomerase II were not differentially clustered in the breakpoints.

Most \textit{Drosophila} species differ in their karyotype. Homossexual species complexes with identical banding patterns in their polytene chromosomes (CARSON \textit{et al.} 1967) are an exception. However, these species may differ in heterochromatic regions. Therefore, most speciation processes seem to be associated with the establishment of chromosomal rearrangements (WHITE 1978); this does not imply that rearrangements have played an essential role in the speciation process.

In the genus \textit{Drosophila}, those chromosomal rearrangements that mainly contribute to karyotype differentiation are paracentric inversions. Other chromosomal changes such as pericentric inversions or translocations are much less common (CLAYTON and GUEST 1986).

\textit{Drosophila melanogaster} and \textit{D. subobscura} are species included in two different groups of \textit{Drosophila} that presumably diverged 30 myr ago (THROCKMORTON 1975). \textit{In situ} hybridization studies indicate that extensive rearrangements have occurred during the divergence of these species changing dramatically the location of different X-linked markers (SEGARRA and AGUADÉ 1992).

The origin of a paracentric inversion implies two breaks and the reversal of the generated chromosomal segment. It seems plausible that along the genome there may be regions with a tendency to suffer spontaneous breaks such as regions initiating recombination. Sites of recombination seem to be correlated with topoisomerase II cleavage sites (SPERRY \textit{et al.} 1989). Therefore, alternating Pur-Pyr sequences (RY repeats) that show considerable similarity to the consensus target sequence for topoisomerase II (SPITZNER \textit{et al.} 1990) may be involved in illegitimate recombination and might cause chromosomal breaks. Moreover, it has been suggested that transposable elements may act as chromosomal breakage agents in eukaryotes (ENGELS and PRESTON 1984).

Independent of its origin, the fate of a newly arisen inversion may depend on factors such as selection, genetic drift and/or meiotic drive. The same conditions that promote speciation might facilitate the fixation of inversions (for instance, the small population size required by the founder event model of speciation). After speciation, paracentric inversions can be fixed and therefore accumulate in the different lineages during species divergence resulting in a progressive differentiation of karyotypes.

Here we describe the identification of one of the breakpoints of an inversion fixed during the divergence of \textit{D. melanogaster} and \textit{D. subobscura} in the recombinant clone Acp12A (SCOTT and LUCCHESI 1991). This identification was fortuitous but allowed us to proceed to its molecular characterization and to isolate both inversion breakpoints in both species. Comparison of sequences of these inversion breakpoints has allowed us to delineate accurately the location of the breakpoints, to characterize features of their sequences and to check for the presence of transposable elements.

MATERIALS AND METHODS

Cloning and sequencing: Random genomic libraries of both \textit{D. subobscura} and \textit{D. melanogaster} were used. Phage
Acpl2A encompassing the Pgd region of D. melanogaster was kindly provided by M. Scott and J. Lucchessi. Conditions for library screening and Southern blotting are as previously described (Aguadé 1988). All fragments including the breakpoints were subcloned in pBlueScript SK+ or KS+. Sets of nested deletions were obtained according to Henikoff (1984). Only in one case (see below) were specially designed oligonucleotides for both strands of the dideoxy method (Sanger et al. 1977) using double-stranded DNA and T7 DNA polymerase. Sequences were assembled using Staden’s programs (Staden 1982). The Wisconsin GCG package (Devereux et al. 1984) was used for sequence alignment and comparison.

In situ hybridization: Probes were labelled with biotin-16-UTP by nick translation. Prehybridization, hybridization and detection were performed as described by Segarra and Aguadé (1992).

Expression assay: Total RNA was prepared from adult flies of D. melanogaster and D. subobscura and used to obtain total cDNA by reverse transcription using either random hexanucleotides or oligo(dT) (Ridley 1993). Two oligonucleotides located in the first and second exons of the putative transactivation unit were used to amplify this region by the polymerase chain reaction (PCR) from the total cDNA. PCR products were separated in a 1.5% (3 NuSieve:1 SeaKem) agarose gel.

RESULTS

Identification and isolation of the breakpoints in D. subobscura and D. melanogaster: In the process of cloning the Pgd region of D. subobscura, two different fragments of the Pgd region of D. melanogaster (Scott and Lucchessi 1991) were used as probes (Figure 1). Three positive phages were isolated in D. subobscura (Figure 1): phages Ps1 and Ps2, obtained when using the telomere-proximal region of the Pgd gene as a probe (4.7RR), that gave a unique and strong signal at band 13A of the A (= X) chromosome of D. subobscura and phage λF6, obtained when using the telomere-distal region of the Pgd gene as a probe (3.5RR), that gave a unique and strong signal at band 16C close to the telomere of the A chromosome of D. subobscura (Figure 2).

Phage λF6 of D. subobscura gave, however, two distant hybridization signals on polytene chromosomes of D. melanogaster (Figure 2): one at band X-2D (where the Pgd region is located) and a second one at band X-5C. On the other hand, when phage Acpl2A or the 3.5-kb EcoRI-EcoRI fragment of D. melanogaster was used as a probe for in situ hybridization (Figure 2), only one signal was observed on chromosomes of D. melanogaster (band X-2D) and two signals on chromosomes of D. subobscura (bands A-13A and A-16C). As shown schematically in Figure 3, these data clearly point to the existence of a chromosomal rearrangement between D. melanogaster and D. subobscura. Phages Acpl2A (Scott and Lucchessi 1991) and λF6 encompass one of the breakpoints in D. melanogaster and D. subobscura, respectively. Phages Ps1 and Ps2 encompass the second breakpoint in D. subobscura (see below), despite its failure to give two signals when hybridized to polytene chromosomes of D. melanogaster. A random genomic library of D. melanogaster was screened with a 2.1-kb HindIII-SalI subclone of phage Ps2 of D. subobscura spanning the breakpoint in this species. One positive phage (ft1) that includes the second breakpoint in D. melanogaster was isolated.

Characterization and subcloning of the breakpoints: After Southern blot analysis of phage λF6 of D. subobscura using the 3.5-kb EcoRI-EcoRI fragment of D. melanogaster as a probe, two overlapping fragments that spanned the breakpoint were subcloned and partially sequenced: a 1.7-kb EcoRI-XbaI fragment and a 2.3-kb XhoI-XbaI fragment (Figure 1). In the case of phage Ps1, which contains the Pgd region (J. M. Martín-Campos, unpublished data), most of the phage was subcloned and sequenced, although the sequence reported here only includes a 1.8-kb terminal fragment of this phage.

In D. melanogaster the original 3.5-kb EcoRI-EcoRI fragment spanned the breakpoint but also included part of the coding region of Pgd (Scott and Lucchessi 1991). Initially a 1.1-kb SalI-EcoRI fragment distal to the coding region was subcloned and sequenced; later the region extending from the published sequence to the SalI site was sequenced directly from the 3.5-kb subclone by using specially designed oligonucleotides for both strands. After Southern blot analysis of phage ft1 using the 2.1-kb HindIII-SalI fragment from the Ps2 phage of D. subobscura as a probe, a 2.2-kb SalI-SalI fragment that spanned the breakpoint in D. melanogaster was sub-
Figure 2.—In situ hybridization of probes Acp12A and λF6 to polytene chromosomes of *D. melanogaster* and *D. subobscura*. (A) Acp12A vs. *D. melanogaster*. (B) Acp12A vs. *D. subobscura*. (C) λF6 vs. *D. subobscura*. (D) λF6 vs. *D. melanogaster*.

cloned and sequenced (Figure 1). The four sequences are deposited in the EMBL data library under accession numbers Z46522 and Z46608 to Z46610.

**Sequence comparison, search for homologies and delimitation of breakpoints:** Sequence comparison between species of the two regions encompassing the inversion breakpoints has shown regions of rather high levels of similarity in three of the four comparisons, as revealed by dot-matrix analysis (Figure 4), allowing identification of homologous regions (Figure 5). Although no consistent pattern of similarity could be detected when comparing the sequenced regions of phages f1 from *D. melanogaster* and λF6 from *D. subobscura*, Southern blot analysis of the f1 positive phage, using the *EcoRI*-HindIII fragment of λF6 (distal to the sequenced region) as probe, revealed homology only in the 2.2-kb *SacI*-SacI fragment. As shown in Figure 5, the location of the breakpoint can be delineated rather accurately both in phage Acp12A from *D. melanogaster* (in a region at most 400 bp long) and in phage Ps1 of *D. subobscura* (also in a region at most 600 bp long) as they are flanked by two rather conserved homologous regions. However, only one of the breakpoint limits can be unambiguously established both in phages λF6 of *D. subobscura* and f1 of *D. melanogaster* (see Figure 5).

Some short thymine-rich regions (30–50 bp long) were present in all the breakpoint regions (Figure 5). In fact, in each of the breakpoint regions, there is such a thymine-rich stretch in the limit of a region of homology between *D. melanogaster* and *D. subobscura*. Additionally, several alternating Pur-Pyr sequences (at least 10 repeats long) could be identified in all the breakpoint regions.

**Detection of a new transcription unit:** An open reading frame (ORF) 276 bp long was detected both in *D. melanogaster* and *D. subobscura*. The expression assay performed both in adult males and females revealed that this ORF is a transcription unit. In fact, when total

Figure 3.—Schematic representation of the expected result of *in situ* hybridization in case that Acp12A and λF6 contained a breakpoint. Striped bars indicate the probe used. Arrows indicate the expected hybridization signal.
cDNA was PCR amplified using primers located in exons 1 and 2 of the putative coding region a single band of 180 bp was detected (Figure 6). Sequencing of this PCR cDNA product confirmed the location of one of the intron/exon boundaries of this gene. As shown in Figure 5 and based on sequence comparison, this new gene (bcn92) would consist of three exons coding for a putative protein 92 amino acids long. Computer search in the DNA (EMBL), protein (Swiss-Prot, release 27.0) and protein motifs (Prosite, release 11.0) databases did not reveal any homology to bcn92. Estimates of silent and nonsilent nucleotide divergence at the bcn92 region between D. melanogaster and D. subobscura were $K_s = 1.08$ and $K_a = 0.08$ (Jukes and Cantor 1956; Nei and Gojobori 1986). These values are intermediate among estimates for other genes sequenced in the melanogaster and obscura groups (Moriyama and Gojobori 1992; Segarra and Aguade 1993).

**DISCUSSION**

The melanogaster and the obscura lineages have been separated for more than 30 myr (Throckmorton 1975). When sequences of a particular genomic region are compared between these two groups, alignment is generally only possible for coding regions; it is very difficult or impossible for some introns or flanking regions. The regions encompassing the inversion breakpoints were therefore not expected to be highly conserved unless subject to some functional constraint. The levels of similarity found in the regions shown in Figures 4 and 5 cannot be attributed to randomness. In fact, a new transcription unit has been identified and a second region corresponds to RNA D in Gandhi et al. (1992). Additionally, there is a region with an average level of similarity of 67.4% in phages Ps1 and ft1 (Figure 5). Although we have not detected any transcription unit in this region of similarity, it could contain regulatory signals for an adjacent unidentified gene.

The homologies detected have allowed us to delimit one of the breakpoints in D. melanogaster to an ~400-bp region (phage Acp12A, Figure 5) and in D. subobscura to an ~600-bp region (phage Ps1, Figure 5). The original similarity in the immediate flanks of the breakpoints has not been retained, indicating the absence of common functional constraints in both species. The transcription units near the breakpoints that are conserved in both species have not been disrupted by the inversion. We cannot, however, assert that the inversion has not altered or separated regulatory regions from the identified coding regions, as we have no knowledge of possi-

**FIGURE 4.**—Dot plots of the three sequence comparisons with regions of similarity. A window of 21 bases, with a stringency of 14 bases has been used.
Drosophila Inversion Breakpoints

**FIGURE 5.**—Comparison between the sequenced regions encompassing the inversion breakpoints. Big shaded boxes include regions of homology between *D. melanogaster* and *D. subobscura* and delimit the breakpoints (white bars between shaded boxes). The telomere-proximal clone of *D. melanogaster* (Acp12A) presents two regions with a high level of similarity with *D. subobscura*, one with the distal clone Ps1 and the other with the proximal clone AF6. Both regions encompass transcription units. The distal clone of *D. subobscura* (Ps1) also presents a region of similarity with the distal clone of *D. melanogaster* (ft1). Within homologous regions, higher levels of similarity are indicated by striped boxes. The exons of *bcn92* and the primary transcript RNA D are shown as bars below the map and the arrowheads indicate the direction of transcription; a discontinuous bar indicates that in *D. subobscura* the RNA D transcript has been established by sequence comparison. *, thymine-rich regions. C, Clal; H, HindIII; R, EcoRI; S, SaI; Sc, SacI; Ss, SspI; Xb, Xbal; X, XhoI. D. mel., *D. melanogaster*; D. sub., *D. subobscura*.

![Diagram of inversion breakpoints](image)

**FIGURE 6.**—Analytical electrophoresis of PCR products obtained when using primers located within the coding region of *bcn92*.

![Electrophoresis image](image)
rich stretches bordering the breakpoints is noteworthy. On the other hand, although several putative topoisomerase II binding sites (RY repeats) could be detected in the regions encompassing the breakpoints of the inversion under study, they were not differentially clustered in the breakpoints.

Recently Wesley and Eanes (1994) have reported the sequences of the two breakpoints of inversion In(3L)Payne of D. melanogaster. This, however, is a polymorphic inversion and not an inversion fixed during species divergence. Contrary to our result that the inversion fixed between D. melanogaster and D. subobscura has most probably not broken any transcription unit, one of the In(3L)Payne breakpoints disrupts three transcripts. However, no transposable elements have been detected in any of the breakpoints of these two inversions; this stands in contrast with the presence of one of these elements in the breakpoint of the Bar duplication in D. melanogaster (Tsubota et al. 1989). On the other hand, we have found a thymine-rich stretch bordering one of the breakpoints of In(3L)Payne (sequence D, Wesley and Eanes 1994).

In D. melanogaster, the order of markers with respect to the telomere is telomere-Pgd-bcn92-RNA D. In contrast in D. subobscura, RNA D is close to the telomere and Pgd-bcn92 map on a more internal position (although their relative orientation to the telomere is unknown). Therefore, in addition to the characterized inversion, at least another inversion is necessary to explain the present distribution of markers in both species. This is not surprising because it has been shown previously by \textit{in situ} hybridization (Segarra and Aguadé 1992) that the X chromosome has suffered extensive reorganization through chromosomal rearrangements since the split of the melanogaster and obscura lineages. However, only an extensive study by \textit{in situ} hybridization using overlapping clones of \textit{D. melanogaster} will allow a good inference of the number of inversions fixed between the two lineages.

We thank M. J. Scott and J. C. Lucchesi for the Acp12A clone and J. Rozas for critical comments on the manuscript. This work was supported by a predoctoral fellowship from Comisión Interdepartamental de Recerca i Tecnologia to S.C. and by grant PB91-245 from Dirección General de Investigación Científica y Técnica to M.A.

\section*{Literature Cited}


Segarra, C., and M. Aguadé, 1992 Molecular organization of the \textit{X} chromosome in different species of the \textit{obscura} group of \textit{Drosophila}. Genetics 130: 513–521.


Communicating editor: A. G. Clark