The phylogeny of nine species of the *Drosophila obscura* group inferred by the banding homologies of the chromosomal regions. IV. Element C

ANTONIO BREHM & COSTAS B. KRIMBAS

Department of Genetics, Agricultural University of Athens, 75 Iera Odos, 11855 Athens, Greece

Homologies of the sections of the polytene chromosomes of element C, among several species of the *obscura* group of the genus *Drosophila*, were established according to the similarity of their banding patterns. The information gathered was used to construct an unrooted phylogenetic tree based on qualitative criteria. This tree is compared to three other similar trees derived from independent information provided by the study of chromosomal elements B, D and E. The general congruence of the patterns of these trees proves the well-foundedness of this approach. A single exception to this congruence is discussed. Finally a consensus tree, encapsulating information from all chromosomal elements, is presented and its topology is compared to those derived from electrophoretic data.

Keywords: chromosome phylogeny, *Drosophila*, element C, *obscura* group, phylogeny *Drosophila*, phylogeny *obscura*.

Introduction

The method of phylogenetically seriating chromosome gene arrangements using overlapping inversions and following the most parsimonious path was long ago proposed by Sturtevant & Dobzhansky (1936) and has been used by several authors since (for a review see Krimbas & Powell, 1992). This method may also be used to seriate closely related species when the identification of the homologous regions in their respective giant chromosomes is established. This move from the phylogeny of intraspecies gene arrangements presupposes the acceptance of a further condition, that of the absence of the gene arrangements' polymorphism transcending the species splitting (Brehm *et al.*, 1991; Krimbas & Powell, 1992).

In addition to the seriation of a triad of gene arrangement, based on the presence of overlapping inversions, there is the possibility of seriating them, following the same principle of parsimony, by a method based on the conservation or disassociation of nearby located chromosomal segments. Thus the presence, in gene arrangement I, of the contiguously located segments A B C in that order, and also the presence of a gene arrangement II showing A and B together but C separated from them, and finally the

presence of gene arrangement III displaying B and C together and A apart leads to the following most parsimonious seriation

II [AB...C] ——— I [ABC] ——— III [A...BC].

Already Wülker *et al.* (1984) has explained this principle, while its origins are to be found in the work of Stalker (1972), who used it to detect phylogenetic relations between distantly related *Drosophila* species.

We have been engaged in the study of nine species belonging to the *obscura* group and subgroup of species of the genus *Drosophila* using these principles. These species have a karyotype composed of five chromosomal elements (acrocentric chromosomes or arms of metacentric ones) and one dot chromosome. There is no indication of transfer of chromosomal material from one to another element by translocation. Thus every element can be studied independently and from each study information may be gathered permitting the construction of an unrooted phylogeny. These phylogenies can then be compared. A congruence among them would corroborate the assumption that phylogenetic relations derived this way are generally valid. This is quite important because only DNA sequence data may provide such a rich amount of information as banding sequences of giant chromo-

somes. Finally phylogenies based on DNA sequence data may be compared to those derived from banding patterns.

This programme of work is nearly completed. From the five chromosomal elements we were able to study only four, namely element B (Brehm *et al.*, 1991). C (the present study), D (A. Brehm & C. B. Krimbas, 1992, in press), and E (Brehm & Krimbas, 1990), the first one, the sex chromosome A, being extremely reorganized and therefore impossible to study by this method.

Thus the present paper deals first with element C, undoubtedly the most difficult of the autosomal elements to be homologized within this species group. It compares the four phylogenies derived independently from the study of the four autosomal elements and discusses their congruence. Finally it compares phylogenetic information derived from a different source, electrophoretic data, with that gathered from the study of the banding pattern of the giant chromosomes.

Materials and methods

The following nine species were examined: *Drosophila subobscura, D. madeirensis, D. guanche, D. kitumensis, D. microlabis, D. subsilvestris, D. obscura, D. ambigua* and *D. tristis*. All species are Palearctic in their geographical distribution except for *D. kitumensis* and *D. microlabis*, which are Afrotropical and were collected in Kenya. *D. madeirensis* and *D. guanche* are endemic to the Madeira and Canary Islands respectively. For a detailed description of the strains used see Brehm *et al.* (1991).

Microscopic preparations of polytene chromosomes were made from early pupae as described by Brehm *et al.* (1991) and Brehm & Krimbas (1990, 1992). Photographs from these slides were used for comparisons. Photomaps of element C from each species were constructed and compared in the search for segments with an identical banding pattern. As it seems that *D. obscura* occupies a central position relative to the remaining species studied, chromosomes from all other

species were compared to its standard sequence. Indeed *D. obscura* displays an important and complicated inversion polymorphism which was the object of a detailed study (Brehm & Krimbas, 1991) and thus a photomap of its standard sequence is already available.

Apart from *D. obscura* only two other species were found to be polymorphic for element C, *D. subobscura* and *D. tristis*. *D. subobscura*'s polymorphism is well studied and recently reviewed by Krimbas (1992). In order to proceed to comparisons we have used the Standard gene arrangement as depicted in the drawings of Kunze-Mühl & Müller (1958) amended by the new photomaps published by Brehm (1992). *D. madeirensis* bears the Standard gene arrangement of *D. subobscura* and *D. guanche* differs from it by two independent inversions involving the sections 59D-66C/D and 67C-72B/C (Krimbas & Loukas, 1984; Molto *et al.*, 1987). According to the new interpretation of Brehm (1992) the breakage points are situated at the locations 59D-66 and 68-71B.

Results

From the study of the similarity of their respective gene arrangements it is possible to distinguish clearly three clusters of related species, the *obscura*, the *subobscura* and the African cluster. Figure 1 is a photomap of element C of *D. tristis*. On the standard gene arrangement of *D. tristis*, shown in this figure, we have indicated by arrows the break points of several inversions, which transform this arrangement to that of the other two species also belonging to the *obscura* cluster, namely *D. obscura* and *D. ambigua*. Thus inversion t transforms this arrangement to C_{2+6} encountered in *D. obscura*. Two consecutive overlapping inversions, 6 and 2, are needed to obtain the Standard arrangement of *D. obscura*. On the other hand the gene arrangement of *D. tristis* may be transformed to that of *D. ambigua* by four additional inversions, some of them overlapping, a_1, a_2, a_3 and a_4. Thus the gene arrangement of *D. ambigua* with respect to that of *D. tristis* is $C_{a1+a2+a3+a4}$. The system of notations in this figure is
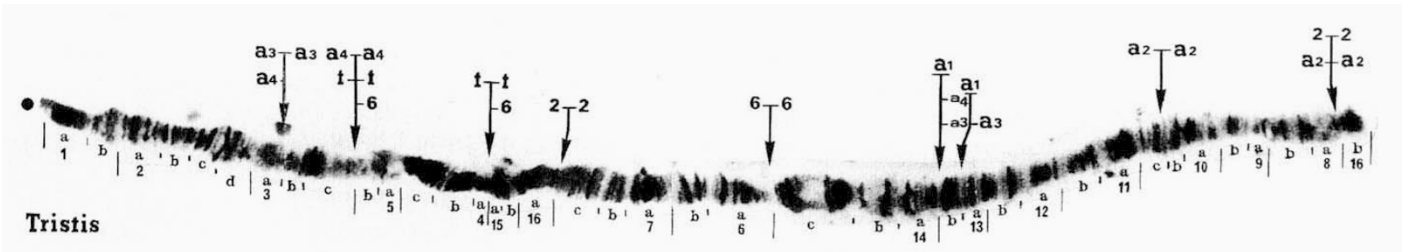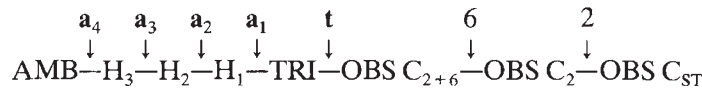

Fig. 1 Element C of *Drosophila tristis*. The breakpoints of inversions t (transforming the *tristis* gene arrangement to that of *obscura* C_{2+6}), a_1, a_2, a_3 and a_4 (transforming it to the gene arrangement of *ambigua*) are indicated by arrows. Further explanations in the text.

the same as the one used by Kunze-Mühl & Müller (1958) for *D. subobscura* and by Brehm & Krimbas (1991) for *D. obscura*.

Thus it is possible to arrange in a linear order the gene arrangement of the three species as follows:

$$\overset{\overset{a_4}{\downarrow}\,\overset{a_3}{\downarrow}\,\overset{a_2}{\downarrow}\,\overset{a_1}{\downarrow}\quad\overset{t}{\downarrow}\qquad\qquad\overset{6}{\downarrow}\qquad\overset{2}{\downarrow}}{AMB—H_3—H_2—H_1—TRI—OBS\ C_{2+6}—OBS\ C_2—OBS\ C_{ST}}$$

where H_1, H_2 and H_3 represent hypothetical intermediate gene arrangements.

A characteristic segment of the gene arrangement C_{2+6} of *D. obscura* is that comprising the sections 7a/6ba/14c; it is indicated in Figures 2, 3 and 4 by a grey bar. As mentioned below, this small characteristic segment is found intact (in one piece) and was identified in all the species studied with the exception of *D. subobscura* and *D. madeirensis*, where it is separated into two parts (Fig. 2a). The presence of this segment permits the identification of the gene arrangement of *D. obscura* which is closer to that of the other species of the *obscura* group (except for *D. ambigua* and *D. tristis*, as already discussed). It must be the C_{2+6} gene arrangement of this species, since only in this arrangement does 7a/6ba/14c remain intact.

D. subsilvestris' C chromosome could not be completely homologized with that of *D. obscura* but much of it has identifiable counterparts, as shown schematically in Fig. 2b. There is no indication whether the *D. subsilvestris* gene arrangement is more closely related to *D. ambigua* or to *D. tristis* than it is to the C_{2+6} gene arrangement of *D. obscura*.

D. kitumensis differs from *D. microlabis* by two overlapping inversions (their breakpoints on the chromosome of *D. kitumensis* are indicated by the numbers 1 for the first inversion and 2 for the second, see Fig. 3). Segment 7a/6ba/14c is the only one homologized in these two species and in *D. obscura*. It is obvious that both African species differ from *D. obscura* by a number of rearrangements in such a way that common banding patterns are no longer discernable, except for 7a/6ba/14c.

The *subobscura* species cluster also shows an important reorganization of gene sequences in comparison to those of *D. obscura* and of all other species in general. However, the link between this cluster and the other species is easily defined. *D. guanche* is the species occupying the intermediate position between on the one hand *D. madeirensis–D. subobscura* and the remaining species on the other. Figure 2a shows that one of the inversions by which *D. guanche* differs from the two remaining species of its cluster, *D. madeirensis* and *D. subobscura*, splits segment 7a/6ba/14c (shown

as a grey bar in Figs 2–4) into two pieces. As this segment is encountered in one piece in all other species the obvious seriation is:
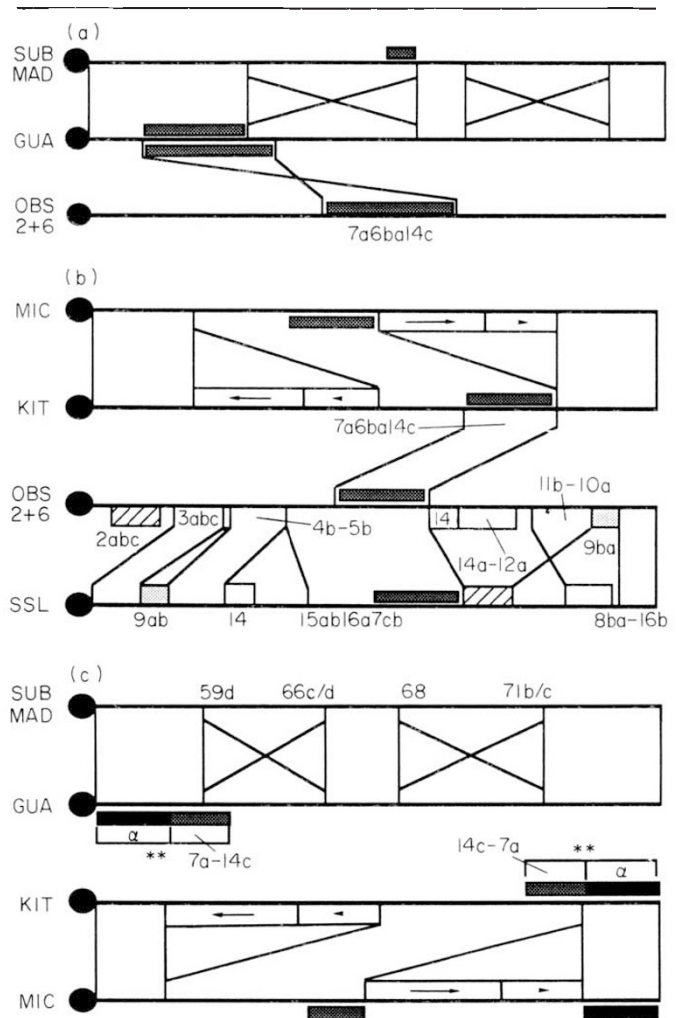
MAD, SUB — GUA — other species.

Fig. 2 Homologous chromosomal segments of element C among some of the species studied are schematically presented. (———) Chromosome, (●) the position of the centromere. Homologous segments have their ends united by lines; when the lines cross they are in inverted order. A grey box indicates the diagnostic segment 7a/6ba/14c (in some cases abbreviated 7a–14c), which in *subobscura* and *madeirensis* is separated into two segments. (*) Compound segment a + 7a − 14c. Its first part (a) is indicated by a black box. Remote homologous segments are symbolized by similar boxes. SUB = *D. subobscura*, MAD = *D. madeirensis*, GUA = *D. guanche*, OBS = *D. obscura*, MIC = *D. microlabis*, KIT = *D. kitumensis*, SSL = *D. subsilvestris*.
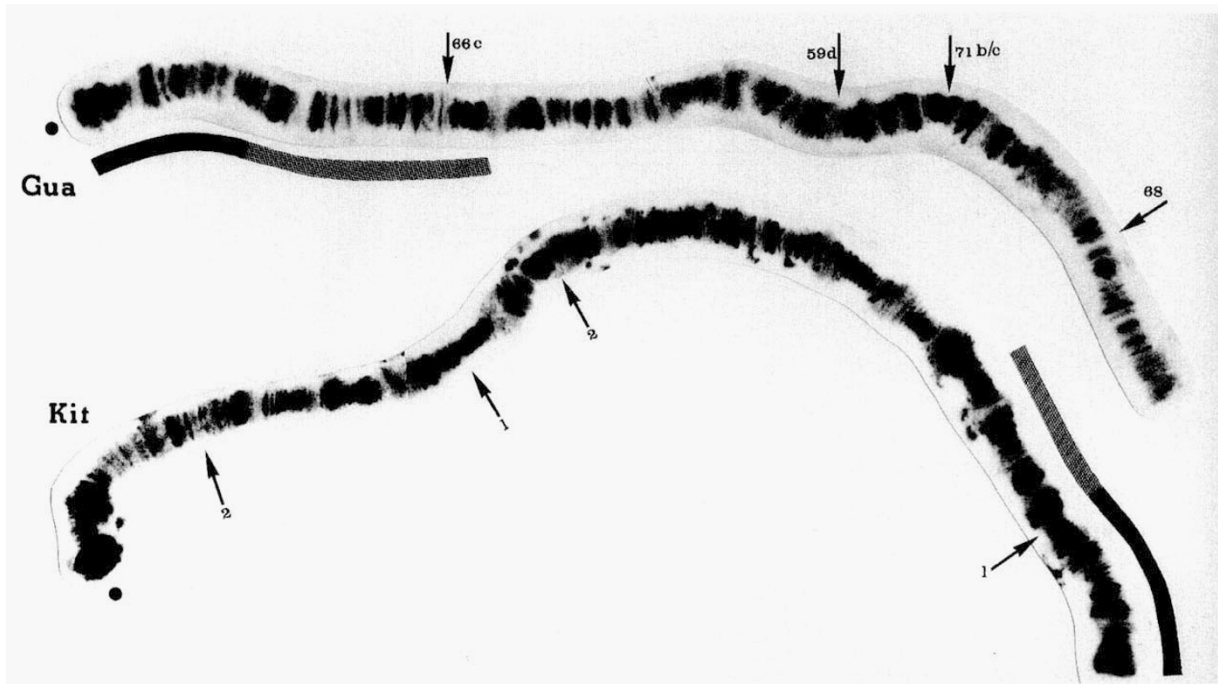
Fig. 3 The position of the characteristic segments in the chromosomes of *D. guanche* (Gua) and *D. kitumensis* (Kit). Segment 7a/6ba/14c is symbolized by a grey bar while a is symbolized by a black bar. (●) The position of the centromere. The positions of the inversion's breakpoints that transform it into *D. madeirensis* and *D. subobscura* are indicated on the chromosome of *D. guanche*, while on the chromosome of *D. kitumensis* the two overlapping inversions which transform it into *D. microlabis* are shown.
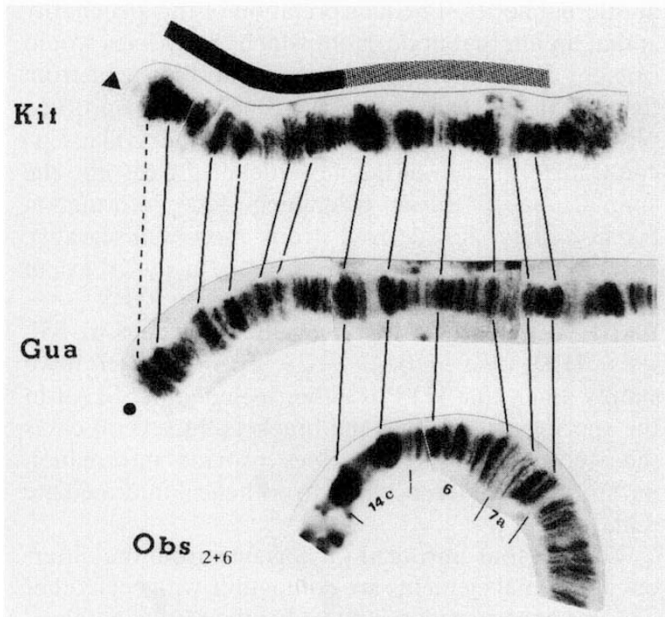

Fig. 4 The characteristic segment in the chromosomes of three species. (▲) The distal chromosomal end. *D.kitumensis* and *D. guanche* have both black and grey bar segments which are contiguous. *D. obscura* does not.

There is another important segment, encountered in its totality only in the species of the African and the *subobscura* clusters (Figs 2c, 3 and 4), which includes the already mentioned: 7a/6ba/14c. As a convenient

description we divide this large segment in two: a (represented by a black bar in the previously mentioned figures) and 7a/6ba/14c. In *D. guanche* it is present at the most proximal to the centromere euchromatic part (a + 7a/6ba/14c). Only *D. kitumensis* carries it intact in the African cluster and it starts at the very distal end and in an inverted order (14c/6ab/7a + a). One of the inversions differentiating this species from *D. microlabis* splits it in two and displaces segment 7a/6ba/14c away from a.

Thus the distribution among these species of segment a/7a/6ba/14c in one or two separate pieces makes the phylogenetic scenario provided be element D very plausible (Brehm & Krimbas, 1992), i.e. the African species are derived somewhere in the pathway between the *subobscura* cluster and *D. obscura*. The fact that segment a is contiguous with 7a/6ba/14c, both in the *subobscura* cluster and in one species of the African cluster but not in *D. subsilvestris*, indicates that this last species might have as its origin a branch located after the African species split. This argument leads to the acceptance that the separation of a from 7a/6ba/14c occurred twice in the *obscura* phylogeny, i.e. two independent breaks took place at the same chromosomal position, one between *D. kitumensis* and *D. subsilvestris* and another between *D. kitumensis* and *D. microlabis*. An alternative scenario would place this last species linearly in between *D. kitumensis* and *D.*
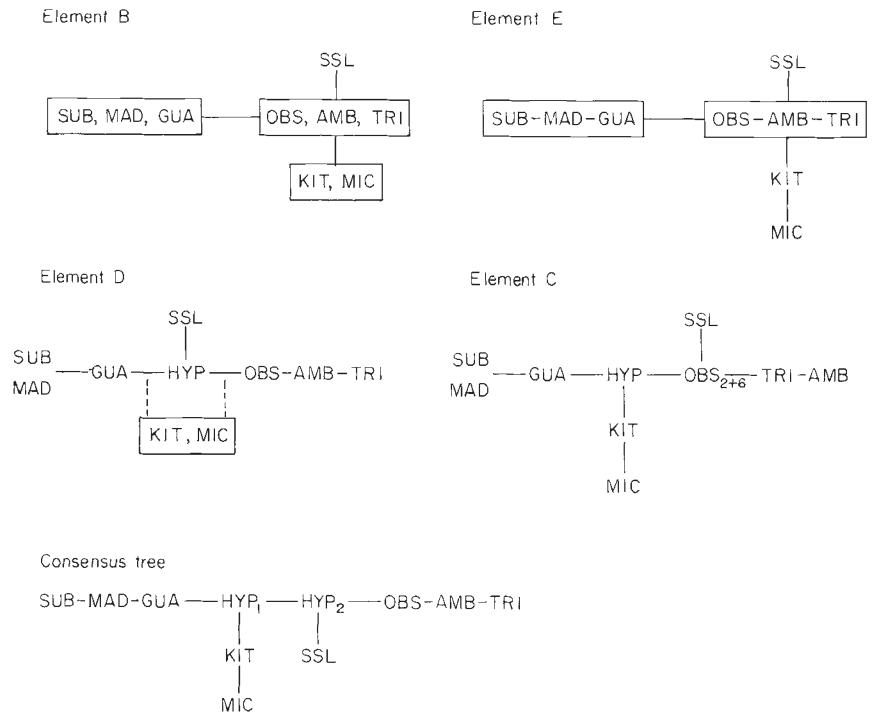
Fig. 5 Unrooted phylogenies derived from elements B, E, D, C and the consensus unrooted tree.

subsilvestris. It is difficult to decide between these two alternatives because an important number of chromosomal segments have not been homologized between *D. microlabis* and *D. subsilvestris*, but according to present research the former seems more plausible.

An unrooted tree comprising all the information gathered from element C is presented in Fig. 5.

Discussion

The four unrooted phylogenies constructed from independent data, provided by the study of the four autosomal elements, are shown in Fig. 5. These phylogenies carry different amounts of information. Boxes containing species indicate the fact that the specific species from those contained, which directly connect this cluster to another species or another box of species, remains unknown. In these boxes species may be ordered (connected by hyphens) or simply mentioned (separated by commas, their seriation remaining unknown). HYP symbolizes a hypothetical gene arrangement needed to connect those of the extant species studied. It might belong to one or the other species it connects (but was not yet encountered in them) or it might indicate the existence of a different, undiscovered or extinct, taxonomic unit.

This kind of schemata representing phylogenetic relations differs from the ones usually presented in binary trees in that extant species may occupy positions at internal nodes. The interpretation of this peculiarity is that an internal node, from which the species would branch, does not differ in its gene arrangement from that of this extant species. In this sense this unrooted phylogenetic tree is a collapsed one. Further collapse(s) can transform each of the more detailed trees into the simpler ones, those containing less information because they are derived from more conservative chromosomes, like element B, or to a lesser extent element E. Thus the seriation SSL —— [OBS —— AMB —— TRI] is topologically equivalent to SSL —— HYP —— OBS —— AMB —— TRI] if we simply admit that HYP is either included in SSL or in the species contained in the brackets. In several cases the study of an element does provide information permitting the inference of a hypothetical intermediate unit.

Thus all four unrooted phylogenies from the different autosomal elements are compatible with each other and this congruence is support for the results obtained and strongly corroborates the validity of the method. However, there is one exception found in the seriation of the species belonging to the *obscura* cluster. This seriation is indicated as OBS —— AMB —— TRI in two instances (in the phylogenies derived from elements E and D) and as OBS —— TRI —— AMB in the case of the phylogeny derived from element C. How is this exception to the general congruence to be explained?

Obviously such an exception indicates that one of the assumptions accepted is not absolutely valid. This is the assumption that gene arrangement polymorphisms do not transcend species separations (and that introgressions are also absent). If not, after their splitting, species may have retained both gene arrangements and subsequently fixed one of them in an order which indicates a phylogeny contrary to the real one. Note that when constructing phylogenies from gene or DNA sequence data, similar assumptions should be made.

Thus this assumption is violated in the case of the species of the *obscura* cluster for the gene arrangements of element C. The congruence of trees derived from independent sources of data indicates the rarity of such exceptions; transpecific polymorphisms are not generally encountered. In their absence the phylogenetic seriations are correct, while when they are present they may lead either to correct or to incorrect phylogenies, depending on the order of fixation of the gene arrangements. *A priori* there is no reason to consider that one or other alternative is favoured. Thus overall, taking into consideration both the cases of absence of transpecific polymorphism, and those of its presence, when studying many independent chromosomes the majority are expected to provide the correct species seriation. In our case the seriation OBS — — — AMB — — — TRI is more often encountered than its alternative, OBS — — — TRI — — — AMB, and although this is not a statistically significant difference, it points to the acceptance of the former seriation as the most probably correct.

Taking into account all the information provided from the four chromosomal elements we were able to construct a consensus tree. This tree is the richest, the most complete unrooted phylogeny encapsulating all the data gathered (Fig. 5).

The consensus tree is generally in good agreement with those derived from electrophoretic data. Several studies have appeared proposing phylogenies based on electrophoretic data; however, only three include most of the species we have studied (Lakovaara *et al.* 1974; Loukas *et al.*, 1984; Cariou *et al*, 1988). In the study by Lakovaara *et al.*, three of the species studied were not included (*D. madeirensis* and the two African species). The others are connected in a similar way to our consensus tree, except for two minor details. A similar disposition is also found in Loukas *et al.* (1984). The tree of Cariou *et al.* (1988) groups together all species of the *subobscura* cluster, but not in the right order; the position of the African species is between the branches of *D. tristis* and *D. ambigua*. *D. tristis* is the one closely related to *D. obscura* and this last is directly connected to the *subobscura* cluster. Electrophoretic data are

generally indicative but the information they provide is depauperated in comparison to gene sequence data or aminoacid sequence data; therefore we should not expect the complete recovery of the correct phylogenetic events. Even when 100 genes are studied the data do not permit the recovery of the correct topology when only eight species are studied (Krimbas & Sourdis, 1987). This indicates the interest that polytene chromosome data as well as molecular data have for phylogenetic tree construction.

References

BREHM, A. 1992. Photographic maps of polytene chromosomes of *Drosophila subobscura*. Appendix to C. B. Krimbas, The inversion polymorphism of *D. subobscura*. In: Krimbas, C. and Powell, J. R. (eds) *Drosophila Inversion Polymorphism*, CRC Press, Boca Raton, FL, pp. 198–204.

BREHM, A. AND KRIMBAS, C. B. 1990. The phylogeny of nine species of the *Drosophila obscura* group inferred by the banding homologies of chromosomal regions. II. Element E. *Hereditas*, **113**, 157–168.

BREHM, A. AND KRIMBAS, C. B. 1991. Inversion polymorphism in *Drosophila obscura*. *J. Hered.*, **82**, 110–117.

BREHM, A. AND KRIMBAS, C. B. 1992. The phylogeny of nine species of the *Drosophila obscura* group inferred by the banding homologies of chromosomal regions. III. Element D. *Genome*, in press.

BREHM, A., KRIMBAS, C. B., SOURDIS, J. AND CARIOU, M. L. 1991. The phylogeny of nine species of the *Drosophila obscura* group inferred by the banding homologies of chromosomal regions. I. Element B. *Genome*, **34**, 464–471.

CARIOU, M. L., LACHAISE, D., TSACAS, L., SOURDIS, J., KRIMBAS, C. AND ASHBURNER, M. 1988. New African species in the *Drosophila obscura* group: genetic variation, differentiation and evolution. *Heredity*, **61**, 73–84.

KRIMBAS, C. B. 1992. The inversion polymorphism of *Drosophila subobscura*. In: Krimbas, C. B. and Powell, J. R. (eds) *Drosophila Inversion Polymorphism*, CRC Press, Boca Raton, FL, pp. 127–220.

KRIMBAS, C. AND LOUKAS, M. 1984. Evolution of the 'obscura' group *Drosophila* species. I. Salivary chromosomes and quantitative characters in *D. subobscura* and two closely related species. *Heredity*, **53**, 469–482.

KRIMBAS, C. B. AND POWELL, J. R. 1992. Introduction. In: Krimbas, C. B. and Powell, J. R. (eds) *Drosophila Inversion Polymorphism*. CRC Press, Boca Raton, FL, pp. 1–52.

KRIMBAS, C. B. AND SOURDIS, J. 1987. Recent improvements on handling allelic isozyme data for tree construction. In: Rattazzi, M. C., Scandalios, J. G. and Whitt, G. S. (eds) *Isozymes: Current Topics in Biological and Medical Research*. A. R. Liss, New York, **15**, 49–62.

KUNZE-MÜHL, E. AND MÜLLER, E. 1958. Weitere Untersuchungen ueber der chromosomale Struktur und naturlichen Struckturtypen von *Drosophila subobscura*. *Chromosoma*, **9**, 559–570.

LAKOVAARA, S., SAURA, A., LANKINEN, P., POHJOLA, L. AND LOKKI, J. 1976. The use of isozymes in tracing evolution and in classifying Drosophilidae. *Zool. Scrip.,* **5**, 173–179.

LOUKAS, M., KRIMBAS, C. B. AND VERGINI, Y. 1984. Evolution of the 'obscura' group of *Drosophila* species. II. Phylogeny of ten species based on electrophoretic data. *Heredity,* **53**, 483–493.

MOLTO, M. D., DE FRUTOS, R. AND MARTINEZ-SEBASTIAN, M. J. 1987. The banding pattern of polytene chromosomes of *Drosophila guanche* compared with that of *D. subobscura. Genetica,* **75**, 55–70.

STALKER, H. D. 1972. Intergroup phylogenies in *Drosophila* as determined by comparison of salivary banding patterns. *Genetics,* **70**, 457–474.

STURTEVANT, A. H. AND DOBZHANSKY, TH. 1936. Inversions in the third chromosome of wild race of *Drosophila pseudo-obscura* and their use in the study of the history of the species. *Proc. Natl. Acad. Sci., U.S.A.,* **22**, 448–450.

WÜLKER, W., LÖRINCZ, G. AND DEVAI, G. 1984. A new computerized method for deducing phylogenetic trees from chromosome inversion data. *Z. Zool. Syst. Evolut/forschung,* **22**, 86–91.