Differing levels of dispersed repetitive DNA among closely related species of *Drosophila*

(cloned DNA library/sibling species/mutation rate/speciation)

ANDREW P. DOWSETT AND MICHAEL W. YOUNG

The Rockefeller University, New York, New York 10021

Communicated by James E. Darnell, Jr., April 26, 1982

The genomic concentrations of certain middle ABSTRACT repetitive DNA sequences vary considerably among closely related species of Drosophila. In fact, the chromosomes of D. melanogaster appear to carry approximately 3 times as much middle repetitive DNA as those of the sibling species D. simulans. Although most of the middle repetitive DNA of D. melanogaster consists of segments of "nomadic" DNA that occupy different dispersed chromosomal locations in different strains of flies, repeated DNA sequences recovered from the D. simulans genome are most often restricted to single chromosomal positions. Apparent differences in the total concentrations of middle repetitive DNA in the two species are most easily attributed to an approximately sevenfold difference in their dispersed repetitive and nomadic DNA contents. These differences may affect the relative mutation rates of these species or contribute to their reproductive isolation or both.

Chromosomally dispersed repeated DNA sequences will probably be found in large numbers in most eukaryotic genomes. Repeated DNA of this sort has been recognized, and in many cases extensively characterized, in a wide variety of organisms having only distantly related phylogenies (1). So ubiquitous are these DNA sequences, that they have justifiably come to be regarded as a fundamental, perhaps even an essential, component of the eukaryotic genome (for noteworthy exceptions, see refs. 2 and 3).

We have been studying this class of DNA sequences in Drosophila melanogaster, in which about 12% of the genomic DNA is repeated in a dispersed manner (4, 5). Four principal conclusions can be drawn from previous investigations. (i) The dispersed repetitive DNA of this species is composed of at least 50-100 different sequence families that, together, make up most of the middle repetitive DNA. (ii) Generally, 10-100 closely related member sequences form each family. (iii) The number of member sequences composing each family is conserved in different strains of this species. (iv) Despite this preservation of numbers, the chromosomal positions occupied by members of a family differ in different strains of the species. Since members of these repeated DNA families do not occupy fixed chromosomal locations, we refer to them as "nomadic."

These conclusions are based on analyses of libraries of cloned D. melanogaster chromosomal DNA. In this paper, similar methods are applied to a study of dispersed repetitive DNA in two sibling species of D. melanogaster, D. simulans, and D. mauritiana. All three species have similar morphologies and can be interbred to form viable, but not necessarily fertile, offspring. In this report, we show that, in contrast to previous findings with different strains of D. melanogaster, the sizes of many families of dispersed repetitive DNA vary among closely related

species. Moreover, some repeat sequence families that are present in D. melanogaster are likely to be missing altogether from the D. simulans and D. mauritiana genomes. D. melanogaster appears to have about 7 times as much dispersed repetitive DNA as D. simulans, so that a gain or loss of a few thousand nomadic DNA segments must have accompanied or followed the divergence of these two species.

MATERIALS AND METHODS

Preparation of Nucleic Acids. pDm and pDs plasmids were constructed by joining fragments of D. melanogaster DNA (Dm segments) or D. simulans DNA (Ds segments) to EcoRI-digested pBR325 (6). Fragments of Drosophila DNA were prepared for joining by digestion of total adult nuclear DNA with EcoRI. Clones were selected at random from two libraries of several thousand hybrid plasmids, and those carrying Dm or Ds segments >3 kilobases (kb) long were retained for further analysis. As indicated earlier, such clones may contain nonrepetitive or middle repetitive Drosophila DNA, but highly repeated satellite DNA is not usually obtained by these methods (7).

Plasmid DNAs were prepared as follows: Cells, grown overnight in 1.5 ml, were centrifuged for 1 min in an Eppendorf Microfuge and washed with 10 mM Tris HCl, pH 7.4/1 mM EDTA. Cells were then suspended in 200 μ l of 15% sucrose/ 50 mM Tris-HCl, pH 9.0/50 mM EDTA, and to this was added 85 μ l of lysozyme (Sigma) solution (5 mg/ml). This mixture was incubated for 30 min at room temperature, 600 μ l of 0.2 M NaOH/1% NaDodSO4 was added, and the solution was placed on ice for 5 min. Then, 300 μ l of 3 M potassium acetate (pH 6.0) was added, and the mixture was centrifuged as above. The supernatant, which contained primarily plasmid DNA, was recovered and plasmid DNA was precipitated with an equal volume of isopropanol. DNA was suspended in 100 μ l of 50 mM Tris·HCl, pH 7.4/25 mM EDTA and extracted once with isoamyl alcohol/chloroform/phenol (1:24:48). DNA was ethanol precipitated from the supernatant.

DNA was prepared from *D. melanogaster*, *D. mauritiana*, and *D. simulans* adults as modified from the method of Endow (8). Flies were ground in a loose fitting glass-glass homogenizer in a solution of 50 mM Tris HCl, pH 8.0/100 mM EDTA/0.35 M sucrose. Nuclei and cells were centrifuged at 4,000 rpm for 10 min in a DuPont HB-4 rotor and then gently suspended in 50 mM Tris HCl, pH 8.0/100 mM EDTA/0.5% NaDodSO₄. The suspension was rocked at room temperature for 5 min and incubated on ice for 15 min. For each milliliter of solution, 0.91 g of CsCl was added, and then ethidium bromide was added to $250 \mu g/ml$, and this solution was centrifuged for 60 hr in a Beckman type 40 rotor at 15°C. Ethidium bromide was removed

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: kb, kilobase(s); Dm, D. melanogaster; Ds, D. simulans.

from the DNA with isoamyl alcohol, and CsCl was removed by dialysis.

Hybridization Procedures. Hybridization of ³²P-labeled nick-translated (9) DNA to denatured restriction fragments that had been transferred to nitrocellulose has been described (10). The method of preparing and hybridizing [³H]cRNA probes to polytene chromosomes has also been described (11).

Enzymes. Restriction endonucleases were obtained from Bethesda Research Laboratories and New England BioLabs. *Escherichia coli* DNA polymerase was purchased from New England Nuclear, and *E. coli* RNA polymerase was a gift from P. Model.

Drosophila Stocks. D. melanogaster strain F9 was used in all experiments. This strain was constructed by A. P. Dowsett in 1979 from several laboratory stocks and should be homozygous for all chromosomes. A single strain of D. mauritiana and one strain of D. simulans were used in these experiments. Both were provided by K. Tartof.

RESULTS

Isolation of Middle Repetitive DNA from D. melanogaster and D. simulans Chromosomes. Ninety-four colonies, each propagating a different pDm plasmid, were selected at random from a library of several thousand clones. The average length of the Dm segments contained in the plasmids chosen for analysis was 6.4 kb. To identify pDm plasmids containing repetitive DNA, each was ³²P labeled by nick-translation and then hybridized to nitrocellulose blots of total genomic D. melanogaster DNA that had been digested to completion with EcoRI and fractionated by agarose gel electrophoresis. pDm plasmids carrying only unique sequence DNA hybridize to restriction fragments of one or a few lengths in accordance with the numbers and sizes of the *Eco*RI fragments contained in the pDm plasmids themselves. Each pDm plasmid containing repetitive DNA hybridizes not only to restriction fragments defining the Dm portion of the plasmid but also to fragments that cannot be found in the plasmid itself. In a few instances, pDm plasmids hybridize very intensely to genomic restriction fragments similar in length to those found in the plasmid. Such pDm plasmids are also assumed to carry repetitive DNA.

In this manner, 26 pDm plasmids containing repetitive DNA and 68 plasmids that apparently contain only nonrepetitive DNA were identified. Thus, about 28% of the plasmids in this D. melanogaster library carry repetitive DNA. This proportion is quite consistent with previous findings (4).

The proportions of plasmids containing repetitive and nonrepetitive DNA in a *D. simulans* library were determined in a similar fashion. A library of cloned *D. simulans* chromosomal DNA segments was constructed in a manner identical to that described for *D. melanogaster*. From this library of several thousand pDs plasmids, 126 were chosen for further analysis. The average length of the Ds segments in this collection of 126 pDs plasmids was 6.6 kb. pDs plasmids that carry repetitive and nonrepetitive DNAs were identified by hybridizing each to *Eco*RI-digested *D. simulans* genomic DNA. By using the methods outlined for pDm plasmids, 13 pDs plasmids containing repetitive DNA and 113 pDs plasmids that appear to carry only nonrepetitive DNA were identified. Accordingly, about 10% of the pDs plasmids in the *D. simulans* library contain repetitive DNA.

Hybridization of Repetitive pDm Plasmids and pDs Plasmids with Genomic DNAs from D. melanogaster, D. mauritiana, and D. simulans. Table 1 shows that repeated DNA iso-

Dm segment	simulans/ melanogaster	mauritiana/ melanogaster	Ds segment	melanogaster/ simulans	mauritiana/ simulans
Dm142	≤0.01	≤0.01	Ds268	0.32	0.51
Dm366	≤0.03	≤0.03	Ds99	0.40	0.18
Dm101	0.10	0.10	Ds258	0.63	0.33
Dm47	0.11	0.07	Ds255	0.67	0.37
Dm321	0.17	0.45	Ds168	0.73	1.13
Dm256	0.22	0.40	Ds181	0.84	0.73
Dm27	0.25	0.19	Ds137	0.97	1.27
Dm34	0.25	0.43	Ds193	1.03	1.09
Dm17	0.27	0.35	Ds262	1.09	0.83
Dm151	0.29	0.33	Ds205	1.15	0.76
Dm298	0.30	0.44	Ds357	1.36	0.68
Dm73	0.32	0.06	Ds272	1.58	0.51
Dm331	0.33	0.65	Ds246	3.57	1.27
Dm67	0.35	0.38			
Dm177	0.37	0.33			
Dm39	0.42	0.18			
Dm377	0.51	0.58			
Dm334	0.58	0.54			
Dm305	0.62	0.47			
Dm54	0.74	0.22			
Dm179	0.88	0.42			
Dm114	0.92	0.70			
Dm180	0.99	0.74			
Dm36	1.23	1.26			
Dm95	1.34	1.41			
Dm185	2.45	2.11			

Table 1. Relative abundance of DNA sequences homologous to repeated Dm and Ds segments in three sibling species of Drosonbilg

Data represent hybridization of Dm and Ds segments to *D. simulans*, *D. mauritiana*, or *D. melanogaster* total genomic DNAs expressed as percentage of Dm or Ds segment hybridization to conspecific DNA. Dm segments are listed in order of increasing hybridization to *D. simulans* DNA, and Ds segments are listed in order of increasing hybridization to *D. melanogaster* DNA:



lated from *D. melanogaster* is often less abundantly represented in *D. simulans* and *D. mauritiana* DNA than in conspecific DNA. Families of repeated Dm DNA can be reduced to 3% or less of their conspecific levels in the genomes of these sibling species, and some of these families may be present only in *D. melanogaster* (see also Fig. 1*a*). About one-half of the Dm families studied are reduced to less than 1/3 of their conspecific. values in one or both sibling species. The data in Table 1 were obtained by quantitative analysis of hybridizations such as those shown in Fig. 1. Each autoradiograph was traced with a microdensitometer according to the method of Lis *et al.* (13).

Table 1 also shows the results obtained when pDs plasmids are hybridized with the genomic DNAs of these three species. In contrast to the findings with repetitive Dm families, families of D. simulans repetitive DNA are usually found to be about as abundantly represented in D. melanogaster and D. mauritiana as they are in D. simulans.

Chromosomal Distribution of *D. simulans* and *D. melanogaster* Repetitive DNA. The experiments described thus far indicate that many families of repeated DNA sequences are more abundant in the chromosomes of *D. melanogaster* than in those of *D. simulans*, but they do not provide information

FIG. 1. Comparison of restriction fragments in D. melanogaster, D. simulans, and D. mauritiana total genomic DNAs homologous to repeated Dm and Ds segments. D. melanogaster (left lanes), D. simulans (center lanes), and D. mauritiana (right lanes) adult DNAs were digested with EcoRI, fractionated by electrophoresis in an 0.8% agarose gel, transferred to nitrocellulose as described (10), and hy-bridized to ³²P-labeled Dm142 sequences (a), Dm47 sequences (b), Ds137 sequences (c), Ds193 (d), and Ds246 sequences (e). Numbers on the right are lengths (kb) of HindIII fragments of phage λ DNA (from ref. 12).

that reveals what kinds of sequences might be disproportionately represented. For this reason, pDm and pDs plasmids carrying repetitive DNA were hybridized to salivary gland polytene chromosomes derived from *D. simulans* and *D. melanogaster*. This procedure distinguishes dispersed repetitive DNA from DNA sequences that are repeated at a single chromosomal location.

Table 2 shows that 12 of the 26 repetitive Dm segments isolated are hybridized to D. melanogaster polytene chromosomes, with the result that 9 (75%) are repeated in a dispersed fashion in this species. A similar frequency of dispersed reptitions (87%) was previously observed for 23 unselected repeated Dm segments (4). Evidently, as shown with two independently constructed D. melanogaster libraries, most Dm segments carrying repetitive DNA hybridize to multiple chromosomal locations. From the present results, it can be calculated that about 21% (0.75 × 0.28) of all pDm plasmids comprising the total D. melanogaster library prepared for these experiments contain dispersed repetitive DNA.

Table 3 shows the results obtained when repetitive Ds segments are hybridized to D. simulans polytene chromosomes. Chromosomal positions were determined for 10 of the 13 re-

Table 2. Localization of *D. melanogaster* polytene chromosomal sites homologous to repetitive Dm segments

Dm segments		
Plasmid	Dm segment length, kb	<i>D. melanogaster</i> genomic hybridization sites, no. and type
pDm27	7.0	21.5 ± 4.8 (6) euchromatic and chromocenter
pDm34	17.0	10.2 ± 2.4 (6) euchromatic and chromocenter
pDm39	14.5	34.0 ± 3.0 (6) euchromatic and chromocenter
pDm54	10.0	59.2 ± 5.1 (5) euchromatic and chromocenter
pDm73	5.5	18.0 ± 1.8 (6) euchromatic and chromocenter
pDm95	3.0	Chromocenter only
pDm151	4.5	41.2 ± 7.0 (6) euchromatic and chromocenter
pDm179	6.5	Chromocenter only
pDm180	4.5	1 euchromatic only
pDm185	6.5	41.2 ± 4.9 (6) euchromatic and chromocenter
pDm298	7.5	19.8 ± 1.8 (6) euchromatic and chromocenter
pDm321	7.5	19.5 ± 2.7 (6) euchromatic and chromocenter

Results represent mean \pm SD. Means were determined by counting several nuclei from a single fly. Values in parentheses represent numbers of nuclei counted.

Plasmid	Ds segment length, kb	D. simulans genomic hybridization sites, no. and type	D. melanogaster genomic hybridization sites, no. and type
pDs99	6.0	39.7 ± 3.8 (10) euchromatic and chromocenter	Not determined
pDs137*	8.5	Nucleolus only	Nucleolus only
pDs168	4.5	64.7 ± 5.6 (6) euchromatic and chromocenter	64.5 ± 4.7 (4) euchromatic and chromocenter
pDs193	3.0	Chromocenter only	Not determined
pDs205	3.0	Chromocenter only	Chromocenter only
pDs246	7.0	$2.5 \pm 1.0 (10)$ euchromatic and chromocenter	Not determined
pDs 255	10.0	Chromocenter only	Chromocenter only
pDs258*	8.0	Nucleolus only	Nucleolus only
pDs268	7.0	1 euchromatic only	Not determined
pDs357	4.0	Chromocenter only	Not determined

Table 3. Localization of polytene chromosomal sites homologous to repetitive Ds segments

Results represent mean \pm SD and were determined as in Table 2. Values in parentheses represent numbers of nuclei counted.

* After chromosomal localization, pDs137 and pDs258 were analyzed for, and found to contain, ribosomal DNA.

petitive Ds segments, but only 3 of these were found at multiple locations. Thus, in a survey of close to 100 pDs plasmids, only 3 have been found that contain dispersed repetitive DNA. Moreover, one of these dispersed repeated sequences occupies only two to three scattered chromosomal sites (see Table 3, pDs246). In contrast to the results for pDm plasmids, we calculate the frequency of plasmids carrying dispersed repetitive DNA to be about 3% (0.30 \times 0.10) in the total library of *D*. *simulans* cloned DNA segments. In summary, not only are fewer repetitive pDs plasmids recovered from the *D*. *simulans* library, but most of the repeats that have been isolated are restricted to single chromosomal regions. Examples of two pDs hybridizations are shown in Fig. 2.

The sevenfold difference in the rate of recovery of dispersed repetitive DNA segments from the D. melanogaster and D. simulans libraries does not apply to repeated DNA sequences that hybridize to single polytene chromosomal positions. As shown in Tables 2 and 3, pDm and pDs plasmids containing repeated DNA sequences that are restricted to a single chromosomal region are isolated from both libraries at the same frequency. Nondispersed repetitive DNA is carried by about 6% of the clones that form both libraries. As shown in Table 3, several pDs plasmids containing repetitive DNA that is restricted

to a single chromosomal region in D. simulans have also been hybridized to D. melanogaster polytene chromosomes. In each case, hybridization is limited to the same chromosomal region in the second species.

DISCUSSION

From previous work, it was concluded that most of the middle repetitive DNA in D. melanogaster is composed of nomadic DNA families, members of which are reiterated in a dispersed fashion (4). It has also been pointed out that a characteristic size is preserved for each family in different strains of this species, so that dispersed repeated and nomadic DNAs should contribute about 12% of the total genomic DNA in all strains of D. melanogaster (4). In this paper, we have reported that specific families of dispersed repetitive DNA can vary severalfold in concentration among the sibling species D. melanogaster, D. simulans, and D. mauritiana. If differences in the frequencies of total dispersed repetitive DNA in the D. melanogaster and D. simulans plasmid libraries accurately reflect variations in the concentrations of this class of DNA sequences in the genomes of these two species, then dispersed repetitive and nomadic DNA makes up only about 2% of the chromosomal DNA of D.



FIG. 2. Localization of repeated Ds segments in *D. simulans* polytene chromosomes by *in situ* hybridization. (a) [³H]cRNA transcribed from pDs255 was hybridized to *D. simulans* chromosomes. (b) Hybridization of pDs205 sequences. Arrows indicate hybridizations to chromocenters.



FIG. 3. Tentative phylogeny of members of the D. melanogaster species subgroup. Relationships are taken from Tartof (14); 1, 2, and 3 represent hypothetical intermediate species. Double-headed arrows indicate that neither the absolute nor the relative ages of these species have been determined

simulans. This seems to be the simplest interpretation of the data.

A phylogenetic tree showing the relationship of the three sibling species, D. melanogaster, D. simulans, and D. mauritiana, is presented in Fig. 3. Three additional species belonging to the species subgroup are also indicated. The phylogeny is based on the polytene chromosome banding patterns of all six species. The chromosome banding patterns of D. simulans and D. mauritiana are essentially homosequential, and those of D. melanogaster differ from these two relatives only by an inversion of part of chromosome arm 3R and four smaller inversions, none of which exceeds 20 chromosome bands in length. More complex rearrangements separate these species from the other members of the subgroup (15). The relatedness of these species can also be demonstrated by interspecies hybridizations (16, 17). D. simulans and D. mauritiana can be crossed to form viable and fertile offspring, while D. melanogaster and D. simulans, and D. melanogaster and D. mauritiana, interbreed to form viable but sterile progeny. D. mauritiana can be successfully crossed to D. teissieri, D. yakuba, and D. erecta, and progeny of these crosses are sterile. As expected, the adult morphologies of the six species are similar and, from the catalogue of gene homologies established by Sturtevant (18), it can be concluded that the genetic maps of at least D. simulans and D. melanogaster are probably equivalent if adjustments are made for the indicated chromosome inversions. A survey of all six species in the species subgroup has shown that, among a large collection of D. melanogaster-derived nomadic DNA sequences (14 cloned segments), not one is shared by all members of the species subgroup (unpublished). These and the observations we have reported here are easily understood if it is assumed that most nomadic DNA provides no function that is essential to the survival of insects or their populations within this species subgroup.

Interspecific differences in the sizes or numbers (or both) of these repeated sequence families may have interesting consequences. For example, there appears to be a remarkable correlation between the incidence of spontaneous mutation in D. melanogaster and the repositioning of segments of nomadic DNA (refs. 19-21, unpublished observations). In fact, spontaneous mutations correlate so well with rearrangements of nomadic DNA that it seems guite possible that a very sizeable fraction, perhaps half or more, of all such mutations recovered by standard genetic procedures in this species will eventually be attributed to changes in the positions of these DNA segments. If rearrangements of nomadic DNA prove to be a primary agent of mutagenesis, spontaneous mutation rates may be found to vary several-fold between species in a fashion that reflects quantitative differences in the nomadic DNA contents of their genomes.

Differing levels of nomadic DNA may also augment the reproductive isolation of closely related species. A large body of genetic data suggests that certain mating incompatibilities can be detected in crosses between different strains of D. melanogaster and that these appear to be linked to the behavior of a few families of dispersed repeated and nomadic genes (22-24). The families are unusual in that they are represented in very different numbers in different strains of this species (see also ref. 21). We suggest that some families of dispersed repeated and nomadic DNA that are abundantly represented in D. melanogaster but not in D. simulans or D. mauritiana may similarly contribute to the infertility seen in hybrids formed between these species.

We thank Irwin Gelman for technical assistance and Norton D. Zinder for enthusiastic discussion. We also thank Norton D. Zinder and James Darnell for helpful criticism of the manuscript. This work was supported by the Andre and Bella Meyer Foundation and the National Institutes of Health. A.P.D. was a Whitney Foundation Fellow.

- Davidson, E. H., Galau, G. A., Angerer, R. C. & Britten, R. J. 1. (1975) Chromosoma 51, 253-259.
- 2
- Timberlake, W. E. (1978) Science 202, 973–975. Krumlauf, R. & Marzluf, G. A. (1979) Biochemistry 18, 3705–3713. 3.
- Young, M. W. (1979) Proc. Natl. Acad. Sci. USA 76, 6274-6278.
- Young, M. W. & Schwartz, H. E. (1980) Cold Spring Harbor 5. Symp. Quant. Biol. 45, 629-640.
- Bolivar, F. (1978) Gene 4, 121-136. 6.
- Finnegan, D. J., Rubin, G. M., Young, M. W. & Hogness, D. S. (1977) Cold Spring Harbor Symp. Quant. Biol. 42, 1053-1063.
- Endow, S. A. (1980) Cell 22, 149–155. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) I. Mol. Biol. 113, 237-251.
- 10. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
- Wensink, P. C., Finnegan, D. J., Donelson, J. E. & Hogness, D. 11. S. (1974) Cell 3, 315-325.
- Philippsen, P., Kramer, R. A. & Davis, R. W. (1978) J. Mol. Biol. 12. 123, 371-386.
- 13. Lis, J. T., Prestidge, L. & Hogness, D. S. (1978) Cell 14, 901-919
- Tartof, K. D. (1979) Cell 17, 607-614 14.
- Lemeunier, F. & Ashburner, M. (1976) Proc. R. Soc. London Ser. 15. B 193, 275-294.
- David, J., Lemeunier, F., Tsacas, L. & Bocquet, C. (1974) Ann. 16. Genet. 17, 235-241.
- Lemeunier, F. (1979) Dissertation (Univ. of Paris, Paris, France). 17.
- Sturtevant, A. H. (1929) Carnegie Inst. Washington Publ. 399, 18. 1 - 62
- 19. Hogness, D. S., Bender, W. W., Akam, M. E., Saint, R. B. & Spierer, P. (1981) J. Supramol. Struct. Cell. Biochem. Suppl. 5, 385
- Marx, J. L. (1981) Science 213, 1485-1488. 20.
- Spradling, A. C. & Rubin, G. M. (1981) Annu. Rev. Genet. 15, 21. 219-264.
- Bregliano, J. C., Picard, G., Bucheton, A., Pelisson, A., Lavige, 22. [. M. & L'Heritier, P. (1980) Science 207, 606–611.
- 23. Kidwell, M. G., Kidwell, J. F. & Sved, J. A. (1977) Genetics 86, 813-833
- Engels, W. R. & Preston, C. R. (1981) Cell 26, 421-428. 24.