Metallothionein Gene Duplications and Metal Tolerance in Natural Populations of Drosophila melanogaster

G. Maroni, J. Wise, J. E. Young and E. Otto

Department of Biology and Curriculum in Genetics, University of North Carolina, Chapel Hill, North Carolina 27599 Manuscript received July 24, 1987 Accepted August 24, 1987

ABSTRACT

A search for duplications of the Drosophila melanogaster metallothionein gene (Mtn) yielded numerous examples of this type of chromosomal rearrangement. These duplications are distributed widely—we found them in samples from four continents, and they are functional—larvae carrying Mtn duplications produce more Mtn RNA and tolerate increased cadmium and copper concentrations. Six different duplication types were characterized by restriction-enzyme analyses using probes from the Mtn region. The restriction maps show that in four cases the sequences, ranging in size between 2.2 and 6.0 kb, are arranged as direct, tandem repeats; in two other cases, this basic pattern is modified by the insertion of a putative transposable element into one of the repeated units. Duplications of the D. melanogaster metallothionein gene such as those that we found in natural populations may represent early stages in the evolution of a gene family.

I T is a widely accepted hypothesis that gene duplication is one of the first steps in the evolution of new genes. Since the original function can be maintained by one of the two copies, the duplicate is able to escape the pressure of natural selection and accumulate the mutations necessary to develop a new function. If the process of duplication and divergence is repeated, there results a family of genes having related but not identical functions. Further divergent evolution could then give rise to genes with totally different functions (BRIDGES 1935; MUELLER 1936; OHNO 1970; OHTA 1987). Studies of this phenomenon have been limited by the scarcity of naturally occurring, functionally meaningful duplications of recent origin (MAEDA and SMITHIES 1986).

For some established gene families, DNA sequence analysis shows evidence of the ancestral duplications; however, these are old events, and mutations of different kinds have changed the sequences considerably. Examples include the γ -globin genes in mammals (SLIGHTOM, BLECHL and SMITHIES 1980) and the human haptoglobin genes (MAEDA *et al.* 1984).

Newly produced duplications of chromosomal segments are rare mutational events that are well-known to cytogeneticists. In general, however, it is not possible to ascertain the functional significance of these duplications outside of the experimental situation in which they were isolated. Studies of various loci in Drosophila, including rosy (GELBART and CHOVNICK 1979), maroon-like and rosy (SHAPIRA and FINNERTY 1986), Alcohol dehydrogenase (CHIA et al. 1985) and white (DAVIS, SHEN and JUDD 1987) are illustrative of attempts to determine duplication frequencies and molecular mechanisms involved in their origin.

Duplications of the Drosophila metallothionein (MT) gene (*Mtn*) may represent a more informative example of the early stages in the development of a new gene family. We first found a *Mtn* duplication in a laboratory population selected for cadmium resistance; larvae homozygous for that duplication produce twice as much MT RNA and tolerate increased concentrations of cadmium chloride (OTTO, YOUNG and MARONI 1986). The discovery of another duplication in an unselected laboratory strain suggested that Drosophila might be polymorphic for *Mtn* duplications and led us to a systematic search in samples from several wild populations. Here we present the results of that search.

MATERIALS AND METHODS

D. melanogaster cultures and stocks: The majority of the lines from natural populations used in this study were derived from single fertilized females ("iso-female" lines) captured in nature (Table 1). The exceptions were 52 of the North American lines represented in part A of Table 1; these were isogenic for the three major chromosomes (LAU-RIE-AHLBERG et al. 1980). The lines in Table 1B are a subset of the lines shown in Table 1A. The sources of the flies tested are as follows:

- Tropical Africa (Table 1A), a total of 153 lines from Botswana and Zambia: 18 lines from each country were chosen at random for the tests reported in Table 1B.
- Botswana (85 lines): collected in 1985 near a fruit and vegetable market in the Okavango River delta by ROBERT DORRIT; supplied to us by MARTIN E. KREITMAN.
- Zambia (68 lines): collected in 1985 near a fruit and vegetable market in the Luangwa River valley by ROBERT DORRIT; supplied to us by MARTIN E. KREITMAN.

- Egypt (15 lines): collected in 1985 at a vineyard in Gianaclis by JEAN DAVID who supplied them to us.
- Australia (10 lines): collected in 1982 at a vineyard in Victoria by I. A. MCKENZIE; supplied to us by MARGARET G. KIDWELL.
- United States (17 lines): collected in 1984 at a fruit and vegetable market in Raleigh, North Carolina, by MARTIN E. KREITMAN, who supplied them to us.
- France (14 lines): collected in 1985 at a vineyard in Sancerre by ANNIE FLEURIET; supplied to us by MARGARET G. KIDWELL.

Ninety-three other lines derived from flies collected in Europe, Japan and the United States were supplied by CATHY LAURIE-AHLBERG and CHARLES AQUADRO.

Tests of tolerance to copper ions: First-instar larvae were transferred to food supplemented with varying $CuSO_4$ concentrations, and viability was determined as the fraction of larvae that reached pupariation. [For details, see MARONI and WATSON (1985) and OTTO, YOUNG and MARONI (1986).]

Molecular analyses: For restriction enzyme analysis, DNA was extracted from samples of 15 flies by the method of LIS, SIMON and SUTTON (1983); it was digested with restriction enzymes, fractionated by gel electrophoresis, and processed for Southern analysis following standard procedures. Restriction maps were constructed on the basis of single and double restriction-enzyme digestions. The probes used were (1) cDm51, a full-length cDNA clone (LASTOWSKI-PERRY, OTTO and MARONI 1985), (2) two fragments derived from cDm51—one of them extending from the 5' end to a BamHI site in the middle of the cDNA and the other extending from that BamHI site to the 3' end, and (3) various genomic fragments of the Mtn region. Total RNA was extracted, fractionated by gel electrophoresis and processed for filter hybridization as previously described (OTTO, YOUNG and MARONI 1986).

Detection of Mtn duplications: A procedure was designed to test for the presence of Mtn duplications in a large number of strains derived from flies captured in the wild. For each strain, DNA was digested with the restriction enzyme BamHI and processed for Southern analysis; hybridization to a radioactive cDNA probe identified fragments bearing Mtn sequences. As can be seen from the restriction map in Figure 1, flies with a single copy of Mtn are expected to yield two BamHI fragments carrying Mtn sequences, a 3.5-kb fragment and either a 2.5-kb or a 7.0-kb fragment (depending on whether the BamHI at position 3.2 kb is present or not). Flies with an Mtn duplication should yield, in addition, a third fragment that joins the two Mtn copies. The size of this "junction fragment" depends on the size of the duplication and the position of its boundaries. For any Drosophila line having extra bands in its restriction pattern, we obtained 8-12 sublines homozygous for specific third chromosomes, the chromosome on which Mtn is located, by standard genetic methods using TM3 (LINDSLEY and GRELL 1968) as a balancer. These sublines were likewise tested for any restriction pattern abnormality; and, once chromosomes with exceptional patterns were identified, analyses with several restriction enzymes were made to confirm or rule out the existence of a duplication. For one duplication type, i35, we were unable to obtain homozygous viable flies; and the analysis was carried out in hemizygotes (over the Mtn deficiency $Df(3R)by^{10}$ (MARONI, OTTO and LASTOWSKI-PERRY 1986).

Because the test described above would not reveal duplications producing a junction fragment of the same size as one of the normal fragments, some of the lines were tested twice, once with BamHI and once with HindIII. This would allow detection of duplications which include Mtn and have at least one boundary between positions -2.5 and +15.0 in the restriction map (Figure 1); *i.e.*, any duplication smaller than 17.5 kb would be detectable.

RESULTS

Occurrence of *Mtn* **duplications:** Altogether we tested, using only *Bam*HI, 302 lines derived from flies captured throughout the world. Among 149 lines, representing all geographic areas except tropical Africa, 20 of them carried duplications; in contrast, among the remaining 153 lines, all from two sites in tropical Africa, we found no duplications (Table 1A).

We did a more extensive analysis, involving separate digestions with *Bam*HI and *Hin*dIII, on samples from locations for which we had at least 10 iso-female lines derived from one collection. This procedure verified the absence of duplications in a subset of 36 lines chosen from among the 153 lines from tropical Africa. At the other extreme, we found the highest incidence of duplications in the French collection: half of those lines carried duplications (Table 1B).

Restriction analyses of Mtn duplications: Duplications were classified into different types on the basis of restriction enzyme analyses. Note that duplications of the same type might be different at the nucleotide sequence level even though they would appear identical at the level of resolution of our restriction maps. In four different duplication types, H22, H46, B60 and H35, (Figure 1) the duplicated segments were arranged as direct tandem repeats. Two other duplication types, i60 and i35, appeared to have been derived from types B60 and H35, respectively (Figure 1); i60 was identical to B60 except for the presence of a 7-kb insertion near the 5' end of the downstream repeat. Insertion sequences were present in many copies in the genome of this strain, and we suspect that they represent a transposable element since they were present in different genomic locations in other strains (data not shown). Likewise, for i35, the duplication boundaries were indistinguishable from those of H35 on the basis of restriction maps, but there was a presumed transposable element insertion. In this case, the insertion was in the middle of the downstream Mtn gene, conceivably disrupting it.

Figure 2 shows *Hin*dIII fragments of the *Mtn* region obtained from wild-type and from duplication-bearing flies. Five of the duplication types show two bands, one corresponding to the normal 15-kb fragment and the other representing the duplicated gene. The intensity of hybridization for the two bands is similar, indicating that these flies carried no more than two copies of *Mtn* per haploid genome. In the sixth duplication, i35, an element bearing a *Hin*dIII site was inserted into one of the *Mtn* genes (Figure 1), leading to the production of three instead of two fragments

TABLE 1

Occurrence of Mtn duplications in natural populations

Origin	No. of lines tested	No. of duplications		No. of duplications of each type					
			Percent	H22	H46	B 60	H35	i 60	i35
A. BamHI digestion								_	
Tropical Africa	153	0	0.0						
Australia and Japan	35	2	5.7			2	(3)		
North America	79	10	12.7	3	1	5	(1)		1
Europe and North Africa	35	8	22.9			6	(2)	2	
Total	302	20		3	1	13	(6)	2	1
B. BamHI and HindIII digestion									
Botswana	18	0	0.0						
Zambia	18	0	0.0						
Egypt	15	2	13.3			2			
Australia	10	3	30.0				3		
United States	17	7	41.2	3		4			
France	14	7	50.0			3	2	2	
Total	92	19		3		9	5	2	

The samples in part B are a subset of those in part A. H35 was not detectable by digestion with BamHI because the junction fragment was the same size as one of the normal BamHI fragments (3.5 kb); the numbers in parenthesis represent strains known to carry H35, but may not include all of the strains in part A with that duplication; for that reason, the values in parenthesis were not added in to obtain the values given in the "No. of duplications" column. Each of the lines shown in part B were derived from a single fertilized female captured in nature. As an approximation, then, if we assume that each line carries four different third chromosomes the frequency of Mtn duplications in the chromosome pool would be one-fourth of the frequencies given in Table 1B.



FIGURE 1.-Restriction maps of the normal Mtn region and of six different duplications. Blocked section represents Min: hatched boxes are exons, white box is the intron and solid boxes are translated sequences (transcription is from left to right). Stippled areas represent uncertainty in the boundaries and junctions of the duplications (boundaries are the transitions between single-copy and duplicated sequences; junction is the transition from one repeat to the next). Backhatched boxes represent insertions. Abbreviations: B, BamHI, E, EcoRI; G, BglII; H. HindIII; P. PstI; S. SstII; kb. 10³ bases. The BamHI site shown in parentheses for the wild type is present only in some strains, among them all those with duplication type H22. The first HindIII site to the right of the gene is at position 15.0 kb. The duplication sizes are (from top to bottom): 2.2, 4.6, 6.0 and 3.5 kb.



Note that the geographic distribution of duplication

types does not appear to be uniform (Table 1, last six columns). B60 was found not only in the wild but also in several laboratory strains, most notably the third



1.16 1.19 1.05 - 0.90 1.12

FIGURE 2.—Southern blots of DNA from each of the duplication strains hybridized to a *Mtn* cDNA probe. The DNA was digested with *Hin*dIII. DNA from the wild-type strain Samarkand has only one band; all other lanes show, in addition, the junction fragments. The numbers at the bottom of the figure are the ratios between radioactive counts in the junction fragments and the normal fragments as obtained from a similar experiment using *Bam*HI. The strains used were the following: wild type, Samarkand; H22, $Dp(3;3)Mtn^{H22}$ (stock No. 350); H46, KA08; B60, R130, i60, K10-3; i35, K17-7. Because the duplication i35 is in a lethal-bearing chromosome, DNA was extracted from flies heterozygous for Df(3) by^{10} (deficient for *Mtn*) (MARONI, OTTO and LASTOWSKI-PERRY 1986) and a chromosome designated Ra1124 (bearing the duplication).

chromosome balancers TM1, TM2 and TM3 (LIN-DSLEY and GRELL 1968). Likewise, H46 was found in laboratory stocks and in samples of laboratory populations that had been kept under NaCl or CuSO₄ selection for several years (WALLACE 1982). H22 was the same duplication type we originally found in a laboratory strain selected for cadmium resistance (OTTO, YOUNG and MARONI 1986).

Functional tests of *Mtn* **duplications:** In order to test whether both copies of a duplicated gene were functional, we measured relative amounts of MT RNA in larvae homozygous for various duplications and in wild-type Samarkand larvae. Duplications H22, H46, B60, and i60 produced 1.7–2.1 times as much MT-RNA as wild-type controls while the strain carrying a duplication disrupted by insertion of a foreign element, i35, was indistinguishable from the wild type (Fig. 3).

With respect to their tolerance to CuSO₄, strains that produced elevated levels of MT RNA, H22, H46, B60 and i60 appeared more tolerant than either the Samarkand control or i35 (Figure 4). Similar results



1.0 2.1 1.7 1.8 2.0 0.8

FIGURE 3.—Autoradiograph of a nucleic acid blot from a wild type and several duplication strains hybridized to a cDNA probe of *Mtn*. The larvae were fed for 24 hr on 0.16 mM Cd²⁺-supplemented medium before extraction of nucleic acids. The numbers at the *bottom* of each lane were obtained by liquid scintillation counting of filters similar to the ones shown here. The value for each strain is relative to the wild type, and each value is the mean of three measurements obtained from separately extracted samples; standard errors were less than 10% of the value of each mean. The duplication H35 was not included in the analysis because homozygotes for that chromosome were not viable. The other strains were as in Figure 2.

were also obtained for tolerance to cadmium (data not shown). This agrees with the results of experiments in which a second functional copy of *Mtn* was introduced into wild-type Drosophila by germline transformation: transformed larvae were approximately 1.5 times more tolerant to cadmium than controls transformed with an inactivated *Mtn* (OTTO *et al.* 1987). These results also agree with our observations on the duplication $Dp(3;3)Mtn^{H22}$ (OTTO, YOUNG and MA-RONI 1986).

DISCUSSION

The observations presented in this paper begin to answer two questions that are crucial to the understanding of early stages in the evolution of gene families: what are the molecular mechanism(s) that give rise to duplications? what forces maintain gene duplications in a population?

Origin of *Mtn* **duplications:** Having found four clearly different types of *Mtn* duplication, we can be certain that at least four independent events were responsible for these duplications. We have no evidence, however, that all of the duplications of any one type were derived from a single ancestral chromosome: that is, when the same type of duplication appears in different areas, it is impossible to say



FIGURE 4.—Cu-tolerance of larvae homozygous for the various duplication types (except H35). Survival (%) was measured as the fraction of individuals reaching pupariation at each salt concentration as compared to larvae reared in medium lacking CuSO₄. Each data point is the average for three cultures, each started with 50 first instar larvae. The same strains as in Figure 2 were used.

whether this is because a *particular duplication* is very widespread or because a certain *type of duplication* tends to arise repeatedly.

Tandem duplications could arise through two general mechanisms: (1) any form of non-homologous breakage and reunion; (2) unequal but homologous recombination between DNA segments that flank the gene and have similar sequences. At this time, mechanism (1) cannot be ruled out for any of the duplications we have studied; and, while mechanism (2), in its simplest form, has been ruled out for one particular duplication, $Dp(3;3)Mtn^{H22}$ (OTTO, YOUNG and MA-RONI 1986), the other duplications presented in this paper could well have originated in this fashion.

A variation of mechanism (2) involves pairing and homologous recombination between transposable elements located on the 5' side of Mtn in one chromosome and on the 3' side in the homologue. This mechanism was apparently responsible for a duplication in the *white* locus of *D. melanogaster* (DAVIS, SHEN and JUDD 1987). However, for *Mtn* we found no sign that any duplication arose in this manner. This sort of unequal crossing-over leads to duplications in which a transposable element is left at the junction between the repeated units; and while we found instances of putative transposable elements associated with *Mtn* duplications, the inserts in question are located *within* one of the repeats rather than between them. Admittedly, if transposable elements did play a role in generating duplications, they could have moved out after the duplications were formed leaving behind traces that would be detectable only by a more detailed analysis of boundary and junction sequences.

Distribution of Mtn duplications: With respect to the forces that lead to a high frequency of Mtn duplications, we can draw two conclusions, the first with more confidence than the second: (1) Those forces are not exerted uniformly in the geographic range of the species. This conclusion is based on the differences in incidence of duplications between Europe and North America on the one hand and Tropical Africa on the other. (2) Those forces were not exerted uniformly during evolutionary time. This conclusion is suggested by the absence of any genes closely related to Mtn (MARONI, OTTO and LASTOWSKI-PERRY 1986): while another Drosophila metallothionein gene (Mto) exists (MOKDAD, DEBEC and WEGNEZ 1987), it is very distantly related to *Mtn*; the sequence similarities between Mtn and Mto are of the same magnitude as the similarities between these genes and Crustacean metallothioneins. The persistence of a high frequency of Mtn duplications should have led, at one time or another, to the fixation of duplications and therefore to the existence of a relatively extended Mtn gene family.

Our data do not elucidate the nature of the forces responsible for the high frequency of duplication. However, given the existence of several duplications of independent origin, it seems unlikely that *Mtn* duplications are maintained by linkage to some other, selected gene. For the same reason we judge it to be improbable that genetic drift accounts for the observed frequency of duplications. We are left with two possible explanations: natural selection and high forward-mutation rate.

Positive selective pressure in certain geographic areas coupled with a constant mutation rate could explain the high frequency of duplications in Europe and North America. The absence of a family of genes closely related to *Mtn* and the scarcity of duplications in African populations could be explained if (1) the selective pressures are historically quite recent and (2) in the absence of positive selection, duplications are efficiently eliminated from a population. Evidence that supports this hypothesis is that all H22 duplications we found are associated with the presence of a *Bam*HI restriction site at 3.3 kb that is absent from all other duplication types; thus suggesting that all H22 duplications are identical by descent. Evidence that argues against this view is the existence of i35. In this case one of the *Mtn* copies seems to be inactive; thus i35 should have no selective advantage and should be destined to be lost. Was it pure chance that we encountered this duplication, and then only in a single line?

Given the phenotype of the duplications we studied, increased tolerance to Cd and Cu, a positive selective pressure could result from historically recent high environmental levels of certain transition metals such as Cu, Cd or Hg. Industrial discharge is one possible source of these metals. Another possible source (one that is perhaps more relevant given the natural history of fruit flies), is the Cu-containing sprays that have been used in vineyards and orchards over the course of the last century, to combat bacterial and fungal infections.

The other force that could maintain duplications at a high frequency is a difference between forward and reverse mutation rates. If so, we would argue, as we did above, that there must be temporal and spatial localization of increases in the rate of occurrence of new duplications. This would imply, however, that most of the *Mtn* duplications isolated are independent events, even those within the same duplication type. The reason for the limited number of types of duplications would then have to be attributed to the molecular processes involved.

Finally, it should be noted that one phenomenon considered common in duplications, the increase in the number of copies beyond two, by unequal crossing-over [KARIN *et al.* (1984) for instance], was not observed in our samples: all strains analyzed had only two copies of Mtn.

We thank N. G. HAIRSTON, D. F. KUBAI and J. C. LUCCHESI for their valuable suggestions and S. WHITFIELD for the artwork. We are especially grateful to the Drosophila workers who made hundreds of strains available for analysis (see Table 1). Supported by U.S. Public Health Services grant ES02654.

LITERATURE CITED

- BRIDGES, C. B., 1935 Salivary chromosome maps. J. Hered. 26: 60-64.
- CHIA, W., C. SAVAKIS, R. KARPS, H. PELHAM and M. ASHBURNER, 1985 A mutation of the Adh gene of Drosophila melanogaster containing an internal tandem duplication. J. Mol. Biol. 186: 679-688.
- DAVIS, P. S., M. W. SHEN and B. H. JUDD, 1987 Asymmetrical pairings of transposons in and proximal to the white locus of *Drosophila* account for four classes of regularly occurring exchange products. Proc. Natl. Acad. Sci. USA 84: 174-178.

- GELBART, W. M., and A. CHOVNICK, 1979 Spontaneous unequal exchange in the rosy region of *Drosophila melanogaster*. Genetics 92: 849-859.
- KARIN, M., R. NAJARIAN, A. HASLINGER, P. VALENZUELA, J. WELCH and S. FOGEL, 1984 Primary structure and transcription of an amplified genetic locus: the CUP1 locus of yeast. Proc. Natl. Acad. Sci. USA 81: 337-341.
- LASTOWSKI-PERRY, D., E. OTTO and G. MARONI, 1985 Nucleotide sequence and expression of a *Drosophila* metallothionein. J. Biol. Chem. **260:** 1527–1530.
- LAURIE-AHLBERG, C. C., G. MARONI, G. C. BEWLEY, J. C. LUCCHESI and B. S. WEIR, 1980 Quantitative genetic variation of enzyme activities in natural populations of *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA 77: 1073-1077.
- LINDSLEY, D. L., and E. H. GRELL, 1968 Genetic variations of Drosophila melanogaster. Carnegie Inst. Wash. Publ. 627.
- LIS, J. T., J. A. SIMON and C. A. SUTTON, 1983 New heat shock puffs and β -galactosidase activity resulting from transformation of Drosophila with an *hsp70-lacZ* hybrid gene. Cell **35:** 403– 410.
- MAEDA, N., and O. SMITHIES, 1986 The evolution of multigene families: human haptoglobin genes. Annu. Rev. Genet. 20: 81– 108.
- MAEDA, N., F. YANG, D. R. BARNETT, B. H. BOWMAN and O. SMITHIES, 1984 Duplication within the haptoglobin Hp² gene. Nature **309**: 131–135.
- MARONI, G., and D. WATSON, 1985 Uptake and binding of cadmium, copper and zinc by *Drosophila melanogaster* larvae. Insect Biochem. **15:** 55-63.
- MARONI, G., E. OTTO and D. LASTOWSKI-PERRY, 1986 Molecular and cytogenetic characterization of a metallothionein gene of Drosophila. Genetics 112: 493–504.
- MOKDAD, R., A. DEBEC and M. WEGNEZ, 1987 Metallothionein genes in *Drosophila* constitute a dual system. Proc. Natl. Acad. Sci. USA 84: 2658-2662.
- MUELLER, H. J., 1936 Bar duplication. Science 83: 528-530.
- OHNO, S., 1970 Evolution by Gene Duplication. Springer-Verlag, New York.
- OHTA, T., 1987 Simulating evolution by gene duplication. Genetics 115: 207-213.
- OTTO, E., J. E. YOUNG and G. MARONI, 1986 Structure and expression of a tandem duplication of the *Drosophila* metallothionein gene. Proc. Natl. Acad. Sci. USA 83: 6025-6029.
- OTTO, E., J. M. ALLEN, J. E. YOUNG, R. D. PALMITER and G. MARONI, 1987 A DNA segment controlling metal-regulated expression of the *Drosophila* metallothionein gene, *Mtn.* Mol. Cell. Biol. 7: 1710–1715.
- SLIGHTOM, J. L., A. E. BLECHL and O. SMITHIES, 1980 Human fetal G₇- and A₇-globin genes: complete nucleotide sequences suggests that DNA can be exchanged between these duplicated genes. Cell **21**: 627–638.
- SHAPIRA, S. K., and V. G. FINNERTY, 1986 The use of genetic complementation in the study of eukaryotic macromolecular evolution: rate of spontaneous gene duplication at two loci of *Drosophila melanogaster*. J. Mol. Evol. 23: 159–167.
- WALLACE, B., 1982 Drosophila melanogaster populations selected for resistance to NaCl and CuSO₄ in both allopatry and sympatry. J. Hered. 73: 35-42.

Communicating editor: A. CHOVNICK