49. A New Family of Mobile Dispersed Middle Repetitive Elements in *Drosophila melanogaster*

By Ko HARADA, Akihiko KOGA, Shin-ichi KUSAKABE, and Terumi MUKAI
Kyushu University 33, Fukuoka

(Communicated by Motoo KIMURA, M. J. A., Sept. 12, 1988)

Restriction fragment variation in the 13 kb region surrounding the *Adh* structural gene was examined among chromosomes sampled from a Raleigh, North Carolina population (Koga *et al.* 1988). In this survey, an insertion element was found to be polymorphic in *In(2L)t*-carrying chromosomes. This element (insertion "i" of Koga *et al.* 1988) is 2.6 kb and is located about 1.7 kb downstream from the proximal end of the *Adh* transcriptional unit. We cloned and sequenced this element. This element was found to occupy dispersed chromosomal locations with 10 to 30 copies per genome. These studies suggest that this sequence is a new family of dispersed middle repetitive sequences or transposable elements. We also have experimental results that this element is actually movable. The results are reported here.

Materials and methods. *Cloning and DNA sequencing of insertion "i".* The bacteriophage including insertion "i" was isolated from the library of R375, one of the six "i"-carrying strains, by the method of plaque hybridization using ³²P-labeled Probe 2 (Koga *et al.* 1988). This probe is a subcloned fragment prepared from the library of an "i"-free strain (C160) and covers the insertion region of "i". The clone isolated from the library of R375 is "B0". Base sequences of the whole range of insertion "i" in B0 and the 400 bp region of Probe 2 corresponding to the insertion site of "i" were sequenced by the dideoxy method as modified by Hattori and Sakaki (1986).

In situ hybridization. Salivary gland chromosomes of mutation accumulation lines of Mukai and Cockerham (1977) were fixed on slides and denatured with 0.07 N NaOH for 2 min. The probe DNA was labeled with biotin-labeled dUTP (ENZO) by nick translation (Rigby *et al.* 1978) and hybridized as described by Pardue and Gall (1975). The probe DNA was detected using the ENZO detection kit (DETEK-If) under a fluorescent microscope.

Result and discussion. *Insertion sequence "i".* Restriction maps were constructed from clone B0 and an "i"-free clone, Probe 2, as shown in Fig. 1. Comparing these maps localizes the insertion to between the Eco RI and Sal I sites in Probe 2 (B0-C160 in Fig. 1). The regions between these two restriction sites in both strains, about 0.4 kb in C160 and about 3.0 kb in R375, were sequenced, and the precise position of the insertion was determined by aligning the two sequences. Insertion "i" is 2581 bp long, and a fraction of the host genome is duplicated at both ends of the insertion. Several alignments are possible for the sequences. In the alignment that maximizes the length of the duplicated sequences, the length of the duplication is 12 base pairs (ATGGGGTGTGAA). The sequence was compared with known sequences of transposable elements and genes in *Drosophila* species using EMBL (Release 10.0) and GENBANK (Release 48.0); however, no conspicuous homology was detected. The complete sequence of insertion "i" also showed that it has no

Fig. 1. Restriction maps of DNA regions showing the position of insertions. B0 is the clone including insertion "i" isolated from R375. B0-C160 is the corresponding chromosomal clone isolated from the "i"-free strain, C160. B4 is a clone isolated from R375 including an insertion partially homologous to insertion "i". B4-C160 is the corresponding chromosomal clone isolated from C160. Subcloned fragments are shown for B0 and B4. Restriction sites are indicated by vertical lines: B, Bam HI; E, Eco RI; H, Hin dIII; P, Pst I; S, Sal I; Sc, Sac I; X, Xho I and Xb, Xba I. Pst I and Xba I sites are shown only in insertions.

apparent open reading frames nor inverted or direct terminal repeats.

Specificity of the host duplications. Using an inner region of insertion "i" (1.6 kb Eco RI-Hin dIII fragment, see Fig. 1), DNA fragments which have homology with insertion "i" were cloned from the library of R375. Three clones (B4, B7 and B9) were chosen randomly from twenty such clones, and restriction maps were constructed for them. Based on the comparison of the restriction maps of these clones, the right end of each insertion was identified and the fragments containing these regions were subcloned. The subclones were sequenced and the regions corresponding to the right end of insertion "i" are aligned in Fig. 2 with that of B0. The sequences from all four clones are identical from the left side to C at -1 (the scale over the sequences was arbitrary). Although B4, B7 and B9 have identical sequences up to G at 11, the remainders of their sequences are different. This observation suggests that the four insertions may be heterogeneous in length and have different host duplication sequences. In order to test this possibility, the insertion site was determined for clone B4. Using the small fragment located at the right side of the insertion in B4 as probe, the chromosomal region corresponding to the sequences covering the insert portion in B4 was cloned from the library of C160 (B4-C160). Regions around the left side of the insertion in B4 and the corresponding region in the genomic clone were subcloned and sequenced. The exact host DNA duplication for B4 was determined by comparing these two sequences, and is ATACTTTTGGGT. It has a maximum length of 12 bp and is different from that of B0. The left duplication is attached to a long poly(A) tail of 29 base pairs. The restriction maps of the subcloned fragment of B4 and the insertion-free clone B4-C160 are also shown in Fig. 1. Comparing these maps reveals that the full length of the insertion in B4 is 5.0 kb. The family of insertion sequences including "i" was called the *BN* element. Heterogeneity in length among copies is often observed

```

      -20      -10      0      +10      +20      +30      +40
      .....*.....*.....*.....*.....*.....*.....*.....
B0  --GAACACGTCCTCGATTGCTCGATGGGGTGTGAAAGAGTAATGACAAAATTTGGTGTGCCCAA--
      .....*.....*.....*.....*.....*.....*.....*.....
B4  --GAACACGTCCTCGATTGCTCATCTCCCATGCGAATACTTTGGGTGGCGGTGGCTGTCGGGCG--
      .....*.....*.....*.....*.....*.....*.....*.....
B7  --GAACACGTCCTCGATTGCTCATCTCCCATGCGAATGATTATAATAATACTTAATTAATT--
      .....*.....*.....*.....*.....*.....*.....*.....
B9  --GAACACGTCCTCGATTGCTCATCTCCCATGCGCTAAATAACTAAGAAGGCAGTGTCGAATTT--

```

Fig. 2. Base sequences of the region around the right end of the *BN* elements. The underlined sequences in B0 and B4 are the verified host genome duplications at maximum size. The scale for the sequence was assigned arbitrarily.

Table I. Insertion sites of the *BN* element in the mutation accumulation lines of Mukai and Cockerham (1977)

AW													
No. lines	21F	22D	24E	25E	29F	31A	39E	39F	43F	45B	45F	50C	50F
	1	29	29	1	29	29	29	29	29	29	2	29	1
	51A	53B	53F	57C	57F	58B							
	1	29	29	29	1	29							
JH													
No. lines	24D	29F	32A	32F	34A	34C	34D	39E	39F	43F	44F	53B	53F
	28	28	28	28	28	28	28	28	28	28	28	27	28

29 AW lines and 28 JH lines were examined. Chromosomal sites were assigned following the chromosome map in Lefevre (1976).

in other well-known transposable element families; for example, *FB* (Potter *et al.* 1980; Truett *et al.* 1981), *I* (Fawcett *et al.* 1986) and *P* (O'Hare and Rubin 1983). Target site specificity is not rigidly required for many transposable elements except for the sequence (TATA) preferred by element 297 (Ikenaga and Saigo 1982). The molecular structure of the *BN* element shares some common features with the *F* element in lacking terminal repeats and having a long poly(A) tail (cf. DiNocera *et al.* 1983).

Mobility of the BN element. The molecular structure of the *BN* element and its distribution in the genome inferred from the restriction fragment analysis indicates that the *BN* element is a middle repetitive sequence and may be a transposable element. To test whether or not this element is movable, we examined the distribution of copies of the *BN* element in the mutation accumulation lines of Mukai and Cockerham (1977). Spontaneous mutations have been accumulated for about 400 generations on two sets of second chromosomes (AW and JH) with minimum pressure of natural selection. In each set of lines, mutations have been accumulated both on the marked inversion chromosome, *Cy* (*In*(2*LR*)*SM1*) and the lethal second chromosomes (AW and JH) each of which originated from a single stem chromosome. The insertion sites of the *BN* element on 29 AW and 28 JH second lethal chromosomes were detected by *in situ* hybridization using clone B0 as probe. The results are listed in Table I.

There are 13 common sites for both the AW and JH sets of lethal chromosomes, and 7 additional sites in a total of 6 lines among the 29 AW lethal chromosomes examined. Two new sites (50F and 57F) occurred in the same line (AW40). No additional sites occurred among the 28 JH lethal chromosomes examined. Thus the insertion rate is $7/(400 \times 57 \times 20 \times 2) = 7.7 \times 10^{-6}$ (/element/generation) (57 lines with an average copy number of 20 *BN* elements per haploid genome). Only one site loss was observed in one of the JH lines (53B in JH362), so the excision rate is $1/(400 \times 57 \times 13) = 3.4 \times 10^{-6}$ (/element/generation) (57 lines with 13 copies per second chromosome). Little is known about the insertion or deletion rate in other transposable elements. Biémont and Aouar (1987) estimated an insertion rate of 10^{-3} (/element/generation) for *mdg-1*. Recalculation of the data of Eggleston *et al.* (1988) suggests an average insertion rate of 5.0×10^{-5} (/element/generation) following P-M dysgenic crosses to 1.3×10^{-5} (/element/generation) following non-dysgenic crosses for 15 different families of transposable elements which exclude *P*, *I* and *hobo* elements. The excision rate was estimated to be 7.5×10^{-5} (/element/generation) for 12 of these transposable elements (Eggleston *et al.* 1988) in the P-M dysgenic lines but no excision was observed in the non-dysgenic lines. The insertion and excision rates obtained for the *BN* element are about one order of magnitudes smaller than these values, so, the *BN* element is probably relatively stable. Although the mode of transposition is not known, sequencing of the longer *BN* element in B4 is in progress and some clues for solving this problem are expected soon.

Acknowledgments. We thank Dr. Y. Sakaki for his technical advice. We also thank Dr. T. F. C. MacKay, North Carolina State University, for reading the manuscript. This experiment was supported by research grants from the Ministry of Education, Science and Culture of Japan and by the Toray Science Foundation. This is paper no. 51 from the Laboratory of Population Genetics, Kyushu University.

References

- 1) Biémont, C., and Aouar, A. (1987): *Heredity*, **58**, 39-47.
- 2) DiNocera, P. P., Digan, M. E., and Dawid, I. B. (1983): *J. Mol. Biol.*, **168**, 715-727.
- 3) Eggleston, W. B., Johnson-Schlitz, D. M., and Engels, W. R. (1988): *Nature*, **331**, 368-370.
- 4) Fawcett, D. H. *et al.* (1986): *Cell*, **47**, 1007-1015.
- 5) Hattori, M., and Sakaki, Y. (1986): *Anal. Biochem.*, **152**, 232-238.
- 6) Ikenaga, H., and Saigo, K. (1982): *Proc. Nat. Acad. Sci. USA*, **79**, 4143-4147.
- 7) Koga, A. *et al.* (1988): (in preparation).
- 8) Lefevre, G. Jr. (1976): *The Genetics and Biology of Drosophila*. vol. 1a, Academic Press, N.Y., pp. 32-66.
- 9) Mukai, T., and Cockerham, C. C. (1977): *Proc. Nat. Acad. Sci. USA*, **74**, 2514-2517.
- 10) O'Hare, K., and Rubin, G. M. (1983): *Cell*, **34**, 25-35.
- 11) Pardue, M. L., and Gall, J. G. (1975): *Methods Cell Biol.*, **10**, 1-16.
- 12) Potter, S. S. (1982): *Nature*, **297**, 201-204.
- 13) Rigby, P. W. J. *et al.* (1977): *J. Mol. Biol.*, **113**, 237-251.