A transposable genetic element associated with positive regulation of G6PD gene expression in *Drosophila melanogaster*

MASANOBU ITOH, MARI IWABUCHI*, NAOKI YORIMOTO† AND SAMUEL H. HORI‡
Department of Zoology, Faculty of Science, Hokkaido University, Sapporo 060, Japan

(*Received 8 December 1987 and in revised form 8 April 1988*)

Summary

The DNA structures around the G6PD coding region in three high-G6PD activity mutants and their low-activity revertants of *Drosophila melanogaster* were analysed by Southern blot using a cloned G6PD gene as a probe. As a result, two kinds of insertion sequences were found; one was present just 5′ to exon I (Ins1), and the other within the intron (Ins2). The Ins1 sequence was 3·5 Kb in two mutants and 2·9 Kb in one mutant. In both cases, it consisted of a core sequence either 1·2 or 0·6 Kb long flanked by terminal repeats. On the other hand, low-activity revertants possessed either a defective Ins1 or no Ins1. The Ins2 sequence was found in all mutants and revertants, but not in Canton S. Although a recombinant phage carrying the DNA fragment spanning the entire Ins1 has not been obtained, sequencing data of the clone containing only the terminal repeats demonstrated that the repeats are defective P elements. Comparison of the genomic DNA structures of mutants and revertants suggested that the element responsible for the positive regulation of the G6PD gene in the mutants would probably be the core sequence, but not the flanking defective P elements. It was also conjectured that the 1·2 Kb core sequence might be composed of two identical elements, which might transpose independently.

1. Introduction

In *Drosophila*, more than 10% of genomic DNA is composed of middle repetitive DNA sequences, and some of them, known as transposable elements, change their positions in the genome. Such transposable elements often alter expression of nearby genes upon transposition (Young, 1979; Spradling & Rubin, 1981), and the alterations are suppressive in almost all cases. In this respect, an *X*-linked mutation recently found in our laboratory appears to be rare in that it causes overproduction of glucose-6-phosphate dehydrogenase (G6PD), probably as a result of insertion of a transposable element in the vicinity of the G6PD coding region (Tanda & Hori, 1983*a*, *b*). It is not certain at present whether such a transposable element has any adaptive and evolutionary significance, but it provides good material for discussing the biological role of transposable elements.

This mutation was found in a strain, 2512H, which had been subjected to artificial selection for high-G6PD activity (Tanda & Hori, 1983*a*, *b*). It appears to be a regulatory mutation, since the copy number of the G6PD gene in this mutant does not differ from that in a low-G6PD activity revertant, 2512L (Iwabuchi, Hori & Yorimoto, 1986). Evidence suggesting that this mutation probably resulted from insertion of a transposable genetic element has been provided by experiments initially aimed at establishing the *X*-linkage of the high-G6PD activity trait in the mutant strain.

When 2512H males were mated to Muller-5 females and the female progeny backcrossed to their fathers, to extract new *X* chromosome lines, the G6PD activity level of some of these lines became unstable, and tests showed that they were segregating *X* chromosomes with and without the functional regulatory factor. When the reciprocal cross was made, mating Muller-5 males to 2512-females and backcrossing male progeny to 2512H females, the resulting *X* chromosome lines lost the regulatory factor less frequently. This implies a contribution of the Muller-5 cytoplasm to the loss of the regulatory factor, the phenomenon thus being analogous to hybrid dysgenesis (Kidwell, Kidwell & Sved, 1977; Bregliano & Kidwell, 1983; Engels, 1983; O'Hare, 1985; Simmons & Karess,

* Present address: Plantech Research Institute, c/o MCI Research Center, 1000 Kamoshida-cho, Midori-ku Yokohama 227, Japan.
† Present address: Kao Corporation Tochigi Research Laboratories, 2606 Akabane, Ichikaimachi, Haga, Tochigi 321-34, Japan.
‡ Corresponding author.

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press

1985). However, detailed examination of the effect of cytotype showed that this phenomenon cannot be simply explained by the P-M system of hybrid dysgenesis and led us to conclude that the *X*-linked, regulatory factor of the mutant strain is stable only within the cytoplasm of this strain, and becomes unstable in the cytoplasm of any other strains including strain 2512L which shares the same *X*, second and third chromosomes with the 2512H strain (Itoh & Hori, 1985).

All these findings suggest that the *X*-linked, high-G6PD activity mutation may have arisen from insertion of a novel class of transposable elements in the vicinity of the G6PD locus. It is worth noting that such an *X*-linked mutation is not a rare instance that we have fortuitously encountered, since similar mutations have been obtained by subjecting wild-caught flies of high G6PD activity to artificial selection for high-G6PD activity (Iwabuchi, Hori & Yorimoto, 1986; this report). The *X*-linked mutations thus obtained share the same characteristics as the 2512H strain in that the high-G6PD activity trait becomes unstable upon outcrossing.

In order to obtain some physical evidence for our notion that such mutations may be associated with insertion of a transposable element, we have examined in the present study the genomic DNA structure around the G6PD coding region of the mutant and revertant strains using a cloned G6PD gene as a probe. As a result, it was found that the mutants do possess insertion sequences upstream of the G6PD gene which are absent in wild-type flies.

2. Materials and methods

(i) Drosophila *strains*

(*a*) Canton S: a standard wild-type laboratory strain classified as an M strain in the P–M system of hybrid dysgenesis (Kidwell, 1979).

(*b*) Harwich: a strong P strain (Kidwell, 1979).

(*c*) Muller-5: an *X* chromosome balancer stock marked by *white apricot* (w^a) and *Bar eyes* (*Bar*) (Lindsley & Grell, 1968). (*d*) M–H: a P strain which carries the *X* chromosome of the Muller-5 strain and the autosomes of the Harwich strain (Itoh & Hori, 1985).

(*e*) *C(1)DX:* an attached *X* stock marked by *yellow body* (*y*), *white eyes* (*w*) and *forked bristles* (*f*).

(*f*) *FM7:* an *X* chromosome balancer stock marked by *Bar* (Merriam, 1968, 1969).

(*g*) 2512H and 2512L: the high- and low-G6PD activity lines derived from strain 2512. This was homozygous for the *X* and third chromosomes, and carried a second chromosome kept balanced against the *SM1 Cy* chromosome. It responded to selection for high-G6PD activity and yielded 2512H. During the course of upward selection, a female of low-G6PD activity was selected and its progeny were subjected to downward selection to yield a low-activity revertant line, 2512L (Tanda & Hori, 1983*a*, *b*). Results of the standard A and A* crosses showed that both H and L lines are Q in the P–M system.

(*h*) 1FH and 1FL: these strains were derived from an isofemale line from a natural population; an *X* chromosome from a wild-caught female exhibiting a high-G6PD activity was extracted by crossing with an *FM7* male, and one of the *X* chromosome lines showing high-G6PD activity was subjected to upward selection for G6PD activity. After 3 generations, a high-activity line was obtained (1FH). At the first generation of the upward selection a female showing low-G6PD activity was selected and its progeny were subjected to downward selection to yield a low-activity line, 1FL.

The high-activity factors of 2512H and 1FH are *X*-linked, *cis*-dominant, and map 0·02 and 0·04 unit to the left of the G6PD locus (Iwabuchi, Hori & Yorimoto, 1986).

(*i*) S44H and S44L: strain S44 was established from a wild-caught female exhibiting a high level of G6PD activity. Upon upward selection during three generations, this isofemale line became very high in G6PD activity. Further bidirectional selection yielded S44H and S44L.

(*j*) rA31310, rA3146, rC155, rC159, rC1510, rC165, rC166, rD1391 and rD13103: these are revertant lines to low-G6PD activity, obtained by outcrossing 2512H males to Muller-5 females (rA lines), M–H females (rC lines) and 2512L females (rD lines), respectively, to establish a number of lines isogenic for 2512H *X* chromosomes (See Itoh & Hori, 1985 for details of the *X* chromosome line construction). These lines originally showed high-G6PD activities without exception, thus indicating the *X*-linkage of the high-activity trait. However, several low-activity revertants soon appeared in the succeeding generations (Itoh & Hori, 1985). Some males of these lines were individually mated with *C(1)DX* females, and low-activity revertants were selected and have been maintained by crossing with *C(1)DX* females.

(ii) *DNA preparation*

About 4 g of frozen flies were ground in a mortar with liquid nitrogen and homogenized in 20 ml of 20 mM Tris-HCl, pH 8·0/10mM EDTA/150 mM-NaCl (SET). The homogenate was centrifuged at 3000 rev/min for 5 min, and then the precipitate was resuspended in 8 ml of SET and mixed with 0·4 ml of 10% SDS. The mixture was twice extracted with phenol/chloroform and once with chloroform. The aqueous phase was extracted with ether and mixed with an equal volume of cold isopropanol. The DNA was spooled out with a glass rod and dissolved in 10 mM Tris-HCl, pH 8·0/1 mM-EDTA.

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press

Phage DNAs were prepared as described by Maniatis, Fritsch & Sambrook (1982).

(iii) *Southern blot analysis*

Southern blot analysis was carried out essentially according to Maniatis, Fritsch & Sambrook (1982) except that in some cases transfer to Zeta-Probe membrane (Bio-Rad) were performed using 1·5 M-NaCl/0·5 M-NaOH as transfer solvent (Reed & Mann, 1985). Probe DNAs were prepared by separating restricted fragments of recombinant phages λDmG6PD14 and λDmG6PD139 by agarose gel electrophoresis and labelled with [α-^{32}P]dCTP (ICN) using a Multiprime DNA labelling system (Amersham).

(iv) *Cloning and sequencing*

The G6PD coding region of strain 2512H was retrieved in cloned form using EMBL3 bacteriophage cloning vectors and *E. coli* NM 539 (*sup*F *hsd*R P2coxP3) (Frischauf *et al.* 1983). DNA fragments were subcloned using M13 or pUC and *E. coli* JM109 (Yanisch-Perron, Vieira & Messing, 1985), and sequenced as described by Sanger *et al.* (1980).

(v) *Enzyme assay*

G6PD activity was assayed by the method described in Hori *et al.* (1982). Protein was quantified by the method of Lowry *et al.* (1951).

3. Results

The restriction enzyme map of the recombinant phages containing the G6PD coding regions is shown in Fig. 1. These two phages are both from a genomic library of wild-type *Drosophila melanogaster* (Maniatis *et al.* 1978); λDmG6PD14 was incorrectly described as from strain 2512H in the previous report (Hori *et al.* 1985). The position of exon II is cited from Ganguly, Ganguly & Manning (1985), and that of exon I from Ito *et al.* (unpublished data).

In order to test whether the high-activity mutations have resulted from insertion of transposable elements in or near the G6PD coding region, we first performed Southern blot analyses of the genomic DNAs of mutants, 2512H, 1FH and S44H and their revertants, 2512L, 1FL and S44L, using probes shown in Fig. 1. The results clearly demonstrate that all the three high-activity mutants and the three revertants have a 4·1 Kb insertion sequence (Ins2) at a similar site within the intron (Fig. 2A), and that all except 2512L have an insertion sequence (Ins1) at a similar site near the 5' side of exon I, whose length varies from strain to strain (Fig. 2B). To study the structures of the Ins1 and Ins2 sequences in detail, Southern blot analyses of the genomic DNAs were further performed with different restriction enzymes and probes. The results are illustrated in Fig. 3.

It will be seen that the Ins1 of 2512H and S44H consists of direct terminal repeats containing ScaI, AvaI and Hin dIII sites, and a core sequence containing two ScaI sites, while one of the ScaI sites in

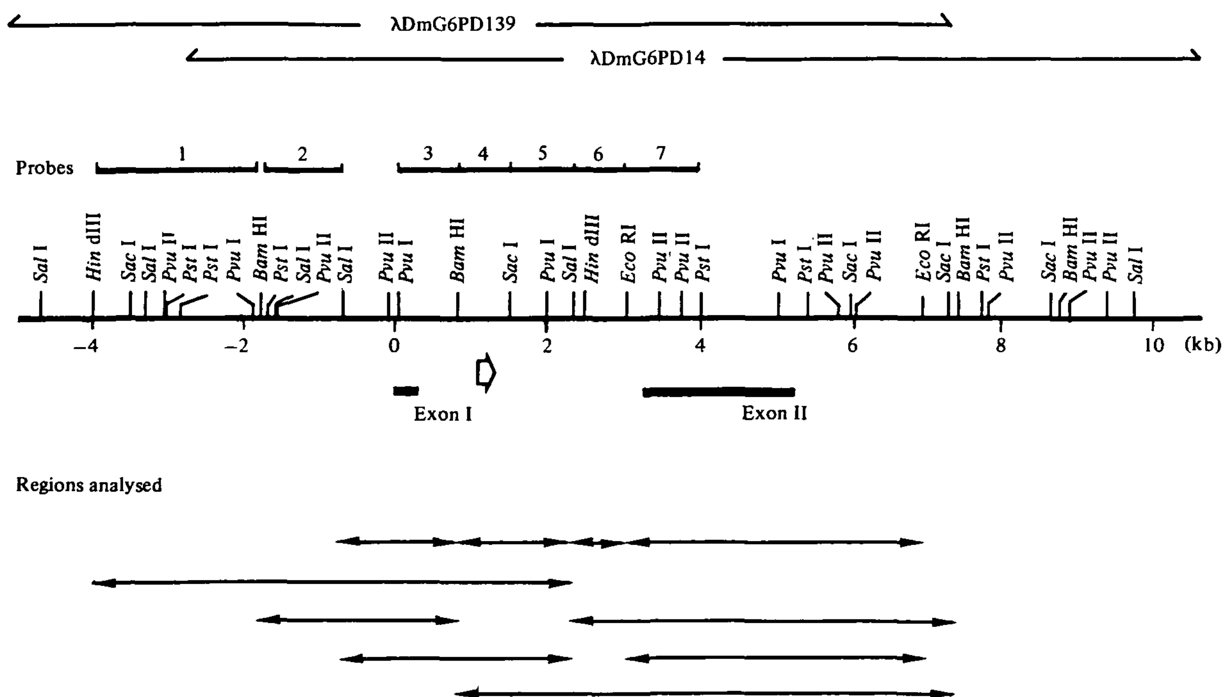
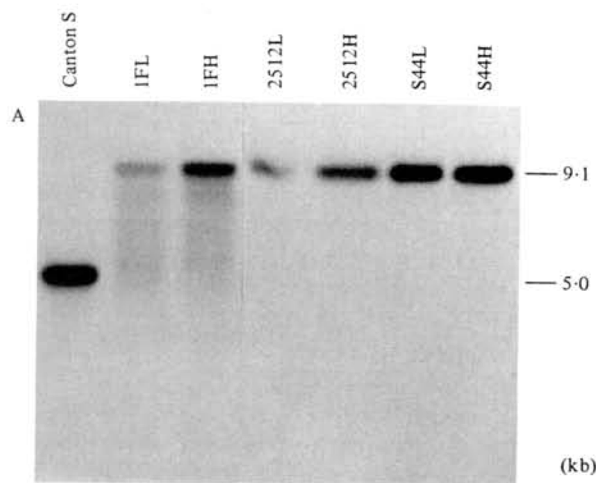
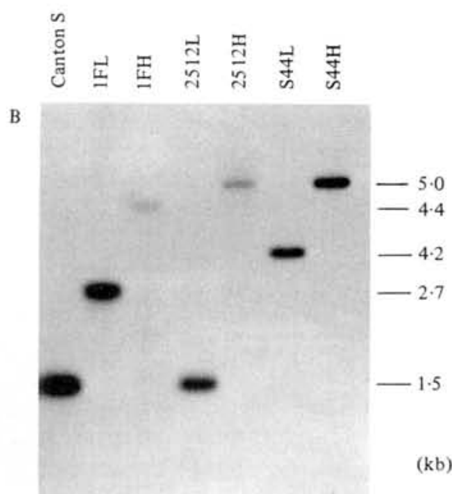


Fig. 1. Restriction map of the cloned G6PD gene (λDmG6PD139 and λDmG6PD14). Regions analysed by Southern blot are indicated. The positions of exon I (H. Ito *et al.*, unpublished data) and of exon II (Ganguly, Ganguly & Manning, 1985) are tentative. Open arrow below the map represents the direction of mRNA synthesis (Hori *et al.* 1985).

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press





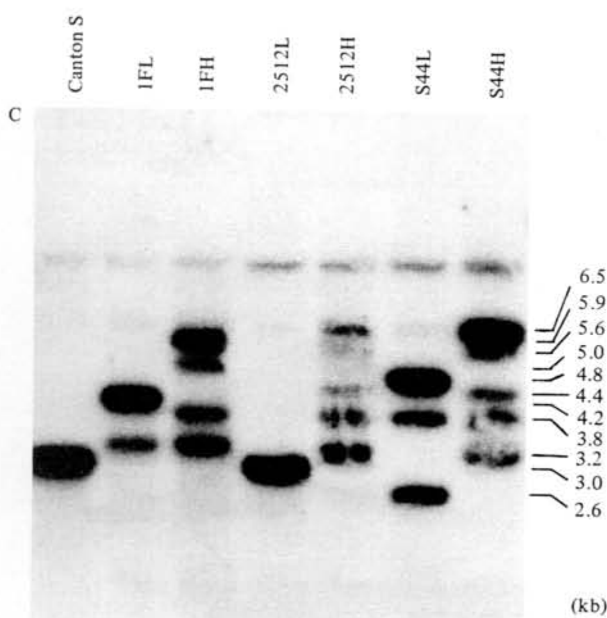


Table 1. *G6PD activities in 10-day-old male imagoes*

Strain	G6PD activity (mean ± S.E.)	No. of individuals
Wild-type		
Canton S	41·7 ± 1·5	12
High-activity mutants		
2512H	168·8 ± 4·5	18
S44H	156·7 ± 8·2	12
1FH	168·5 ± 1·2	84
Low-activity derivatives		
2512L	36·3 ± 1·1	12
S44L	75·9 ± 5·9	12
1FL	28·1 ± 1·1	18
Revertants		
rA31310	40·9 ± 3·2	12
rA3146	37·9 ± 2·8	16
rC155	46·9 ± 3·4	12
rC159	34·4 ± 2·0	12
rC1510	30·8 ± 3·0	12
rC165	79·6 ± 6·3	12
rC166	36·8 ± 2·7	16
rD1391	39·7 ± 2·7	12
rD13103	32·9 ± 2·5	12

G6PD activity was assayed with single head and thorax portions and is expressed as milliunits per milligram proteins.

the core sequence is missing from the Ins1 of 1FH. The ScaI sites were determined with partly digested samples (Fig. 2C). On the other hand, the right terminal repeat is missing in S44L, and the core sequence and one terminal repeat are both deleted in 1FL. 2512L has no Ins1. These findings suggest that the Ins1 in revertant strains may be deletion variants of the Ins1 in high-activity mutants. The structures of the Ins2 sequences were the same in all strains examined.

Table 1 shows the G6PD activities of sixteen strains used in the present study. A comparison of the G6PD activities and the structures of Ins1 sequences suggests that the Ins1 sequence, consisting of two direct repeats and a core sequence with at least one ScaI site, is involved in the high-G6PD activity mutations, and that deletion of one terminal repeat and the core sequence nullifies the stimulative effect of Ins1 on the

Fig. 2. Blot hybridization of genomic DNAs from high-activity mutants and their revertants: (A) *Bam* HI and *Sal* I digests probed with the *Eco* RI–*Pst* I 1·0 kb segment (probe 7 in Fig. 1). (B) *Bam* HI and *Sal* I digests probed with the *Pvu* I–*Bam* HI 0·9 kb segment (probe 3 in Fig. 1). (C) *Sal* I and *Sca* I digests probed with the *Sca* I–*Sal* I 0·9 kb segment (probe 5 in Fig. 1). These genomic DNAs were digested completely by *Sal* I, but partially by *Sca* I. Thus, the length of detected bands shows the leftward distance of *Sca* I sites from the *Sal* I site at the right end of probe 5. This allows us to map the position of internal *Sca* I restriction sites within the *Ins*1 sequences. Fragment sizes were determined by comparison to a *Hin* dIII digest of λ DNA run on the same gel.

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press

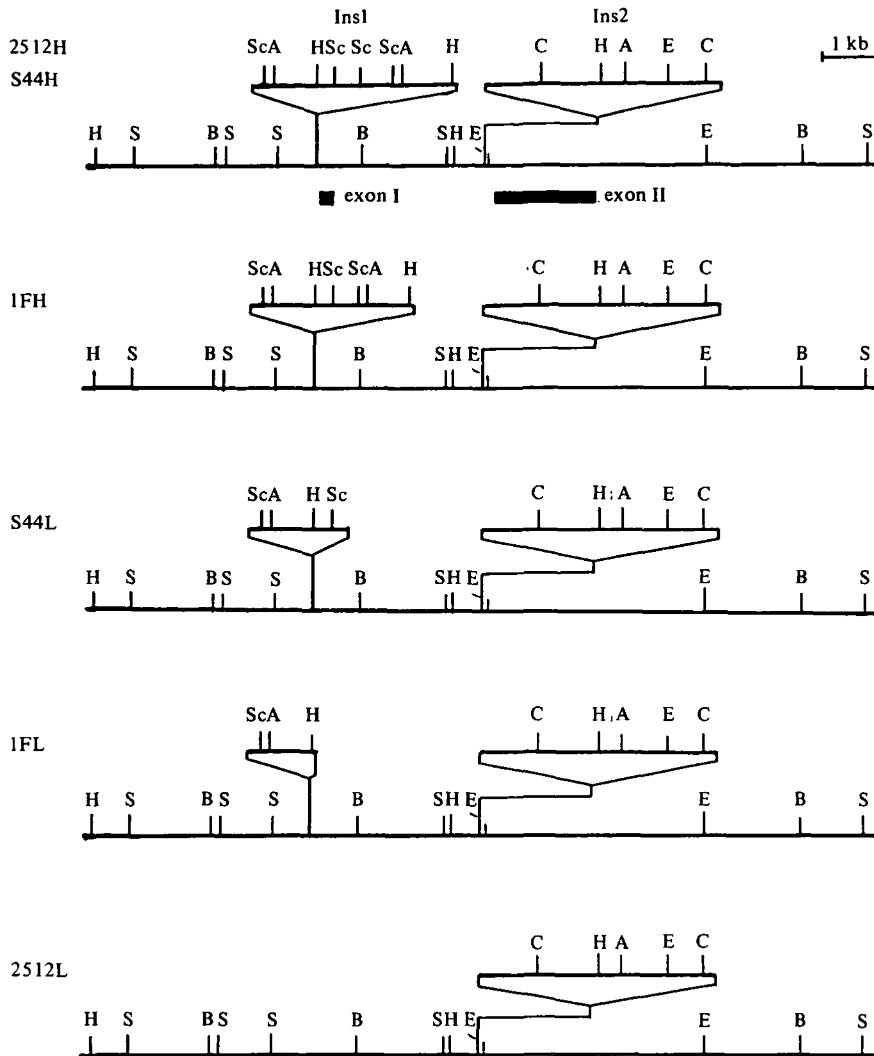


Fig. 3. Restriction maps around the G6PD coding region in high-G6PD activity mutants and their revertants. Abbreviations: *Ava*I, A, *Bam* HI, B, *Cla*I, C, *Eco* RI, E, *Hin* dIII, H, *Sal*I, S and *Sca*I, Sc.

G6PD gene expression as seen in 1FL, although part of the core sequence plus one terminal repeat appear to have a slightly stimulative effect as seen in S44L. On the other hand, the Ins2 sequence seems to have no effect on the G6PD gene activity.

The above notion of the relationship of the structure of the Ins1 sequence to the G6PD gene activity has been tested with nine revertants which had been obtained from crosses of 2512H with low-activity flies. We have previously reported that reversion to wild type could be induced by crossing wild-type females with the mutant males (Itoh & Hori, 1985); the reversion rates in the progeny from crosses Muller-5 × 2512H (M × Q), M–H × 2512H (P × Q) and 2152L × 2512H (Q × Q) were 2·5, 1·9 and 3·8%, respectively (2·7% on the average). Two revertants (rA lines) from the cross, Muller-5 × 2512H, five revertants (rC lines) from the cross, M-H × 2512H, and two revertants (rD lines) from the cross, 2512L × 2512H, were analysed by Southern blot (see Table 1 for G6PD activity of the revertants). For this purpose, the genomic DNAs were digested with several enzymes and probed with probes 1, 3 and 5, and the restriction maps shown in Fig. 4 were constructed. In support of our notion, one terminal repeat only in rA31310, rA3146, rC155, rC159, rC1510, rD1391 and rD13103 does not stimulate the G6PD gene expression at all as in 1FL, and one terminal repeat plus part of the core sequence in rC165 slightly stimulates the gene expression as in S44L. In addition, another important finding was obtained that the two terminal repeats have no effect unless accompanied by the core sequence as seen in rC166. It is thus probable that the core sequence is responsible for the stimulation of G6PD gene expression.

To confirm the results of Southern blot analyses, we next attempted to clone the Ins1 sequence from strain 2512H. A recombinant EMBL3 library was constructed by ligating Sau3AI-digested genomic DNA into the BamHI site of the vector. From 60000 recombinants screened with probe 3, one positive clone (λDmNG007) was obtained. Southern analysis of this clone showed that it carries two insertions; one was 2·3 kb long and located between the PvuII and

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press

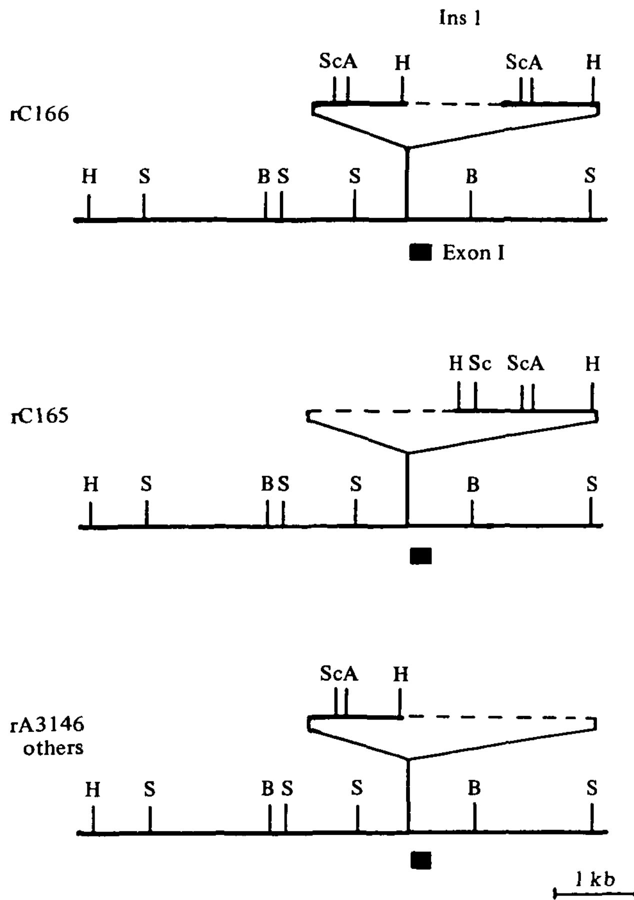


Fig. 4. Restriction maps of the Ins1 in the wild-type revertants from the mutant 2512H. Refer to the legend of Fig. 2 for abbreviations.

the *Pvu*I sites near the left end of exon I, and the other was 4·1 kb long and present between the *Hin* dIII and the *Eco* RI sites within the intron (data not shown). The insertion sites of these sequences were the same as in the genomic DNA of 2512H (cf. Figs. 3 and 5). However, the Ins1 was 1·2 kb shorter than that in the genomic DNA, and this 1·2 kb deletion matched the central core sequence of the genomic Ins1. Namely, the Ins1 sequence present in λDmNG007 consists of two terminal repeats only (Fig. 5). Another library of the genomic DNA of 2512H was also constructed with EMBL3, and 130000 recombinants were screened with probes 2 or 3. Of four positive clones obtained, three contained the Ins1 sequences 1·2 kb shorter than the genomic counterpart as in λDmNG007, and the remaining one had no Ins1, although all the four contained the 4·1 kb Ins2 in the same position. Other attempts to retrieve clones carrying a DNA segment spanning the entire Ins1 sequence have so far been unsuccessful. This suggests that the Ins1 sequence, particularly the core sequence, may be quite unstable and easily deleted from recombinant phages during propagation in *E. coli.*

The cloned Ins1 in phage λDmNG007 was sequenced. The results (Fig. 5) are summarized as follows: (1) the Ins1 consists of two defective P elements, each with the characteristic 31 bp inverted repeats as reported by O'Hare & Rubin (1983), (2) the sequence of the right P element perfectly matched that of the KP element reported by Black *et al.* (1987) (the internal deletion ranged from base 808 to base 2560 and the 32nd base, adenine, was replaced by thymine), (3) the left P element also had the same nucleotide sequence as the KP element except for the 32nd base, guanine, (4) the orientation of the P elements was opposite to the G6PD gene (the direction of the G6PD transcription was previously reported by Hori *et al.* 1985; it was recently confirmed by Ito *et al.* who determined the initiation site of G6PD transcription by S1 mapping and primer extension techniques, unpublished data), (5) the left and right P elements were flanked by duplicated target sequences of CTTGGGCG and CTGCGGCG, respectively, (6) there was a linking sequence, TTTGTTTG between the two defective P elements.

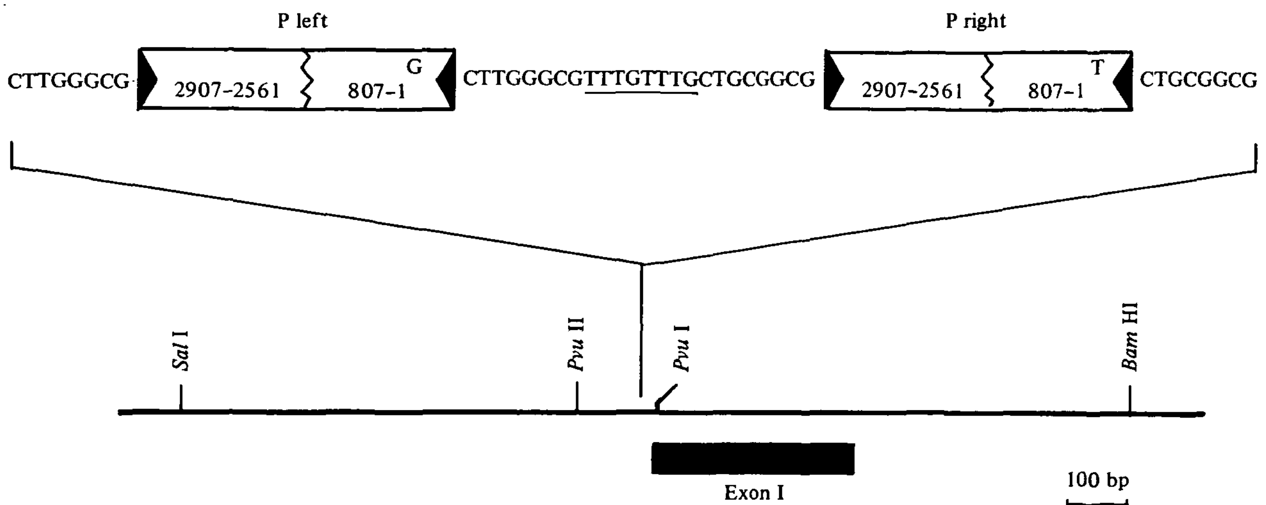


Fig. 5. Structure of Ins1 in a recombinant phage, λDmNG007. Ins1 in λDmNG007 consists of two defective P elements inserted upstream of exon I. Target duplications and 8 bp linking sequence between the two defective P elements are also shown. ▲, 31 bp inverted repeats. The 32nd base, A, in the complete P element was replaced by G and T in the left and right P element, respectively.

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press

4. Discussion

Southern blot and sequencing data clearly demonstrated that two sorts of insertion sequences are associated with the *X*-linked, high-G6PD activity mutations; one, Ins1, is 2·9 or 3·5 kb long and located upstream of exon I, and the other, Ins2, resides within an intron. However, the Ins2 sequences were found at similar positions in several low-G6PD activity revertants as in the high-activity mutants. It is therefore most probable that this sequence is not directly involved in stimulating the G6PD gene expression.

The structure of the Ins1 sequence varied in different strains, and its prototype appeared to be the one found in two mutations, 2512H and S44H. It consists of a core sequence flanked by two defective P elements as illustrated in Fig. 6. The core sequence in one of the three high-activity mutations (1FH), is 0·6 kb long and has one ScaI site, while that in the other two mutations (2512H and S44H) is about twice as long as in 1FH and has two ScaI sites. We therefore suggest that the core sequence in the latter two mutations consists of two identical units, and that 1FH has one such unit as the core sequence.

A number of P-element-mediated mutations have been reported in *D. melanogaster*, but they are all suppressive except w^c (Collins & Rubin, 1982; Levis, Collins & Rubin, 1982). In addition, there appears to be no precedent for a P element providing a new transcription initiation site or containing a promoter or enhancer sequence. Inference is therefore possible that the sequence responsible for stimulating G6PD gene expression in the high-activity mutations may be the core sequence found between the two defective P elements. This inference is supported by analyses of the genomic DNA structures of several revertants; one revertant, rC166, has two defective P elements, and 1FL and six other revertants have only one defective P element upstream of exon I (the sites of P insertions in these revertants seem to be the same in the three high-activity mutations) as shown in Figs. 3 and 4. In none of these cases were the G6PD activities higher than in the wild-type control, Canton S. Furthermore, a partly deleted unit of the core sequence found in S44L and rC165 seems to have a slightly stimulative effect on G6PD gene expression (Fig 6). This is also consistent with the above inference. The G6PD activity of strain 1FH is not statistically different from that of either strain 2512H or S44H. It appears therefore that one unit of the core sequence is as effective as the two units in stimulating gene expression.

It is important to know the exact structure of the core sequence in order to analyse the mechanism underlying its stimulation of gene expression. However, this sequence is highly unstable, and every attempt to retrieve recombinant phages or plasmids carrying the intact Ins1 sequence has so far been unsuccessful. Similar difficulties have been reported with other

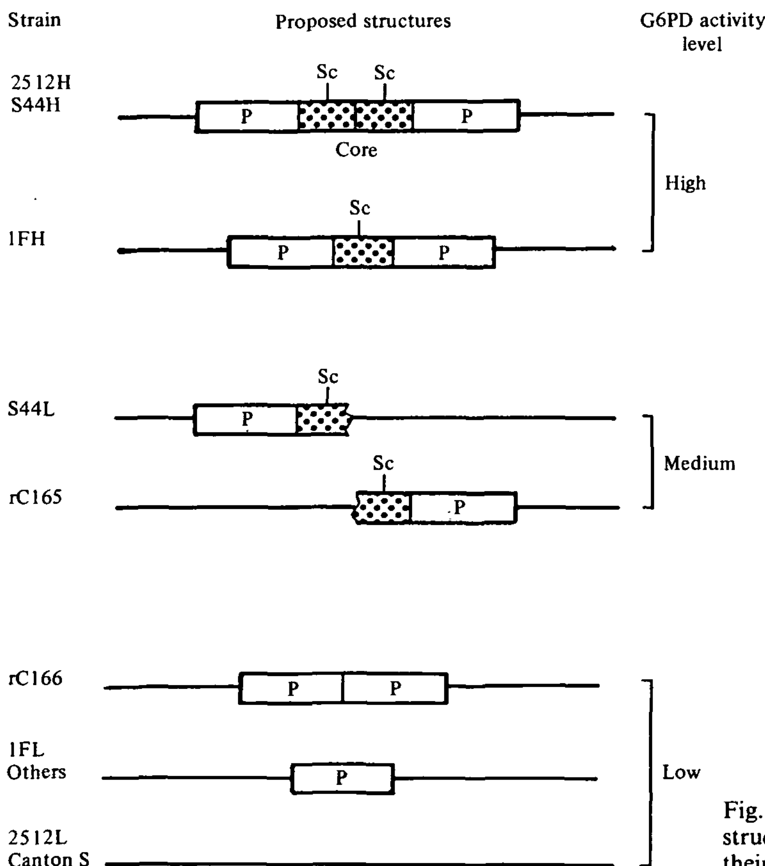


Fig. 6. Schematic representation of the proposed structures of the Ins1 sequences in various strains and their G6PD activity levels.

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press

transposable elements (Levis & Rubin, 1982; Collins & Rubin, 1982; Paro, Goldberg & Gehring, 1983; Bucheton *et al.* 1984).

The only available clue to the mechanism of positive regulation by the core sequence is the sequencing data on the site of P element insertion. A recombinant phage, λDmNG007 isolated from a genomic DNA library of strain 2512H carried two defective P elements arranged in tandem upstream of the G6PD coding region without the core sequence in between. The core sequence might in this case have been precisely excised from the Ins1 sequence during propagation of the phages in host bacteria (Bellett, Busse & Baldwin, 1971). In this phage, the left P element is flanked by target repeats of CTTGGGCG and the right by CTGCGGCG, and the two repeats are connected with an 8 bp sequence TTTGTTTG. We have recently sequenced a 5 kb region upstream of exon I using a recombinant phage retrieved from a Canton S library, and found that the TTTGTTTG sequence flanked by CTTGGGCG and CTGCGGCG is located about 40 bp upstream of the PvuI site (close to the 5' side of exon I; unpublished data). Although no canonical or non-canonical TATA sequences are found around this TTTGTTTG sequence, it is interesting that this TTTGTTTG sequence is followed by a GC-rich sequence about 20 bp long which has homology with a consensus sequence for a DNA region between the TATA box or its equivalent and the transcription initiation site described for several house keeping genes (Martini *et al.* 1986). Insertion of one or two P elements in such a region has no discernible effect on G6PD gene expression, as stated earlier, and the direction of transcription of the inserted P element is opposite to that of the G6PD gene. Taking these into account, one can speculate that the core sequence may carry a sequence having a strong promoter or enhancer activity or that insertion of the core sequence just 5' to exon I alters chromatin structures in such a way as to facilitate acceleration of transcription.

The Ins1 sequence described in this report seems to be functionally analogous to yeast transposons, Ty1 (Cameron, Loh & Davis, 1979) which preferentially integrate into promoter regions of various genes and cause constitutive expression of the affected genes (Roeder & Fink, 1983; Williamson, 1983). Ty1 has a central region of about 5·2 kb of DNA flanked by direct repeats of a δ sequence about 330 bp long, thus being structurally similar to retroviruses (Eibel *et al.* 1980; Fink, Boeke & Garfinkel, 1986). Each δ sequence has transcription start and stop signals, and a transcript initiating in the left δ could proceed out of the δ into the flanking DNA, although this would not necessarily explain the stimulative effects of Ty1 (Elder, Loh & Davis, 1983). Since the P element has no structural homology to the δ sequence of Ty1 elements (O'Hare & Rubin, 1983; Clare & Farabaugh, 1985; Hauber, Nelböck-Hochstetter & Feldmann, 1985; Warmington *et al.* 1985), the mechanism of stimulation of nearby genes might be different in the Ins1 sequence and the Ty1 element.

The Ins2 sequence was found in the same sites of the G6PD gene in the three unrelated, high-G6PD activity strains. That this may not be due to contamination was suggested by our recent survey of Japanese populations; the frequency of flies carrying the Ins2 sequence within the intron of the G6PD gene is significantly higher in Sapporo than in other places remote from Sapporo (unpublished data). Hamabata *et al.* recently sequenced the Ins2 sequence and found that it has no homology to any transposable elements so far reported (unpublished data).

A *Drosophila melanogaster* genomic DNA library (Maniatis) was obtained by courtesy of Dr J. A. Lengyel. We thank our colleagues for discussion and experimental assistance, particularly Dr M. T. Kimura, H. Ito, K. Yoshida, T. Ohtsu and T. Hamabata. This research was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan and by a grant from Takeda Science Foundation.

References

Bellett, A. J. D., Busse, H. G. & Baldwin, R. L. (1971). Tandem genetic duplications in a derivative of phage lambda. In *Bacteriophage Lambda* (ed. A. D. Hershey), pp. 501–513. New York: Cold Spring Harbor Laboratory.

Black, D. M., Jackson, M. S., Kidwell, M. G. & Dover, G. A. (1987). KP element repress P-induced hybrid dysgenesis in *Drosophila melanogaster*. *EMBO Journal* **6**, 4125–4135.

Bregliano, J. C. & Kidwell, M. G. (1983). Hybrid dysgenesis determinants. In *Mobile Genetic Elements* (ed. J. A. Shapiro), pp. 363–410. New York: Academic Press.

Bucheton, A., Paro, R., Sang, H. M., Pelisson, A. & Finnegan, D. J. (1984). The molecular basis of I-R hybrid dysgenesis in *Drosophila melanogaster*: identification, cloning, and properties of the I factor. *Cell* **38**, 153–163.

Cameron, J. R., Loh, E. Y. & Davis, R. W. (1979). Evidence for transposition of dispersed repetitive DNA families in yeast. *Cell* **16**, 739–751.

Clare, J. & Farabaugh, P. (1985). Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. *Proceedings of the National Academy of Sciences, USA* **82**, 2829–2833.

Collins, M. & Rubin, G. M. (1982). Structure of the Drosophila mutable allele, *white-crimson*, and its *white-ivory* and wild-type derivatives. *Cell* **30**, 71–79.

Eibel, H., Gafner, J., Stotz, A. & Philippsen, P. (1980). Characterization of the yeast mobile element Tyl. *Cold Spring Harbor Symposia on Quantitative Biology* **45**, 609–617.

Elder, R. T., Loh, E. Y. & Davis, R. W. (1983). RNA from the yeast transposable element Ty1 has both ends in the direct repeats, a structure similar to retrovirus RNA. *Proceedings of the National Academy of Sciences, USA* **80**, 2432–2436.

Engels, W. R. (1983). The P family of transposable elements in Drosophila. *Annual Review of Genetics* **17**, 315–344.

Fink, G. R., Boeke, J. D. & Garfinkel, D. J. (1986). The mechanism and consequences of retrotransposition. *Trends in Genetics* **2**, 118–123.

Frischauf, A. M., Lehrach, H., Poustka, A. & Murray, N.

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press

(1983). Lambda replacement vectors carrying polylinker sequences. *Journal of Molecular Biology* **170**, 827–842.

Ganguly, R., Ganguly, N. & Manning, J. E. (1985). Isolation and characterization of the glucose-6-phosphate dehydrogenase gene of *Drosophila melanogaster*. *Gene* **35** 91–101.

Hauber, J., Nelböck-Hochstetter, P. & Feldmann, H. (1985). Nucleotide sequence and characteristics of a Ty element from yeast. *Nucleic Acids Research* **13**, 2745–2758.

Hori, S. H., Akasaka, M., Ito, H., Hanaoka, T., Tanda, S., Ohtsuka, E., Miura, K., Takahashi, T. & Tang, J. J. N. (1985). Cloning of the glucose-6-phosphate dehydrogenase gene of *Drosophila melanogaster* using 17-base oligonucleotide mixtures as probes. *Japanese Journal of Genetics* **60**, 455–463.

Hori, S. H., Tanda, S., Fukazawa, K. & Hanaoka, T. (1982). Further studies on the modifier gene system regulating activities of *X*-linked enzymes in *Drosophila melanogaster*. *Japanese Journal of Genetics* **57**, 535–550.

Itoh, M. & Hori, S. H. (1985). An *X*-linked genetic factor that affects the activity of glucose-6-phosphate dehydrogenase (G6PD) in *Drosophila melanogaster*: effect of cytoplasm on its loss from the *X* chromosome. *Japanese Journal of Genetics* **60**, 441–453.

Iwabuchi, M., Hori, S. H. & Yorimoto, N. (1986). *X*-linked mutations that give rise to overproduction of glucose-6-phosphate dehydrogenase in *Drosophila melanogaster*. *Biochemical Genetics* **24** 319–327.

Kidwell, M. G. (1979). Hybrid dysgenesis in *Drosophila melanogaster*: the relationship between the *P-M* and *I-R* interaction systems. *Genetical Research* **33**, 205–217.

Kidwell, M. G., Kidwell, J. F. & Sved, J. A. (1977). Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility and male recombination. *Genetics* **86**, 813–833

Levis, R., Collins, M. & Rubin, G. M. (1982). FB elements are the common basis for the instability of the w^{DZL} and w^{c} Drosophila mutations. *Cell* **30**, 551–565.

Levis, R. & Rubin, G. M. (1982). The unstable w^{DZL} mutation of Drosophila is caused by a 13 kilobase insertion that is imprecisely excised in phenotypic revertants. *Cell* **30**, 543–550.

Lindsley, D. L. & Grell, E. H. (1968). *Genetic Variations of Drosophila melanogaster*. Carnegie Institution of Washington Publication no. 627.

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. *Journal of Biological Chemistry* **193**, 265–275.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning, a Laboratory Manual*. New York: Cold Spring Harbor Laboratory.

Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C. & Quon, D. (1978). The isolation of structural genes from libraries of eucaryotic DNA. *Cell* **15**, 687–701.

Martini, G., Toniolo, D., Vulliamy, T., Luzzatto, L., Dono, R., Viglietto, G., Paonessa, G., D'Urso, M. & Persico, M. G. (1986). Structural analysis of the *X*-linked gene encoding human glucose-6-phosphate dehydrogenase. *EMBO Journal* **5**, 1849–1855.

Merriam, J. R. (1968). *FM7*: first multiple seven. *Drosophila Information Service* **43**, 64.

Merriam, J. R. (1969). *FM7*: a 'new' first chromosome balancer. *Drosophila Information Service* **44**, 101.

O'Hare, K. (1985). The mechanism and control of P element transposition in *Drosophila melanogaster*. *Trends in Genetics* **1**, 250–254.

O'Hare, K. & Rubin, G. M. (1983). Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. *Cell* **34**, 25–35.

Paro, R., Goldberg, M. L. & Gehring, W. J. (1983). Molecular analysis of large transposable elements carrying the *white* locus of *Drosophila melanogaster*. *EMBO Journal* **2**, 853–860.

Reed, K. C. & Mann, D. A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Research* **13** 7207–7221.

Roeder, G. S. & Fink, G. R. (1983). Transposable elements in yeast. In *Mobile Genetic Elements* (ed. J. A. Shapiro), pp. 299–328. New York: Academic Press.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *Journal of Molecular Biology* **143**, 161–178.

Simmons, M. J. & Karess, R. E. (1985). Molecular and population biology of hybrid dysgenesis. *Drosophila Information Service* **61**, 2–7.

Spradling, A. C. & Rubin, G. M. (1981). Drosophila genome organization: conserved and dynamic aspects. *Annual Review of Genetics* **15**, 219–264.

Tanda, S. & Hori, S. H. (1983*a*). Regulation of glucose-6-phosphate dehydrogenase activity by a presumptive extrachromosomal factor in *Drosophila melanogaster*. *Japanese Journal of Genetics* **58**, 531–537.

Tanda, S. & Hori, S. H. (1983*b*). Modifier gene that affects glucose-6-phosphate dehydrogenase activity in *Drosophila melanogaster*. *Japanese Journal of Genetics* **58**, 591–606.

Warmington, J. R., Waring, R. B., Newlon, C. S., Indge, K. J. & Oliver, S. G. (1985). Nucleotide sequence characterization of Ty 1–17, a class I transposon from yeast. *Nucleic Acids Research* **13**, 6679–6693.

Williamson, V. M. (1983). Transposable elements in yeast. *International Review of Cytology* **83**, 1–25.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.

Young, M. W. (1979). Middle repetitive DNA: a fluid component of the *Drosophila* genome. *Proceedings of the National Academy of Sciences, USA*, **78**, 6274–6278.

https://doi.org/10.1017/S0016672300027622 Published online by Cambridge University Press