Variation at the 87A heat shock locus in *Drosophila melanogaster*

(restriction sites/DNA insertion/heterozygosity/transposable elements)

ANDREW J. LEIGH BROWN*

Imperial Cancer Research Fund, Mill Hill Laboratories, Burtonhole Lane, London NW7 1AD, United Kingdom

ABSTRACT   Restriction maps for 25 kilobases of DNA around the 87A7 heat shock locus have been determined in 29 chromosomes isolated from a natural population. The heterozygosity per nucleotide and the proportion of polymorphic nucleotide sites were estimated to be 0.0024 and 0.007, respectively. The mean number of insertional differences in this region between random pairs of chromosomes was 0.95. A significant amount of this variation was due to the insertion of large transposable elements. All the insertion/deletion events were found in a region less than 2 kilobases in size. This could either be due to nonrandom integration or to differences in the intensity of selection against DNA insertion at different sites.

Several estimates of the level of variation in mitochondrial DNA from natural populations have now been made (1–5). This work has generally involved mapping restriction site polymorphisms and various approaches to the analysis of such data have now been published (6–10). The study of DNA-level variation in nuclear genes is technically more complex and to date three systematic population surveys have been completed—one on the β-globin region in man (11) and two on the alcohol dehydrogenase (*Adh*) locus in *Drosophila melanogaster* (12, 13). Much information has been collected at other loci in man (14–18) but, in general, those studies were conditioned by the prior discovery of variation and do not provide an unbiased estimate of its overall level. In this paper I present the results of a survey on variation in restriction map at the 87A7 heat shock locus in a single natural population of *D. melanogaster*.

The major, 70,000-dalton, heat-inducible protein (hsp70) in *D. melanogaster* is encoded at two cytogenetic loci, 87A7 and 87C1 (19–22). The coding sequence for the hsp70 mRNA is 2.1 kilobases (kb) long, contains no introns, and is duplicated at both loci. The duplicated sequences ("z element") include 400 bases 5' to the transcribed region which is conserved in all copies (20, 23–25). At 87A7, two z elements are arranged in a diverging orientation separated by a 1.2-kb spacer region (26, 27). The same arrangement is found in the two closely related species *D. simulans* and *D. mauritiana* (28). Comparison of the restriction maps of these species provides estimates of sequence divergence of 3% between *D. melanogaster* and *D. simulans*, 4% between *D. melanogaster* and *D. mauritiana*, and 1.3% between *D. simulans* and *D. mauritiana*. Although no restriction map is available, it was also shown that the hsp70 gene arrangement in the more distantly related species *D. teisseri* and *D. yakuba* is similar to that found in *D. simulans*, *D. mauritiana*, and at 87A7 in *D. melanogaster* (28).

In this paper I present data that indicate that the conservation of gene arrangement observed in this species subgroup has been achieved in spite of high levels of intraspecific insertion/deletion variation. Although such events may occur

outside the regions required for correct expression of heat shock sequences (29, 30), they nevertheless appear to affect fitness sufficiently for them to be screened out by natural selection.

MATERIALS AND METHODS

Twenty-nine *D. melanogaster* third chromosomes were obtained from a large collection made in June 1977 from a natural population in Raleigh, N.C. Since that time the chromosomes have been maintained over *In (3LR)TM6 Ubx*. Methods used for setting up these lines have been described elsewhere (31). Before extracting DNA, the 87A region of each wild-derived chromosome was uncovered by crossing the balanced stock to a deficiency stock *Df(3R)kar*^{D3} and selecting +/deficiency progeny from which to make DNA. This deficiency extends from 86E16,18 to 87F3,4 and includes approximately 30 lethal complementation groups (32). One wild line gave no progeny of the desired class. It is likely that the deficiency uncovered a lethal mutation on the wild chromosome.

DNA purification, restriction digests, and electrophoresis were performed as described in ref. 22. DNA fragments were visualized by the method of Southern (33) as adapted by Wahl *et al.* (34). Plasmids pPW232.1 and pPW229.1 (14) were used as probes for the 5' and 3' halves of the z element and plasmids 56H8/C (28) and 87A/5 (a gift from A. Udvardy) were used to detect fragments from distal and proximal flanking regions, respectively.

In 4 of the 29 lines Southern transfer hybridization indicated major structural changes at the 87A heat shock locus. To elucidate these changes the 87A hsp70 region was cloned from the four lines in a total of three different vectors. The method of Ish-Horowicz and Burke (35) was used to make a cosmid library from line NC128 from which a recombinant clone, containing 35 kb of *Drosophila* DNA around the 87A heat shock locus, was purified by using plasmid 56H8/C as a hybridization probe. Bacteriophage λ vectors were used to clone the other three strains: λL47 (36) was used to clone a 13-kb *Eco*RI fragment containing the 87A hsp70 genes from strain NC121. Genomic DNA from this strain was digested to completion and ligated to *Eco*RI-cut λL47 DNA by using conditions described by Ish-Horowicz and Burke (35). The ligated mix was packaged *in vitro* following Scalenghe *et al.* (37) and screened as above, omitting the chloramphenicol amplification step. The 87A hsp70 loci from lines NC903 and NC960 were cloned in λEMBL4, a derivative of 1059 (38) developed by H. Lehrach, G. Cesaeris, and N. Murray. Genomic DNA from these strains was partially digested with *Mbo* I under conditions that resulted in a high proportion of fragments of about 20 kb in size, estimated by electrophoresis against *Eco*RI-digested λ DNA. The DNA was then

Abbreviations: hsp70, 70,000-dalton heat-inducible protein; kb, kilobase(s).
* Present address: School of Biological Sciences, Univ. of Sussex, Falmer, Brighton BN1 9QG, U.K.

ligated to *Bam*HI-digested EMBL4 DNA and packaged as described above. Packaged phage were plated on *Escherichia coli* strain WL95, a P2 lysogen on which only recombinants will give plaques, in the first screen. Replating picked plaques was carried out on the isogeneic nonlysogenic strain WL87. After purification of the recombinant phage, DNA minipreps were prepared as described by Davis *et al.* (39). To test for homology with other elements DNA "dot blots" were performed as follows: a solution of plasmid DNA at 5 μg/ml containing a cloned *Drosophila* transposable element was denatured by being made 0.2 M in NaOH and neutralized with 1 M Tris·HCl (pH 7.5). One microliter was spotted onto a nitrocellulose filter, baked at 80°C, and hybridized to nick-translated DNA from the genomic λ clones prepared as above. Five nanograms of pBR322 DNA and approximately 0.5 ng of the recombinant λ clones themselves were included as negative and positive controls, respectively.

Restriction enzymes were obtained from New England BioLabs and Bethesda Research Laboratories; nitrocellulose filters, from Millipore; [^{32}P]dCTP (2,000–3,000 Ci/mmol; 1 Ci = 3.7 × 10^{10} Bq) from Amersham International; and dextran sulfate, from Pharmacia.

Survey Method. Genomic DNA (0.2–0.4 μg) was digested with one of the following enzymes: *Bam*HI, *Bgl* I, *Bgl* II, *Pst* I, *Xba* I, and *Xho* I. Genomic digests were electrophoresed in parallel, transferred to nitrocellulose, and probed initially with plasmid pPW232.1. After exposure, the filter was allowed to decay and rehybridized to pPW229.1, 56H8/C, and 87A/5 in turn according to the digest. DNA could be detected up to 18 months from the date of transfer. For immediate rehybridization filters were immersed in 0.2 M NaOH for 20 min and neutralized in 0.5 M Tris·HCl (pH 7.5) to remove old signal.

A restriction map for each line was constructed showing all sites cut by the above enzymes. Variation in restriction fragment length could thus be assigned either to variable sites or to insertion/deletion events. All of the enzymes used recognize a hexanucleotide sequence.

Outside of the probed regions it is necessary to have flanking sites for a second enzyme if restriction site variation is to be distinguished from insertion/deletion events. Both variable sites discovered in the course of this survey were outside of the areas probed. In the case of the proximal *Xho* I site, an *Xba* I site 5.5 kb further from the coding sequences was used. Although the resolution of this band was not sufficient to use it for scoring *Xba* I site variation, it was sufficient to establish that the 2-kb variation in *Xho* I fragment length was not due to DNA insertion. For similar reasons, a distal *Bgl* II fragment whose length was either 10.5 or 11.5 kb in different strains was not included in the analyses because of the lack of flanking sites. In addition, two *Pst* I sites, each located very close to the central *Bam*HI site in the z elements (a total of four sites), were omitted from the

analysis. Variation in each of these could not be reliably scored at the resolution used in this survey.

RESULTS

A restriction map of the region around the 87A hsp70 genes showing all restriction sites scored in this survey and the sites of polymorphic insertion/deletions is given in Fig. 1. The polymorphic restriction sites discovered in this survey and the extent of the DNA probes used to identify the restriction fragments are also indicated. Table 1 lists the nine configurations of variants found.

Two variable restriction sites were discovered out of 25 screened in 29 lines. Both appear to be located in noncoding DNA. The proportion of polymorphic nucleotide sites is estimated by two methods to be 0.007 (8, 10) with a standard error of 0.005 (10). The heterozygosity per nucleotide is estimated to be 0.0024 (7, 9).

Insertion/deletion events ("insertional variation") contribute a substantial proportion of the variation discovered in the survey. The heterozygosity due to such events can be expressed as the mean number of insertional differences between two randomly picked chromosomes over the region studied. This is estimated from $H = 2\Sigma n_i n_j v_{ij}/n(n-1)$, in which v_{ij} = number of insertional differences between the ith and jth chromosome types and n_i, n_j ($n_i \neq n_j$) are the number of times each type occurs ($n = \Sigma n_i$) in the sample. From the data in Table 1, H (insertion/deletions) = 0.95.

Two size classes of DNA insertion/deletions were found. Four lines contained insertions greater than 0.5 kb in size. The region of DNA containing the insertion was purified from each of these by molecular cloning. Fig. 2 gives brief restriction maps of the clones obtained. In the case of line NC121 the exact site and identity of the insertion has been established by DNA sequence analysis (unpublished data). The first 30 nucleotides from the 5' end of the inserted DNA are identical to the terminal nucleotides of a P element (K. O'Hare and G. M. Rubin, personal communication).

The DNA sequences inserted into the 87A7 loci in lines NC903 and NC960 were tested for homology with seven *Drosophila* transposable elements, by dot blot hybridizations. Strong homology was detected between clone λ903b and mdg-4 DNA. This is a copia-like element 7.5 kb long. It was originally isolated as a partial clone, Dm11, by virtue of its homology to double-stranded RNA (ref. 40; G. P. Georgiev, personal communication). Dm11 hybridizes to five to seven sites in Oregon-R polytene chromosomes (40, 41). Note that the restriction map of the complete mdg-4 element for the enzymes *Eco*RI, *Hin*dIII, *Xho* I, and *Bgl* II is identical to that of gypsy (42), an element responsible for several visible mutants suppressible by suppressor of Hairy-wing [*su(Hw)*]. No homology was detected be-
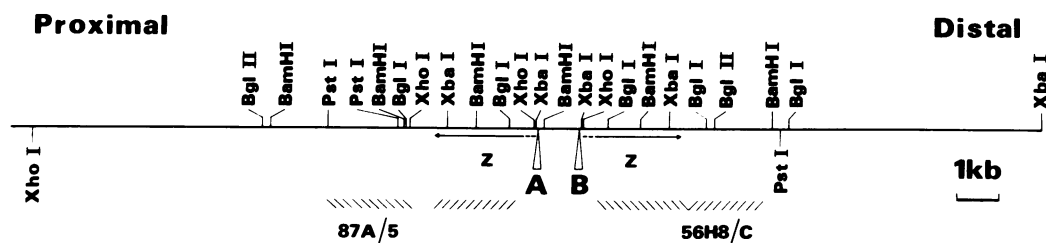

FIG. 1. Restriction map variation at the 87A7 heat shock locus. Twenty-five restriction sites were screened for variation in 29 lines of *D. melanogaster* from a wild population. Monomorphic sites are shown above the line; polymorphic ones are shown below. Hsp70 coding elements (z) and their direction of transcription are indicated by the arrows. The shaded areas depict the extent of homology of the probes used. Plasmids pPW232.1 and pPW229.1 were used to probe the coding region. The sites of two small polymorphic insertion/deletion events are also shown. See text for methods and explanation.

Table 1. Restriction map variants at the 87A7 loci from 29 lines of *D. melanogaster*

		Spacer region				
Configuration	*Xho* I proximal	Insertion A	Insertion B	*Pst* I distal	Lines	Total
1	+	−	+	+	171, 341, 423, 984, 461, 520, 620, 420, 439, 336, 946, 156	12
2	+	+	+	+	420, 206, 77, 478, 224	5
3	+	−	+	−	120, 901	2
4	+	−	−	−	69, 71, 137	3
5	−	−	−	−	969, 912, 981,	3
5a	−	−	−	−	121	1*
6	+	−	+	+	128	1†
7	+	−	?	+	903	1‡
8	+	−	?	−	960	1§

See text and Figs. 1 and 2 for explanation. Question marks indicate that some small insertions could not be scored because of the presence of large ones. This has a negligible effect on the estimates of heterozygosity.
* 0.8-kb insertion to 5′ end of distal z element.
† 5.5-kb insertion to 5′ of proximal z element.
‡ 7-kb insertion to distal *Bam*HI fragment in spacer region.
§ 1.8-kb insertion to distal *Bam*HI fragment in spacer.

tween λ960 and mdg-4 or any of the following elements: *copia, 412, 297* (43), P (44), mdg-1, and mdg-3 (45, 46).

The second class of insertional variation consists of two events of between 100 and 200 base pairs in size. This class is highly polymorphic. Within the limits of resolution of the techniques used, the same variants at each site occur in at least five lines.
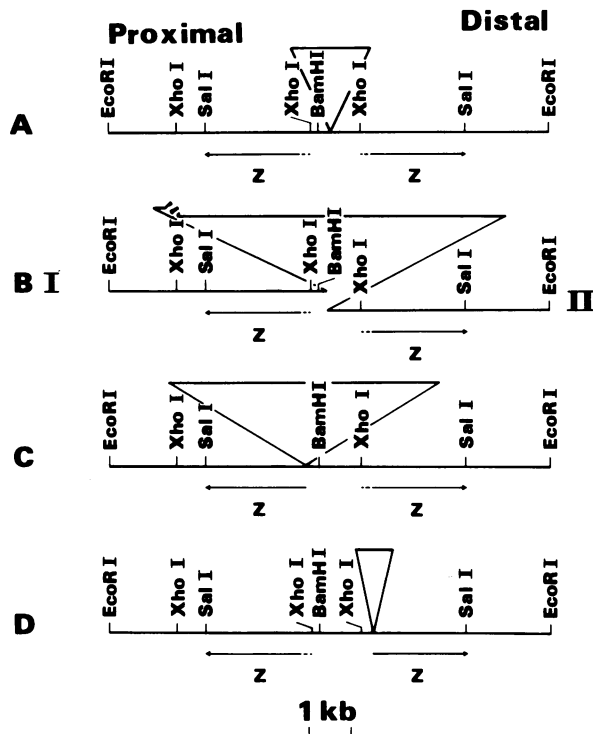

FIG. 2. Four major insertions at the 87A7 heat shock locus. Schematic restriction maps of the 87A7 regions from four lines containing large DNA insertions are shown. The organization of the 87A7 loci in these lines was determined by molecular cloning as described in the text. (*A*) λ960a cloned from NC960 in λEMBL4; (*B*) λ903a (I) and λ903b (II) cloned from NC903 in λEMBL4; (*C*) cosmid pMH128 cloned from NC128 in pJB8; (*D*) λ121 cloned from NC121 in λL47. Nick-translation and hybridization of these clones to restricted genomic DNA from various strains of *D. melanogaster* gives complex, strain-specific patterns, which indicate that the inserted elements are transposable (unpublished data).

The heterozygosity due to these insertions is estimated to be 0.67. A total of 13 kb of DNA, of which 9 kb was noncoding, was screened in sufficient detail to detect events of this size. The heterozygosity per kilobase of noncoding DNA is thus 0.07, but this figure does not reflect the distribution of the insertion events along the DNA.

Figs. 1 and 2 show that the distribution of the sites of the insertional variants is highly nonrandom. All four large insertions have occurred within 1.5 kb and none has been found elsewhere in 25 kb screened. Similarly, both small insertional variants occur within 1.2 kb of each other and in the same region as the large insertions. The DNA involved in insertion B is known not to be repetitive (ref. 47; see *Discussion*) but is found at one other region in the genome. Insertion A has not been characterized.

Strong associations were found between variants at the two polymorphic restriction sites and the two small insertion/deletion polymorphisms. Insertion A in the spacer region occurred only with the commonest configuration of variants at the other polymorphic sites. Similarly, the absence of the proximal *Xho* I site occurred in only one configuration at the other three sites. There was also a strong association between variants at the distal *Pst* I site and insertion B in the spacer region. This association was highly significant (Fisher's exact test). The value of r, the product moment correlation, was 0.836.

DISCUSSION

Estimates of genetic variability have been obtained for two other regions of nuclear DNA. From a survey of restriction site variation around the alcohol dehydrogenase locus of *D. melanogaster*, Langley *et al.* (12) estimate the heterozygosity per nucleotide to be 0.006 and the proportion of polymorphic nucleotide sites to be 0.015. DNA sequence analysis has been used to estimate variability in the alcohol dehydrogenase coding region (13). The proportion of polymorphic nucleotide sites was 0.02, in 1.7 kb of DNA sequenced in each of 11 lines. Estimating heterozygosity per silent nucleotide site from those data, a mean value of 0.011 is obtained, which is higher than that obtained by restriction mapping of the surrounding region.

The second region of nuclear DNA for which comparable data are available is the human β-globin region (11). Heterozygosity per nucleotide has been estimated to be 0.001, with 0.5% of the nucleotide sites polymorphic (8). The estimate of heterozy-

Genetics: Leigh Brown

Proc. Natl. Acad. Sci. USA 80 (1983) 5353

gosity per nucleotide presented here for the 87A7 heat shock region in *D. melanogaster*, 0.0024, thus falls between the earlier estimates in *Drosophila* and in man. Comparing these figures with similar estimates for mitochondrial DNA "heterozygosity," we find that in man nuclear DNA is considerably less variable (8, 9) but in *Drosophila* the two appeared to be similar (7, 12). Although the current estimate for the 87A7 region would draw a closer parallel between the two organisms in this respect, it may be that differences between the estimates reflect differences in constraints imposed by selection, or even in mutation rates, between different regions of nuclear DNA.

A prominent feature of this survey is the high level of variation due to insertion/deletion events. Two classes of event were found, distinguished by size. Two smaller insertion/deletions of 100–200 base pairs were highly polymorphic. Several examples of both were found among the 29 lines (Table 1). The four large insertions, 0.7–7 kb in length, were each found only once in this sample.

One of the two small insertional variants reported here has been previously described. Ish-Horowicz and Pinchin (25) reported a strain difference in the size of the spacer region at the 87A7 hsp70 locus. This difference corresponds to insertion B in Fig. 1 (unpublished data). In a detailed study of the hsp70 genes from 87A7 and 87C1 using DNA sequence analysis, Mason *et al.* (48) have precisely identified what appears to be the same variant. It involves the insertion or deletion of 140 nucleotides some 60 bases upstream of the 5′ *Xba* I site in the distal z element. They report that the DNA involved is virtually identical to an element in a similar position upstream of the proximal hsp70 gene at 87C1. There is evidence that this element is not repeated elsewhere in the genome (47). An adjacent element, some 120 base pairs in length, has been described as being repetitious (48), but this was apparently not involved in the variation found among these strains. Mason *et al.* (48) also suggest that stretches of "simple sequence" DNA in the spacer region might be involved in recombination or similar processes. However, the insertional variation detected in this survey did not occur at these sites.

The associations that were found between variants in this survey are expected under standard population genetic theory. The association between a newly arising neutral variant and flanking markers will persist for a number of generations inversely proportional to the recombination distance which separates them. As the distances between the variable sites are all very small in genetic terms, nonrandom associations are to be expected. There are several possible interpretations of the origins and relationships of the haplotypes observed.

It is possible that insertional variants could have been generated while the balanced lines used in this survey were being constructed. Crosses between wild males and females of many laboratory strains of *D. melanogaster*, such as those used to establish the lines studied here, frequently give rise to a syndrome of germ line aberrations known as hybrid dysgenesis (49, 50). There is now good evidence that in many cases the defects arise as a consequence of the induction of a high rate of transposition of a mobile element—the P factor (44). One of the insertional variants described here, the 0.7-kb insertion into the 5′ end of the transcribed region of the distal z element in line NC121, has already been identified as a P element. It is possible that it arose in this position during the isolation procedure.

In the case of the two small insertional variants it is unlikely that they arose through dysgenesis during chromosomal isolation. First, the sample includes what appear to be several examples of each form at these two sites. Second, from the discovery of very similar variants in different laboratory strains

with no known connection with North Carolina (see above), it is likely that at least one of the polymorphisms is widespread in different populations. Third, if the five examples of insertion A were all derived from a single event in the population, it would be expected that they should be found in the commonest haplotype with respect to the other variable sites as is observed (Table 1). However, if they arose independently during isolation, this distribution becomes one-fifth as probable.

The situation regarding the other three large insertions is less clear. Some mutants induced in a dysgenic cross have been shown to be due to the insertion of *copia* (51), but there is no evidence yet that this extends to other elements. The induction of three insertions in a sample of 29 lines would require an extremely high rate of transposition. It is at least equally plausible that some of these were present in the population and that a significant amount of variation is due to the insertion of transposable elements. This contention is supported by the presence of a transposable element in the spacer region of a variant 87A7 locus, which was cloned from a polymorphic laboratory Oregon-R strain (23, 52). However, the element in question does not seem to be related to those isolated in this study on the basis of size and restriction sites (unpublished data).

The most striking characteristic of the data presented here on insertional variation in the 87A7 region is the marked clustering of the insertion sites. All were found in the spacer region between the two genes within some 1.5 kb of DNA. This represents about 12% of the noncoding DNA surveyed in detail and 7% of the total length of DNA studied. This nonrandom distribution could either have arisen through nonrandom insertion or be the consequence of natural selection. Insertion of transposons Tn3 and Tn10 in prokaryotes has been shown to be highly nonrandom (53, 54), although Tn5 appears to insert at random (55). The clustering of insertions observed in the 87A7 region involves elements that differ considerably in structure. Although the P element, with its 31 base-pair terminal inverted repeats, resembles a prokaryotic transposon (56), mdg-4 (gypsy) has terminal direct repeats 500 base-pairs long and thus more closely resembles a vertebrate retrovirus (G. P. Georgiev, personal communication). It would be surprising if two so different elements often selected the same target for integration.

There is an alternative explanation for the distribution of insertions. Despite the high level of intraspecific variation due to DNA insertion, the 87A7 regions from the closely related species *D. simulans* and *D. mauritiana* show no major disturbance in restriction map from that found in *D. melanogaster* over 20 kb (ref. 28; also unpublished data). Together with similar results in the alcohol dehydrogenase locus in *Drosophila* (12), this suggests that selection can detect the effects of DNA insertion at some distance from coding elements and prevent such variants from becoming fixed in the species. In the β-globin gene cluster of higher primates it has recently been suggested that even the rate of nucleotide substitution may be lower outside the coding regions than that for silent substitutions within them (57). If DNA insertions in noncoding DNA have a significant effect on fitness, it is clearly possible that this effect may vary in severity according to the site. Therefore, it is possible that insertions outside of the spacer in the 87A7 region might have a more severe effect and be screened out more rapidly by natural selection. It is worth noting that the spacer region of the 87C1 locus in *D. melanogaster* is the only place where a major structural difference was found to have occurred during the evolution of the hsp70 coding loci in the *D. melanogaster* species group (28). Certainly only 350 bases 5′ of the transcript are required for correct expression *in vivo* (29). However, the chromosomal puffing response involves a large region around the locus and there might be signals in flanking DNA—as pos-

tulated for the β-globin gene cluster (58, 59)—whose spacing from the coding region is important. Though the scope for speculation is currently considerable, it is reasonable to expect that the acquisition of more data from different loci and different populations would help in discriminating between some of the possibilities.

1. Avise, J. C., Lansman, R. A. & Shade, R. O. (1979) *Genetics* 92, 279–295.
2. Avise, J. C., Giblin-Davidson, C., Laerm, J., Patton, J. C. & Lansman, R. A. (1979) *Proc. Natl. Acad. Sci. USA* 76, 6694–6698.
3. Shah, D. M. & Langley, C. H. (1979) *Nature (London)* 281, 696–699.
4. Brown, W. M. (1980) *Proc. Natl. Acad. Sci. USA* 77, 3605–3609.
5. Brown, G. G. & Simpson, M. V. (1981) *Genetics* 97, 125–143.
6. Kaplan, N. & Langley, C. H. (1979) *J. Mol. Evol.* 13, 295–304.
7. Nei, M. & Tajima, F. (1981) *Genetics* 97, 145–163.
8. Ewens, W. J., Spielman, R. S. & Harris, H. (1981) *Proc. Natl. Acad. Sci. USA* 78, 3748–3750.
9. Engels, W. R. (1981) *Proc. Natl. Acad. Sci. USA* 78, 6329–6333.
10. Hudson, R. R. (1982) *Genetics* 100, 711–719.
11. Jeffreys, A. J. (1979) *Cell* 18, 1–10.
12. Langley, C. H., Montgomery, E. & Quattlebaum, W. F. (1982) *Proc. Natl. Acad. Sci. USA* 79, 5631–5635.
13. Kreitman, M. (1983) *Nature (London)*, in press.
14. Bell, G. I., Karam, J. H. & Rutter, W. J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 5759–5763.
15. Rotwein, P., Chyn, R., Chigwin, J., Cordell, B., Goodman, H. & Permutt, M. A. (1981) *Science* 213, 1116–1120.
16. Wyman, A. R. & White, R. (1980) *Proc. Natl. Acad. Sci. USA* 77, 6754–6758.
17. Higgs, D. R., Goodbourn, S. E. Y., Wainscoat, J. S., Clegg, J. B. & Weatherall, D. J. (1981) *Nucleic Acids Res.* 9, 4213–4224.
18. Antonarakis, S. E., Orkin, S. H., Kazazian, H. H., Jr., Goff, S. C., Boehm, C. D., Waber, P. G., Sexton, J. P., Ostrer, H., Fairbanks, V. F. & Chakravarti, A. (1982) *Proc. Natl. Acad. Sci. USA* 79, 6608–6611.
19. Livak, K., Freund, R., Schweber, M., Wensink, P. C. & Meselson, M. (1978) *Proc. Natl. Acad. Sci. USA* 75, 5613–5617.
20. Artavanis-Tsakonas, S., Schedl, P., Mirault, M. E., Moran, L. & Lis, J. (1979) *Cell* 17, 9–18.
21. Craig, E. A., McCarthy, B. J. & Wadsworth, S. C. (1979) *Cell* 16, 575–588.
22. Ish-Horowicz, D., Pinchin, S., Gausz, J., Gyurkovics, H., Bencze, G., Goldschmidt-Clermont, M. & Holden, J. J. (1979) *Cell* 17, 565–571.
23. Ish-Horowicz, D., Pinchin, S. M., Schedl, P., Artavanis-Tsakonas, S. & Mirault, M.-E. (1979) *Cell* 18, 1351–1358.
24. Mirault, M.-E., Goldschmidt-Clermont, M., Artavanis-Tsakonas, S. & Schedl, P. (1979) *Proc. Natl. Acad. Sci. USA* 76, 5254–5258.
25. Moran, L., Mirault, M.-E., Tissieres, A., Lis, J., Schedl, P., Artavanis-Tsakonas, S. & Gehring, W. J. (1979) *Cell* 17, 1–18.
26. Ish-Horowicz, D. & Pinchin, S. M. (1980) *J. Mol. Biol.* 142, 231–245.
27. Goldschmidt-Clermont, M. (1980) *Nucleic Acids Res.* 8, 235–252.
28. Leigh Brown, A. J. & Ish-Horowicz, D. (1981) *Nature (London)* 290, 677–682.
29. Udvardy, A., Sumegi, J., Toth, E. C., Gausz, J., Gyurkovics, H., Schedl, P. & Ish-Horowicz, D. (1982) *J. Mol. Biol.* 155, 267–280.
30. Pelham, H. R. B. (1982) *Cell* 30, 517–528.
31. Voelker, R. A., Langley, C. H., Leigh Brown, A. J., Ohnishi, S., Dickson, B., Montgomery, E. & Smith, S. C. (1980) *Proc. Natl. Acad. Sci. USA* 77, 1091–1095.
32. Gausz, J., Gyurkovics, H., Bencze, G., Awad, A. M., Holden, J. J. & Ish-Horowicz, D. (1981) *Genetics* 98, 775–789.
33. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503–517.
34. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* 76, 3683–3687.
35. Ish-Horowicz, D. & Burke, J. F. (1981) *Nucleic Acids Res.* 9, 2989–2998.
36. Loenen, W. A. M. & Brammar, W. J. (1980) *Gene* 10, 249–259.
37. Scalenghe, F., Turco, E., Edstrom, J. E., Pirotta, V. & Melli, M. (1981) *Chromosoma* 82, 205–216.
38. Karn, J., Brenner, S., Barnett, L. & Cesareni, G. (1980) *Proc. Natl. Acad. Sci. USA* 77, 5172–5176.
39. Davis, R. W., Thomas, M., Cameron, J., St. John, T. P., Scherer, S. & Padgett, R. A. (1980) *Methods Enzymol.* 65, 404–411.
40. Ilyin, Y. V., Chmeliauskaite, V. G. & Georgiev, G. P. (1980) *Nucleic Acids Res.* 8, 3439–3458.
41. Tchurikov, N. A., Ilyin, Y. V., Skryabin, K. G., Ananiev, E. V., Bayev, A. A., Jr., Krayev, A. S., Zelenstova, E. S., Kulguskin, V. V., Lyubomirskaya, N. V. & Georgiev, G. P. (1981) *Cold Spring Harbor Symp. Quant. Biol.* 45, 655–665.
42. Modolell, J., Bender, W. & Meselson, M. (1983) *Proc. Natl. Acad. Sci. USA* 80, 1678–1682.
43. Finnegan, D. J., Rubin, G. M., Young, M. W. & Hogness, D. S. (1978) *Cold Spring Harbor Symp. Quant. Biol.* 42, 1053–1063.
44. Bingham, P. M., Kidwell, M. G. & Rubin, G. M. (1982) *Cell* 29, 995–1004.
45. Kulguskin, V. V., Ilyin, Y. V. & Georgiev, G. P. (1981) *Nucleic Acids Res.* 9, 3451–3464.
46. Bayev, A. A., Jr., Krayev, A. S., Lyubomirskaya, N. V., Ilyin, Y. V., Skryabin, K. G. & Georgiev, G. P. (1980) *Nucleic Acids Res.* 8, 3263–3273.
47. Lis, J., Neckamey, W., Mirault, M.-E., Artavanis-Tsakonas, S., Lall, P., Martin, G. & Schedl, P. (1981) *Dev. Biol.* 83, 291–300.
48. Mason, P. J., Torok, I., Kiss, I., Karch, F. & Udvardy, A. (1982) *J. Mol. Biol.* 156, 21–35.
49. Kidwell, M. G., Kidwell, J. F. & Sved, J. A. (1977) *Genetics* 86, 813–833.
50. Engels, W. R. (1981) *Cold Spring Harbor Symp. Quant. Biol.* 45, 561–565.
51. Rubin, G. M., Kidwell, M. G. & Bingham, P. M. (1982) *Cell* 29, 987–994.
52. Schedl, P., Artavanis-Tsakonas, S., Steward, R., Gehring, W. J., Mirault, M.-E., Goldschmidt-Clermont, M., Moran, L. & Tissieres, A. (1978) *Cell* 14, 921–929.
53. Botstein, D. & Kleckner, N. (1977) in *DNA Insertion Elements, Plasmids and Episomes*, eds. Bukhari, A. I., Shapiro, J. A. & Adhya, S. L. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 185–203.
54. Tu, C.-P. D. & Cohen, S. N. (1980) *Cell* 19, 151–160.
55. Berg, D. (1977) in *DNA Insertion Elements, Plasmids and Episomes*, eds. Bukhari, A. I., Shapiro, J. A. & Adhya, S. L. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 205–212.
56. Spradling, A. C. & Rubin, G. M. (1982) *Science* 218, 341–347.
57. Barrie, P. A., Jeffreys, A. J. & Scott, A. F. (1981) *J. Mol. Biol.* 149, 319–336.
58. Bernards, R. & Flavell, R. A. (1980) *Nucleic Acids Res.* 8, 1521–1534.
59. Old, J. M., Ayub, H., Wood, W. G., Clegg, J. B. & Weatherall, D. J. (1982) *Science* 215, 981–982.