FORMATION OF CHROMOSOME REARRANGEMENTS BY *P* FACTORS IN DROSOPHILA

WILLIAM R. ENGELS AND CHRISTINE R. PRESTON

Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin 53706

Manuscript received March 13, 1984
Accepted April 21, 1984

ABSTRACT

We studied a collection of 746 chromosome rearrangements all induced by the activity of members of the *P* family of transposable elements in *Drosophila melanogaster*. The chromosomes ranged from simple inversions to complex rearrangements. The distribution of complex rearrangement classes was of the kind expected if each rearrangement came about from a single multibreak event followed by random rejoining of chromosome segments, as opposed to a series of two-break events. Most breakpoints occurred at or very near (within a few hundred nucleotide pairs) the sites of preexisting *P* elements, but these elements were often lost during the rearrangement event. There were also a few cases of apparent gain of *P* elements. In cases in which both breakpoints of an inversion retained *P* elements, that inversion was capable of reverting at high frequencies to the original sequence or something close to it. This reversion occurred with sufficient precision to restore the function of a gene, held-up-b, which had been mutated by the breakpoint. However, some of the reversions had acquired irregularities at the former breakpoints that were detectable either by standard cytology or by molecular methods. The revertants themselves retained the ability to undergo further rearrangements depending on the presence of *P* elements. We interpret these results to rule out the simplest hypotheses of rearrangement formation that involve cointegrate structures or homologous recombination. The data provide a general picture of the rearrangement process and its possible relationship to transposition.

THE production of site-specific chromosome rearrangements is one of the hallmarks of transposable element activity in eukaryotes. This was first indicated by MCCLINTOCK'S studies of maize controlling elements [reviewed by FINCHAM and SASTRY (1974) and FEDEROFF (1983)] and later extended to *Drosophila melanogaster* where high frequencies of rearrangements are generated by the activity of *P* factors (ENGELS and PRESTON 1981a; YAMAGUCHI and MUKAI 1974; YANNOPOULOS, ZACHAROPOULOU and STAMITIS 1982), *L* factors (LIM 1979, 1981) and foldback elements (BINGHAM 1981; LEWIS, COLLINS and RUBIN 1982). Various unexplained cases of rearrangement formation (*e.g.*, LEVITAN 1963; HINTON 1979), especially those in which the rearrangement is observed to revert to wild type (GRUNBERG 1936; NOVITSKI 1961; KALISCH 1970), also seem likely to involve transposable elements. In prokaryotes rearrangements often appear as intermediate stages in the transposition process

(reviewed by CALOS and MILLER 1980), but less is known about these processes in higher organisms.

P factors are members of a family of transposable genetic elements present in many scattered and variable locations in the genomes of some *D. melanogaster* strains (*P* strains), but they are absent in others (*M* strains). Molecular studies of *P* factors (BINGHAM, KIDWELL and RUBIN 1982; RUBIN, KIDWELL and BINGHAM 1982; O'HARE and RUBIN 1983) reveal that they actually compose a heterogeneous family of elements. The group we will call major elements has 2907 base pairs with precise inverted terminal repeats of 31 base pairs and the capacity to encode two or more gene products. Other *P* elements are shorter and appear to outnumber the major elements by approximately 2:1 in at least one *P* strain. Sequencing indicates that these shorter elements are derivable from the major elements by deletions internal to the end repeats (O'HARE and RUBIN 1983). *P* elements usually display only low levels of transpositional activity, but, under certain conditions, they switch to a highly active state in which they cause a syndrome of germline abnormalities known as hybrid dysgenesis. This syndrome, which includes mutability and temperature-sensitive sterility, is thought to be associated with high transposition frequencies. The cellular environment that supports this activity is known as the *M* cytotype since it is normally present in *M* strains. The unusual inheritance of cytotype (ENGELS 1979a, 1981; ENGELS and PRESTON 1981b) is such that the hybrid progeny of *M* strain females and *P* strain males have the *M* cytotype. They also have *P* factors inherited from their fathers that can become active early in the development of the germ cells and cause the hybrid dysgenesis syndrome. [See BREGLIANO and KIDWELL (1983), RUBIN (1983), ENGELS (1983) for recent reviews of *P* factors.]

P factors can produce chromosome rearrangements at extremely high frequencies: approximately 10% per chromosome arm per generation in dysgenic hybrid males (BERG, ENGELS and KREBER 1980). Approximately 85% of the breakpoints of these rearrangements occur at the positions of *P* elements, and the remainder ("sporadics") have an apparently uniform distribution along the chromosomes (ENGELS and PRESTON 1981a). Moreover, these rearrangements occur preferentially in the *M* cytotype, suggesting that they are somehow related to transpositional activity.

In this paper, we present results of genetic, cytological and molecular analyses of a large collection of *P* factor-induced chromosome rearrangements. The results provide information on the kinds of simple and complex rearrangements produced, the developmental and cell cycle timing of the events and some aspects of the molecular nature of the rearrangement breakpoints. Contrary to our expectations from prokaryotic models, we find that *P* elements are not necessarily replicated or even conserved in the process of rearrangement formation, nor is there any tendency for rearrangement events to occur in unit steps involving only two chromosomal breakpoints at a time. There is probably some relationship between rearrangement formation and transposition, but it does not appear to be analogous to cases previously studied in other organisms.

MATERIALS AND METHODS

Stocks and crosses: All *X* chromosomes used in this study were derived from a *P* strain known as π_2. The π_2 strain came from a local wild population in 1975 and was inbred for many generations as described previously (ENGELS and PRESTON 1979, 1980). Compound-*X* strains of the *P* and *M* cytotype and standard-*X* tester stocks are described in ENGELS and PRESTON (1981a).

All matings were set up individually so that independent events could be distinguished from premeiotic clusters. Chromosome rearrangements and revertants were obtained by crossing males from a *P* stock to compound-*X*, *M*-cytotype females and crossing the resulting dysgenic sons to an appropriate tester stock so that rearrangements could be identified in the next generation.

If a rearrangement was recovered in a male, he was doubly mated to standard-*X* tester females to yield larvae for cytology and compound-*X* females, usually of the *P* cytotype, to establish a stock in which the chromosome of interest would not be expected to undergo further rearrangements. In establishing the stock, one or two additional generations of backcrossing to compound-*X*, *P* cytotype females were usually carried out to ensure long-term stability of the *P* cytotype.

Cytology: Salivary glands from third instar larvae were dissected in Ringer's solution, fixed in ethanol and acetic acid, stained in aceto-orcein and squashed on gelatinized slides with siliconized coverslips. Each rearrangement was examined separately by each of us; if we did not agree on the placement of breakpoints, we recorded the range of uncertainty to include both placements.

In situ hybridization: Chromosomes were maintained in *P* cytotype stocks for ten to 50 generations before the *in situ* hybridization slides were made. For all viable rearranged chromosomes, larvae for *in situ* hybridization were homozygous females or, in a minority of cases, males. For the deficiencies and the standard-sequence chromosomes, we used female larvae heterozygous for the chromosome of interest and an *M* chromosome of standard cytological sequence. Slides were prepared as described earlier except that 45% acetic acid was used in place of aceto-orcein. Hybridization was according to the modification by BINGHAM, LEVIS and RUBIN (1981) of PARDUE and GALL'S (1975) procedure except that labeled RNA was used in place of DNA, hybridization was at 37° rather than 25° and slides were treated with RNase following hybridization. The DNA templates were provided by G. M. RUBIN. The probe described as the internal sequences of the *P* factor consists of a 1:1 mixture of the left *Hin*dIII fragment and the *Hin*dIII-*Sal*I fragment as shown in Figure 1 of O'HARE and RUBIN (1983).

EXPERIMENTS AND RESULTS

Classes of rearrangements: We examined 746 independently derived rearrangements by cytological analysis of salivary chromosomes. These rearrangements were recovered by a variety of methods, but most were detected by screening the progeny of dysgenic flies for mutations at the *hdp-b* (heldup wings, formerly *hdp*) locus at cytological position 17C of the *X* chromosome. This method takes advantage of the earlier finding (ENGELS and PRESTON 1981a) that a particular *P* strain, π_2, has a *P* factor site at cytological position 17C2-3 which appears to coincide with the *hdp-b* locus [see LEFEVRE (1976) and LINDSLEY and GRELL (1968) for cytological and genetic terminology]. A mutation at the *hdp-b* locus appears whenever a breakpoint occurs at the site of the *P* factor. Thus, we generated a set of rearrangements with one selected breakpoint at the *hdp-b* locus and one or more nonselected breakpoints elsewhere in the genome. This collection is pooled from experiments involving several *X* chromosomes that differ primarily in the chromosomal positions of their *P* elements. The pooling is justified since, for the present purposes, we are primarily concerned with the general process of *P*-induced rearrangement formation rather than the specific breakpoint positions. Thus, in considering

TABLE 1

Cytological analysis of chromosome rearrangements

	Conservative rearrangements					
	Two-point	Three-point	Four-point	Five-point	Duplications	Deficiencies
Observed	650[a]	60[b]	13[b]	4[b]	8[c]	12[d]
Expected[e]	637.1	82.3	7.1	0.5		
Expected[f]	640.7	75.8	9.3	1.5		

[a] One nonselected breakpoint was at 22B, and all the rest were on the *X* chromosome.

[b] All breakpoints were on the *X* chromosome.

[c] Chromosome designations and cytological sequences are: *Dp(1,1),F14f*:[tip-17C | 12F-13E | 17C-base]; *Dp(1),B395.2*:[tip-17C | (unidentified region of one or several bands) | 17C-base]; *Dp(1,1): L40*:[tip-7A | 19F-17C | 5A-17C | 19F-base]; *Dp(1,1),N71*:[tip-17C | (unidentified sequence of several bands) | 17C-base]; *Dp(1,1),N224*:[tip-5E | 17C-4F | 17C-base]; *Dp(1,1),N254*:[tip-17C | 5E-4F | 17C-base]; *Dp(1,1),O19*:[tip-5E | 7A-5E | 5E-base]; *Dp(1,1),N86*:[tip-17C | (short region of one or more unidentified bands) | 17C-base].

[d] Deletions recovered as revertants of Beadex mutation. Chromosome designations and breakpoints are: *Df(1)E76*,(17B;17E); *Df(1)E88.2*,(17C;18A); *Df(1)E92*,(17A;18AB); *DF(1)E107*,(17B; 18BC); *Df(1)E128*,(17C;18A); *Df(1)E132*,(17B3-5;18A5-7); *Df(1)E160.1*,(17A;18A2); *Df(1)E160.2*, (17B3-6;18A5-7); *Df(1)E131.2*,(17B3-5;18B5-10); *Df(1)E133*,(17B3-5;18B5-10); *Df(1)E171*,(17B1-2; 18A4-7); *Df(1)E187*,(17B;18A).

[e] Expected numbers if nonselected breakpoints are assumed to come from a Poisson distribution with parameter 0.26 and with the zero class unobservable.

[f] Expected numbers if k nonselected breakage events are generated from a Poisson process with parameter 0.12 (determined by maximum likelihood) followed by random rejoining of segments to produce a rearrangement with $k + 1$ or fewer breakpoints. Note that this rejoining process can restore the chromosome to its original sequence. Rearrangements with zero nonselected breakpoints are assumed unobservable.

frequencies of rearrangement classes irrespective of breakage sites, the collection can be treated as statistically homogeneous.

Since most of these rearrangements were recovered in males, the collection is heavily biased against recessive lethal or male-sterile rearrangements. Thus, deficiencies and large duplications are underrepresented in this collection. In almost all cases the cytology was carried out in the generation immediately following recovery to minimize the possibility of a series of rearrangements being misinterpreted as a single event. When it was necessary to wait longer, we maintained the chromosomes in *P* cytotype stocks where *P* factor activity is greatly reduced.

The results are summarized in Table 1. Among the conservative rearrangements (those without duplications or deficiencies) the great majority involved only two breakpoints (649 inversions and one reciprocal translocation). Three-, four- and five-point rearrangements follow in decreasing frequencies. The great excess of inversions over translocations is at least partly explained by our screening procedure; many *X*-autosome translocations are male sterile, as reviewed by LINDSLEY and TOKUYASU (1980) and would, therefore, be eliminated. In addition, some of the rearrangements were recovered from experiments in which the autosomes were *M* derived and therefore lacking in breakage hotspots.

If we assume that the number of nonselected breakpoints has a Poisson

distribution, then the maximum likelihood estimate of the average number of nonselected breaks per *X* chromosome per generation is 0.26. This implies that the realized frequency of rearrangements (number of chromosomes with two or more breaks as a fraction of the total minus the one-break events that would not be observed) is 3.6%, approximately one-third the value obtained from earlier results in which this frequency was measured directly (BERG, ENGELS and KREBER 1980) but in which recessive lethal and sterile rearrangements were not eliminated. The expected numbers of each class under this hypothesis are in Table 1. They differ from the observed numbers by their overestimation of the frequency of the three-point rearrangements and underestimation of the other classes. A much better fit can be obtained if we assume that the breakage events themselves are Poisson distributed, but that this process is followed by a random rejoining of the segments to produce either the standard sequence or a rearrangement involving any subset of the breaks. From combinatorial considerations, the details of which appear in APPENDIX, we obtained the maximum likelihood expectations given in Table 1. Our method of analysis differs from that of KAUFMANN (1941), FANO (1941) and LEA (1962) who studied radiation-induced rearrangements. Those authors did not allow for multiple breakage events on the same chromosome arm. Such events compose the majority of the cases studied here and were, therefore, included in our calculations. In addition, FANO and LEA included rejoinings that led to dominant lethals such as large deletions or dicentric chromosomes, whereas we find that a better fit to our data is obtained by considering only viable rearrangements. The resulting expectations are still distinguishable from the data by a standard goodness-of-fit test, but they are clearly close enough to indicate the usefulness of this model. We conclude that to a first approximation the breakpoints of *P*-induced rearrangements can be considered as generated by a Poisson process.

The eight duplications in Table 1 all occurred on the *X* chromosome in the germlines of dysgenic males. Since males have only one *X* chromosome, we can infer that these events happened during or after the DNA synthesis phase of the cell cycle. One of these duplications is shown in Figure 1. This is a four-point rearrangement in which the segment from 17C to 19F is reversed in polarity and transposed to a distal position where the segment from 5A to 7A is duplicated in direct orientation on either side of the inserted sequence. The generation of this rearrangement probably involved a break at 5A on one sister chromatid and 7A on the other.

Deficiencies are also generated at high frequencies in dysgenic flies, but most of them are eliminated by male lethality. The 12 deficiencies shown in Table 1 were recovered by a special procedure in which daughters of dysgenic males carrying the dominant *Bx* (Beadex wings) mutation were screened for reversion to Bx^+. These revertants were all found to have short deletions covering the *Bx* locus, with breakpoints as given in Table 1. The lack of any clear tendency toward common breakpoints is expected in this case, since *in situ* hybridization showed that the parental *Bx* chromosome had no *P* elements near *Bx*. Thus, the breakpoints in these deficiencies are all of the sporadic type. Much larger deficiencies along with *X-Y* translocations can also be recovered from dysgenic

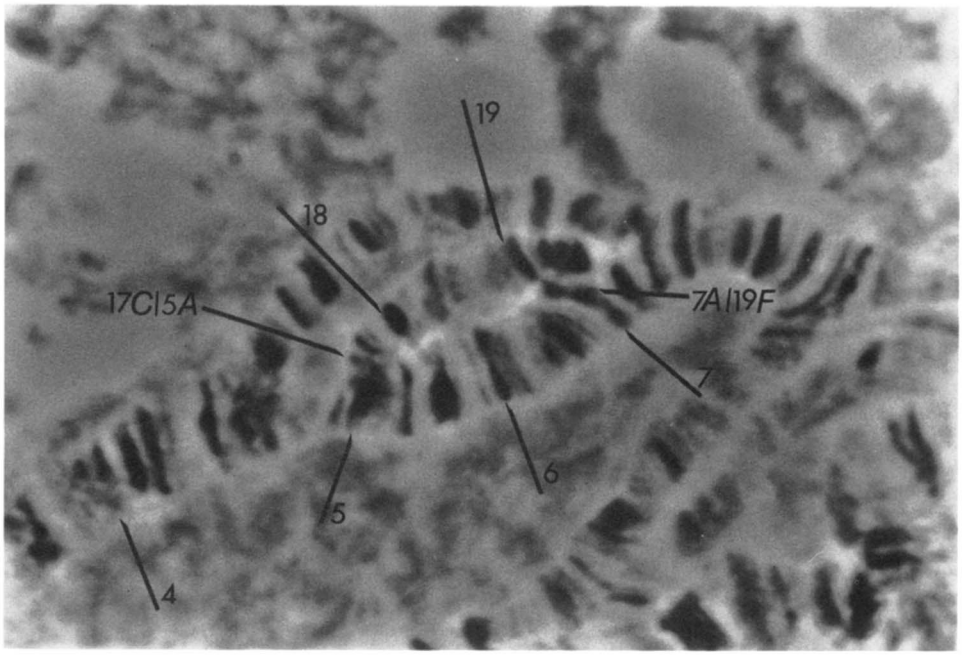

FIGURE 1.—Heterozygote for rearrangement L40 with sequence: tip-7A|19F-17C|5A-17C|19F-base. The loop is caused by the 5A-7A duplication pairing with itself on the rearranged chromosome. The upper strand is the normal homologue. The proximal two breakpoints are not shown.

males as free duplications using the method of PAINTER and MULLER (1929). Such events usually involve one breakpoint at the most distal *P* element of the *X* chromosome and, in the case of the deficiencies, the other breakpoint is at a proximal element, deleting everything between these points (BERG, ENGELS and KREBER 1980). These occur at much higher frequencies than the sporadic events mentioned earlier (approximately 1 *vs.* 0.09%), but they are not included in Table 1 since we made no systematic attempt to analyze them.

Precision of breakage at P *factor sites*: Genetic evidence that breakage hotspots such as positions 5E and 17C of the π_2 *X* chromosome correspond to *P* factor positions (ENGELS and PRESTON 1981a) was well supported by the observation (BINGHAM, KIDWELL and RUBIN 1982) that *in situ* hybridization of *P* factor sequences occurs at these sites. However, since *in situ* hybridization is only able to resolve positions to within 10–1000 kb, it was of interest to refine this observation using DNA probes cloned directly from the regions flanking the *P* elements postulated to be at the breakage hotspots.

Two plasmid clones, designated pπ25.1 and pπ12.20, were derived and analyzed in detail by O'HARE and RUBIN (1983). Both clones contain *Bam*HI fragments with *P* factors from the π_2 genome. These fragments also have genomic sequences on both sides of the elements. O'HARE and RUBIN also obtained the corresponding "empty" fragments, designated pS25.1 and pS12.20, which were derived by homology from an *M* strain genomic library. Thus, pS25.1 and pS12.20 contain only the restriction fragment into which the *P* factor is inserted,

but not the element itself. By *in situ* hybridization of pπ25.1 to *M* strain chromosomes, SPRADLING and RUBIN (1982) determined that this element came from position 17C. We performed the same test for pπ12.20 and found its position to be 5E, which is also one of the breakage hotspots on the π_2 *X* chromosome.

To determine whether our 17C2-3 and 5E3-7 breakage hotspots coincide with these cloned *P* factors, we hybridized labeled pS25.1 and pS12.20 to rearranged chromosomes with breaks at these two sites. For each of 11 rearrangements involving the 5E3-7 site, we found that the pS12.20 probe labeled both of the rearrangement breakpoints. Similarly, we examined eight rearrangements involving 17C2-3 and found that in each case the pS25.1 probe labeled both breakpoints. These results were consistent in complex rearrangements. For example, the sequence

tip-2F|5E-2F|17C-5E|17C-base

was labeled by the pS12.20 probe at the 2F|5E junction and at the 5E|17C junction, whereas the pS25.1 probe hybridized to 2F|17C and 5E|17C. These results show that breakage of the 5E3-7 and 17C2-3 breakage hotspots occurs within the *Bam*HI fragments that contain the *P* factors at those cytological positions.

We determined the orientation of the pπ25.1 sequence with respect to the chromosome centromere by *in situ* hybridization to a 7D1-2;17C2-3 inversion using as probes two subclones obtained from K. O'HARE (personal communication). These clones contain sequences homologous to one side or the other of the 7D1-2 breakpoint. We found that the left end of the fragment, as pictured in O'HARE and RUBIN'S (1983) Figures 1–5, is proximal to the centromere. The probable chromosomal orientation of the pπ12.20 clone was inferred from the tendency for heavier labeling of the proximal endpoints of inversions analyzed by the *in situ* hybridizations described earlier. We, therefore, tentatively assign right side pπ12.20, as shown in O'HARE and RUBIN'S diagrams, to the proximal position, since it has greater (3.7 kb) homology. This assignment, however, will require confirmation.

If our orientation assignments are correct, then this experiment places the 17C2-3 breakpoints within 0.8 kb distal and 1.0 kb proximal to the pπ25.1 *P* factor. For the 5E3-7 hotspot, the limits are within 1.1 kb on the distal side and 3.7 kb on the proximal side of the pπ12.20 element. Since at least several hundred base pairs of homology are required for detection by our *in situ* hybridization procedure, the actual breaks must be well within these bounds. A likely interpretation is that the breakpoints are at or within the boundaries of the *P* elements themselves.

Clusters of rearrangements from premeiotic events: Rearrangement events sometimes occur prior to meiosis as indicated by clusters of two or more in the progeny of a single dysgenic fly (SIMMONS and LIM 1980). Table 2 shows some examples. Cluster sizes cover a wide range, suggesting that the events are liable to occur at many stages during development of the germline. The majority of clusters are homogeneous, as in the family designated F5i of Table 2. Given the high frequency of recurrence of breakage sites, it is likely that some of the homoge-

TABLE 2

Examples of clusters from premeiotic events

Chromosome[a]	Cytological sequence
F5i.1	tip-5E | 17C-5E | 17C-base
F5i.2	tip-5E | 17C-5E | 17C-base
F5i.3	tip-5E | 17C-5E | 17C-base
F5e.1	tip-5E | 17C-5E | 17C-base
F5e.2	tip-4F | 5E-4F | 17C-5E | 17C-base
F3b.1	tip-2F | 17C-2F | 17C-base
F3b.2	tip-2F | 17C-2F | 17C-base
F3b.3	tip-2F | 17C-2F | 17C-base
F3b.4	tip-5E | 17C-5E | 17C-base
F3b.5	tip-2F | 5E-2F | 5E-11A | 17C-11A | 17C-base
F4g.1	tip-2F | 17C-2F | 17C-base
F4g.2	tip-2F | 17C-2F | 17C-base
F4g.3	tip-2F | 20A-17C | 2F-17C | 20A-base
D83f.1	tip-5E | 17C-5E | 17C-base
D83f.2	tip-2F | 17C-18D | 9E-17C | 5E-9E | 5E-2F | 18D-base

[a] Symbols preceding the decimal point designate a given dysgenic parent, and the numeral following the decimal point denotes one offspring of that parent.

neous clusters actually represent more than one independent event. However, in analyzing all rearrangement data, we assume that homogeneous clusters came from one premeiotic event and count them as such.

Heterogeneous clusters were also common. Often members of a heterogeneous cluster shared one or more breakpoints in common, such as the families F5e, F3b, F4g and D83f in Table 2. In a few cases, such as in F5e, one member of a cluster differs from another only by the addition of a single inversion, raising the possibility that complex rearrangements can come about through a series of two-point rearrangements at different developmental stages. In most cases, however, such as the last three clusters of Table 2, there is no simple transition that will convert one cluster member to another. Furthermore, a sequential process would not lead to the Poisson distribution of breakage events indicated by the data. It appears, therefore, that examination of clusters is not sufficient to determine whether complex rearrangements occur through a sequence of individual events. This question will be addressed in the next section.

The occurrence of complex rearrangements as single events: To designate complex rearrangements, we will use the letters a, b, c, . . ., to represent the internal segments in order of their position in the standard sequence. Reversal of polarity relative to the standard sequence is indicated by an overbar. To examine the question of whether rearrangements come about by a sequential process, note that some kinds of rearrangements are easily generated sequentially by two-point exchanges, but others are not. For example, consider the four possible kinds of three-point rearrangements. Three of these rearrangements ($\overline{ab}$, $\overline{b}a$, and $b\overline{a}$, see

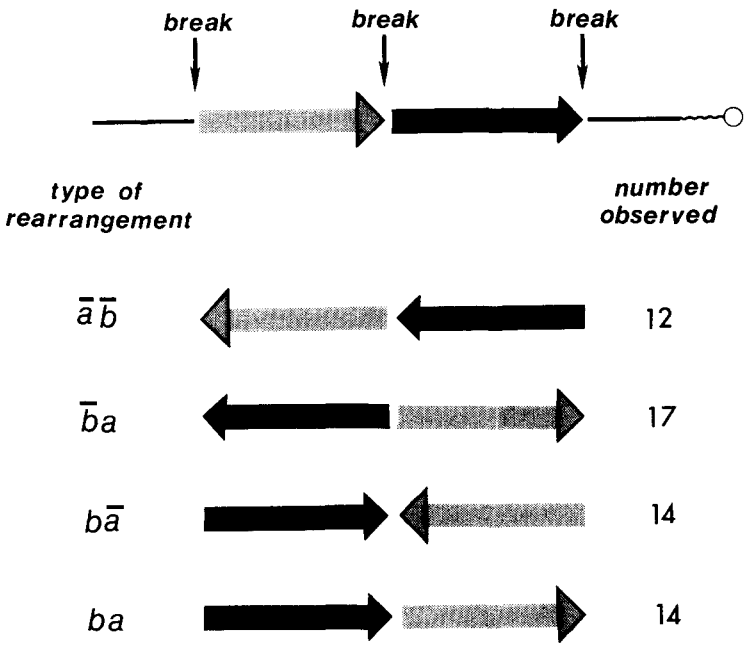

FIGURE 2.—The four possible kinds of three-point rearrangments and the observed number of each.

Figure 2) can be obtained by two inversion steps, but the fourth one (ba) cannot. It requires at least three inversion steps and would, therefore, be much less likely to occur by a sequential process. If all rearrangements occurred by two breakpoints at a time, we would expect one type (ba) to be much less frequent than the other types.

To test this prediction, we determined the band sequence of 57 of the 60 three-point rearrangements of Table 1 and classified them according to these four categories. The results (Figure 2) show that the four types actually occur at approximately equal frequencies, as would be expected from a single multibreak event in which the fragments are rejoined at random as opposed to a sequential process. Data from four- and five-point rearrangements are less extensive, but the same conclusion is indicated. For four-point rearrangements, the only ones that can occur in two steps are $\bar{a}b\bar{c}$, $\bar{a}cb$, $\bar{b}c\bar{a}$, $ba\bar{c}$, $c\bar{a}\bar{b}$ and $\bar{c}b\bar{a}$. There are 19 other possible kinds, all requiring at least three steps. Therefore, we would expect essentially all four-point rearrangements that occurred by a sequential process to be one of these six two-step kinds. However, only eight of 11 whose band sequences we determined fell into these six categories. This frequency is intermediate between the expectations based on purely random and purely sequential processes, suggesting that both possibilities can occur. The five-point rearrangement in chromosome D83f.2 of Table 2 is of type dcb$\bar{a}$, which would require a minimum of four inversion steps instead of three as would be expected of five-point rearrangements if they occurred sequentially. Moreover, if the starting point is assumed to be the simple inversion present as a member of the same cluster, then this rearrangement is equivalent to type $d\bar{b}c\bar{a}$, which still

requires at least four inversion steps. The other three five-point rearrangements (not shown) were of types $b\bar{a}c\bar{d}$, $\bar{a}cb\bar{d}$ and $\bar{c}dba$, which require a minimum of three, four and three inversion steps, respectively. (Of the 208 distinguishable kinds of five-point rearrangements, 57 can be obtained in three inversion steps, and the others require four or five.) Under the hypothesis of sequential inversions we would expect that the vast majority of five-point rearrangements would be of the kind that can come about in three steps, and our observation that two of four such rearrangements require four steps would be highly unlikely.

This analysis does not imply, of course, that sequential rearrangements cannot occur. In fact, chromosomes kept in the *M* cytotype for many generations would certainly be expected to have an excess of complex rearrangements of the stepwise types. Even for the present data in which the cytotype was *M* for usually no more than two generations, it is likely that some of the chromosomes came from two or more separate events. This is particularly true for the four-point rearrangements in which one of the two-step categories, $\bar{c}b\bar{a}$, was observed five times. However, our overall conclusion is that complex rearrangements can, and normally do, come about through a single multibreak event.

Lack of a proximity effect in reunion of chromosome segments: Another alternative worth considering is that the rejoining process preferentially connects points that are close to each other on the chromosome. (Note that this hypothesis is different from that of breakage sites themselves occurring preferentially in close proximity.) To examine this possibility, we will characterize each novel junction by the total distance that would normally separate these points in a standard chromosome and score the rearrangements according to the total of these distances summed over all breakpoints. Our measure of distance will be the number of chromosome segments in the given rearrangement rather than absolute physical distance. In this way we can distinguish between rearrangement types irrespective of the actual positions of the breakpoints. For example, in the three-point rearrangement $\bar{b}a$ the leftmost breakpoint connects two loci that are normally separated by two segments (a and b), and the other two breakpoints connect loci normally separated by one segment. Therefore, the total distance for this class of rearrangement is four segment lengths. No relevant information can be obtained from examining the two- or three-point rearrangements since they necessarily result in total distances of one and four segments, respectively, regardless of how the segments are rejoined. However, for the 25 possible four-point rearrangements, one type ($\bar{a}b\bar{c}$) has a distance of four segment lengths, 12 types have a distance of six and the remainder have a distance of eight. Accordingly, we would expect under this hypothesis that the actual rearrangements would favor the types with distances of four or six. In fact, of the 11 four-point rearrangements whose sequences we determined, only two had a total distance of four, one had a distance of six and eight had a distance of eight. Therefore, there is no indication of nonrandom junctions favoring nearby breakpoints. If anything, there is a slight tendency in the opposite direction.

The fate of P *elements at rearrangement breakpoints*: A labeled probe comprising most of the *P* factor internal sequences but no genomic flanking DNA (see MATERIALS AND METHODS) was hybridized *in situ* to a series of chromosome

rearrangements to determine whether *P* sequences are present at the breakpoints. For each rearrangement we also determined the *P*-labeling sites on the parental chromosome in order to detect any gain or loss of an element in the process of rearrangement formation. Since stocks were maintained in the *P* cytotype, we assume that these results almost always reflect the original state of the chromosomes, even though some generations have occurred since the rearrangement. The set of chromosomes was chosen to include cases in which all breakpoints occurred at parental *P* factor sites and also cases of sporadic breaks in previously empty sites.

The results (Figure 3) indicate that both gains and losses of *P* sites can occur in the formation of rearrangements, but the net tendency is for *P* elements to be lost from the breakpoint. In our sample there were 14 cases of rearrangement events that resulted in a net loss of *P*-labeling sites. In five of these chromosomes, there were no *P* elements remaining at any of the breakpoints, thus ruling out the possibility that the apparent loss is merely the juxtaposition of two elements. The most extreme example is chromosome D10a.3 in which a three-point rearrangement coincided with the loss of all three *P* elements at the breakpoints, even though other sites on the parental chromosome (not shown) remained. Three of the rearrangements (D83f.2, N79.2 and N715.3) had a net gain of one site each, and in the remaining 14 chromosomes there was no net gain or loss. It was not possible to quantify objectively the intensity of labeling at each site, but there were three cases (M62, N79.2 and D83f.2) in which labeling at one breakpoint was clearly and consistently weaker than other reference sites in the same nucleus.

Heterozygotes for two of the deficiencies (E132 and E128) were also examined by *in situ* hybridization. E132 is shown in Figure 4. As mentioned earlier, these deficiencies were derived in such a way that the breakpoints are necessarily sporadic. In each case there was a *P* element present at the deficiency breakpoint, suggesting that the transposition of a *P* element from another site coincided with the deletion of several bands of genomic DNA.

Finally, we tested four of the duplications for hybridization to the internal *P* sequences. Two of them had *P* elements at both endpoints of the duplication, and the other two were not labeled at either site.

It should be noted that some very short *P* elements might not be detected by this procedure, especially if the element had homology only to the right side of the major *P* factor sequence. Thus, some of the cases of apparent loss of *P* elements in this experiment might actually be only partial losses. However, all *P* elements that have been characterized to date have sufficient homology with the probe to be easily detected, indicating that such "invisible" elements are at most rare.

We conclude that the process of rearrangement formation involves in different instances the gain, loss or conservation of *P* elements at the breakpoints. There is no apparent relationship between the rearrangement type and the behavior of *P* elements at the breakpoints. Events that might appear to be similar at the level of cytology reveal a wide variety of types at the DNA level.

Reversion of rearrangements to restore hdp-b$^+$ *phenotype*: To detect reversions of

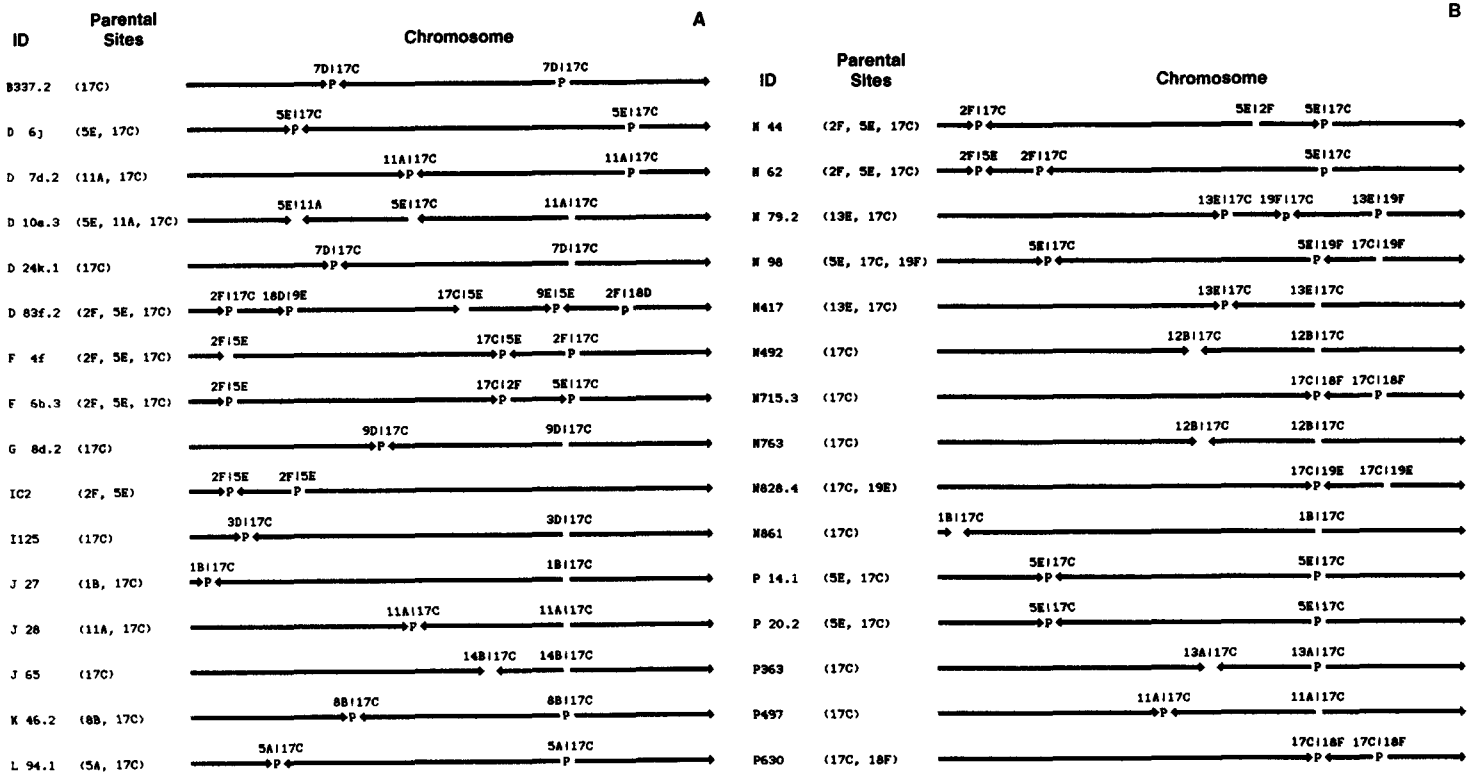

FIGURE 3.—The fate of *P* elements at rearrangement breakpoints. Each row shows the sequence of an *X* chromosomal rearrangement, the relevant *P* element sites on the parental chromosome before the rearrangement and the presence or absence of *P* sites at the rearrangement breakpoints. The polarity of each chromosome segment is indicated by an arrowhead at the point that is normally proximal to the centromere. Each breakpoint is characterized as either *P*, indicating a strongly labeling site, *p* designating a weakly labeling site or blank if there was no label at all. Only sites at breakpoints are shown. Chromosomes P14.1 through P630 were derived from reverted rearrangements.

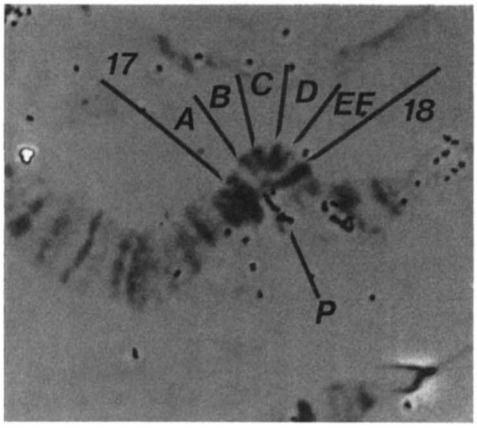

FIGURE 4.—*In situ* hybridization of internal *P* element sequences to a heterozygote for deficiency E132. The breakpoint 17B3-5|18A5-7 is indicated by *P*.

TABLE 3

Reversion of rearrangements

ID	Chromosome type	Stock males/dysgenic males	*hdp*$^+$ revertants
D6j	tip-5E|P|17C-5E|P|17C-base	9/150	16[a]
D7d	tip-11A|P|17C-11A|P|17C-base	7/80	2
D133i	tip-2F|P|17C-2F|P|17C-base	15/220	1
D24k[b]	tip-7D|P|17C-7D| |17C-base	3/43	1
B337.2[c]	tip-7D|P|17C-7D|P|17C-base	7/197	0
F4f	tip-2F| |5E-17C|P|5E-2F|P|17C-base	1/38	0
D10a.3	tip-5E| |11A-5E| |17C-11A| |17C-base	8/178	0

[a] One of these is the duplication O19 shown in Figure 5.
[b] Distal breakpoint is 7D14-18 and therefore to the right of singed.
[c] Distal breakpoint is 7D1-2, the singed locus.

rearrangements we selected seven of our rearrangement stocks with breaks at 17C2-3 and, therefore, the *hdp-b* phenotype. The selected chromosomes also had one or more visible markers to guard against the possibility of contamination. For each rearrangement we screened the progeny of a series of dysgenic males for reversion to *hdp-b*$^+$. As summarized in Table 3, some of the chromosomes, especially D6j, yielded revertants at high frequencies, whereas others reverted rarely or not at all.

Some of this variation in revertability can be explained in terms of the nature of the rearrangement. Thus, the D6j chromosome that reverted at a high frequency possesses *P* factor sites at both breakpoints so that these two sites are expected to behave as breakage hotspots. On the other hand, rearrangement D10a.3 has no visible *P* elements at its breakpoints, and it failed to revert. Note that this is a three-point rearrangement (see Figure 3), but it only requires a simple reversion to restore the 17C region. There were also no revertants of the three-point rearrangement F4f, even though both halves of the *hdp-b* locus

possess *P* elements. In this case, however, reversion would require a three-point rearrangement and would, therefore, not be expected at high frequencies. Surprisingly, one revertant was derived from chromosome D24k which lacks an observable *P* element at its proximal breakpoint. The revertant chromosome has *P* sites at both its former breakpoints. In this case, there might be a small *P* element (fewer than 300 bases of overlap with the probe) that would not be visible by our procedure.

The B337.2 chromosome is a special case. It was derived as we described previously (ENGELS and PRESTON 1981a) from the sn^w chromosome which has a highly unstable *P* insertion at the singed bristle locus. The formation of the B337.2 inversion occurred simultaneously with a mutation to a more extreme singed phenotype, and its distal breakpoint is at the cytological position of singed. It was, therefore, a double mutant at singed and heldup corresponding to the two breakpoints of the inversion. Despite the presence of *P* elements at both endpoints of this inversion, we found no revertants of either *sn* or *hdp-b* in the progeny of 197 dysgenic males. A possible explanation for the discordance between B337.2 and D6j comes from the recent restriction mapping by H. ROIHA, K. O'HARE and G. RUBIN (personal communication) of the B337.2 breakpoints. The *P* elements at these points were found to differ from other *P* elements described by O'HARE and RUBIN (1983), and preliminary indications are that one of the *P* element's termini is missing.

Each of the 20 phenotypic revertants in Table 3 was examined cytologically, and all were found to have reverted to the standard cytological sequence in the neighborhood of *hdp-b*. One of the revertants, designated O19 in the table, also had a new duplication present at the site of the nonselected breakpoint. Figure 5 shows the double-hairpin structure resulting from the palindromic sequence in a homozygote for this reverse-tandem duplication. The other 19 revertants were cytologically normal at the site of the nonselected breakpoint, indicating that phenotypic reversion occurred by simple reinversion of the rearrangement.

It is possible that some additional reinversions occurred but were not detected in our experiment because the reinversion failed to restore the $hdp\text{-}b^+$ phenotype. However, for at least some of the inversions used in this study, the frequency of recovery of $hdp\text{-}b^+$ reversions is approximately the same as the frequency of forward inversions with the same breakpoints. Therefore, unless the total reversion frequency is considerably greater than that of equivalent new inversions, most of the reinversions that occur do restore the $hdp\text{-}b^+$ phenotype and were, therefore, recovered in our screen.

By *in situ* hybridization of the *P* sequences to nine of the revertants, we observed that both former breakpoints remained labeled in six cases, and only one former breakpoint was labeled in the other three cases. Thus, the reversions appear similar to the original inversion events regarding the net tendency for element loss and in the lack of consistency in this respect.

We also tested several of the revertants for hybridization to the flanking DNA probes. For each of five revertants to which the pS25.1 probe was hybridized, the labeled site at the distal former breakpoint, which was present on the parental chromosome, had vanished leaving only the normally labeled site at 17C. This observation suggests that the reversions were precise or nearly so. However, only

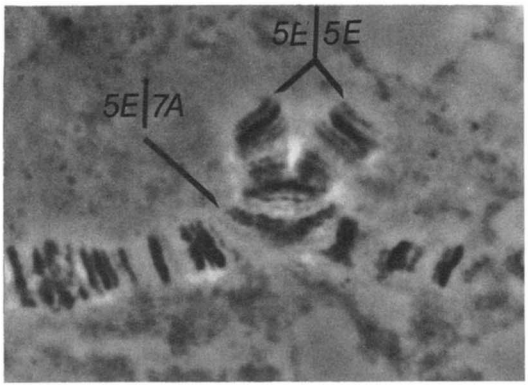

FIGURE 5.—Homozygote for the reverse tandem duplication in chromosome O19 showing double-hairpin loop. The sequence of this chromosome is tip-5E|7A-5E|5E-base.

one of two revertants of rearrangements involving the 5E site tested with the pS12.20 probe appeared to be precise. The other chromosome (O16, see Table 4) had a lightly labeled site at 17C in addition to its normal and more heavily labeled site at 5E. Therefore, the event leading to revertant O16 involved either imprecise rebreakage or a duplication of the unique DNA flanking the *P* factor. Since the reversions were selected for their *hdp-b*$^+$ phenotype, this result also indicates that the extra 5E DNA at the *hdp-b* locus does not prevent its expression.

In conclusion, some, but not all, *P*-induced chromosome rearrangements revert at high frequencies to the original sequence and restore the function of the *hdp-b* gene at one of the breakpoints. We can explain at least some of the variation of revertability in terms of the presence or absence of *P* elements at the breakpoints. These reversions are not always precise with respect to the genomic DNA flanking the *P* element at the breakpoint.

Return of inversions on reverted chromosomes: To further characterize the nature of the reverted chromosomes, we selected six of the 20 *hdp-b*$^+$ revertants, including O19 with its cytologically visible duplication and O16 with its extra S12.20-labeled site at 17C, and returned them to the dysgenic condition to search for new *hdp-b* rearrangements among their offspring. As the results in Table 4 show, all but one of these chromosomes produced fresh breaks at 17C resulting in *hdp-b* mutations. The only exception was O156.3 which was also the only one that lacked a *P* hybridization site at 17C. Given the number of *P* factors present on the parental chromosomes, the frequency of *hdp-b* rearrangements is comparable to that of the original π_2 chromosome. The *P* factor at 17C, therefore, can retain its behavior despite having passed through an inversion-reinversion cycle.

In the case of O16 there was a strong tendency for 5E, the site of the original breakpoint and the source of the extra DNA now present at 17C, to be especially susceptible to further breakage. This chromosome yielded 11 new rearrangements, and all of them had 5E as the nonselected breakpoint even though there were several other *P* elements visible by *in situ* hybridization (not shown). The extra 5E DNA at position 17C might be responsible for this behavior.

TABLE 4

Reinversion of inversions

ID	Parental chromosome	Former break	Stock males/ dysgenic males	Independent *hdp* rearrangements	Nonselected breakpoints
O16	D6j	5E	5/150	11	5E,5E,5E,5E,5E,5E,5E,5E,5E,5E,5E
O82	D6j	5E	5/50	7	5E,5E,5E,7C,7C,7F,7F
O19[a]	D6j	5E	6/150	1	18D
O162.1	D7d.2	11A	6/100	7[b]	1D,1E,5E,11A,12B,18A,18D,20A
O156.3	D7d.2	11A	5/100	0	
O238	D24k	7D	11/250	3	18F,18F,19F

[a] See Figure 5.
[b] Includes one three-point rearrangement.

In situ hybridization of *P* sequences to these new rearrangements (five of them are included in Figure 3) revealed no qualitative differences from events occurring on the original chromosomes; *P* elements are sometimes, but not always, conserved at the breakpoints.

We did not continue this cycle to re-revert these inversions back to *hdp-b*$^+$, but all of the data indicate that the *hdp-b* locus remains a reliable indicator of whether there is a breakpoint at the 17C *P* factor, and that the inversion-reinversion cycle is limited only by the experimenter's patience.

DISCUSSION

Time and place of rearrangement formation: *P* factor activity appears to be primarily limited to the germline, as indicated by studies of other traits, especially *sn*^{w} hypermutability in which somatic mosaicism is rare relative to germline mosaicism. The latter is essentially 100% as opposed to only 0.1% for somatic mosaics (ENGELS 1979b). By analogy with these other traits, we suggest that rearrangement formation is also germline specific, although there is no direct evidence for this.

The tendency for clustering of rearrangements indicates that at least some of these events occur prior to meiosis. This observation is in good agreement with studies of other traits of the hybrid dysgenesis syndrome in which there are many lines of evidence indicating premeiotic occurrence (reviewed by ENGELS 1983). It is also possible that some of the events occur at the time of meiosis as indicated by observations (HENDERSON, WOODRUFF and THOMPSON 1978; YANNOPOLOUS 1978) of chromosomal bridges and fragments at meiotic anaphase of males from crosses which (in light of present knowledge) are likely to activate *P* factors. Our finding, that some rearrangements appear to have come about through participation by both sister strands of a hemizygous *X* chromosome, indicates that at least some rearrangement events follow the onset of the DNA synthesis phase of the cell cycle. The possibility that other rearrangements occur earlier in the cell cycle cannot be ruled out.

The mechanism of rearrangement formation: It is possible that there is no single

mechanism of *P*-induced rearrangement formation and that the collection of rearrangements that we studied represents several qualitatively different kinds of events. However, in interpreting the data, it is useful to postulate that a unified mechanism exists. The first step in considering such a mechanism is to rule out some hypotheses that seemed plausible prior to the present results. The earlier finding (ENGELS and PRESTON 1981a) that most rearrangements have breakpoints only at *P* factor sites strongly suggests homologous recombination between *P* elements as the primary mechanism. However, homologous recombination can cause only two-point rearrangements, and our finding, that most complex rearrangements do not come about by a sequence of two-point exchanges, argues against this mechanism. Furthermore, results from *in situ* hybridization show that there is no necessity for *P* elements to be conserved at the endpoints, as would be expected from the homologous exchange hypothesis. Another possibility was that rearrangements are similar to the cointegrate structures produced as transposition intermediates by prokaryotic transposable elements (reviewed by CALOS and MILLER 1980). Failure to resolve such a structure would result in an inversion if donor and recipient sites are on the same chromosome. This possibility also appears to be ruled out since it cannot account for complex rearrangements occurring by single multibreak events. Moreover, we found that sporadic breakpoints do not necessarily result in new *P* element hybridization sites, as would be expected under the cointegrate hypothesis, which necessarily involves element replication. Neither hypothesis can account for the surprising indeterminacy that we observed regarding the presence of *P* elements at breakpoints.

Based on the present data, we propose the following sequence of events leading to rearrangement formation. This is not meant to be a detailed molecular model but only a description of the major processes that are involved and their order of occurrence:

The first step is probably the determination of a set of breakage sites. This determination can be thought of as actual chromosome breaks or merely preconditions for breaks, such as single-strand nicks or the binding of an endonuclease. If DNA replication has occurred, then these breakage sites can form on both sister chromatids. We postulate this step to be the initial one because of our finding that the sites are generated by a Poisson process or something close to it. A Poisson process implies that each breakage site occurs independently of each other site and, therefore, (probably) independently of any previous step.

The most likely position for breakage is at the terminus of a *P* element. The present finding that *P* elements often vanish in the process of rearrangement formation suggests that breakage sometimes occurs at both termini. The sporadic breaks probably come about by several processes, including breakage at the site of a new *P* element generated by transposition earlier in development, spontaneous breakage of any kind and breaks generated at the target site of aberrant transposition events at the time of rearrangement formation.

The propensity to act as a breakage site might depend on whether a given element is a complete or partial *P* factor. The *P* elements at positions 5E and 17C of the π_2 *X* chromosome are both known to be complete and are both highly

active in forming chromosome rearrangements. Therefore, if such a relationship exists, it would most likely be the complete elements that are more active.

The next step is reunion of the segments delineated by these breakpoints to form a new arrangement. The array of complex rearrangements that we observe is consistent with all segments being rejoined in a single event rather than sequentially. This observation must be reconciled with the clustering of rearrangement events in germline development, which indicates that events can occur at many developmental stages. One possibility is that the breakage sites accumulate sequentially during development, but the reunion comes about as one event. We also argue from the present data that this rejoining is not necessarily precise and might involve detectable DNA duplication or deletion. The nature and extent of this imprecision are highly variable among the different rearrangement events examined. Therefore, a more complete description would necessitate the sequencing of a large number of rearrangement breakpoints.

If n breakage sites have been selected on a chromosome, there are $(n - 1)!$ ways that the segments can be ordered and 2^{n-1} ways to assign the polarities in the final arrangement. Therefore, the simplest hypothesis is that of random reunion in which all of the $(n - 1)!2^{n-1}$ possible ways of rejoining the segments, including the original sequence, are equally likely. The very large number of rearrangements examined in this study should allow detection of any significant deviation from this hypothesis. Yet, our analysis of the frequency of occurrence of each rearrangement category as well as specific breakdowns within categories to test special alternative hypotheses failed to reveal any substantial nonrandomness.

If this description is accurate, then the excision of a *P* element would be a special case of a rearrangement, one in which there was a break on each side of the element and reunion to form the original cytological sequence but without the *P* element.

The relationship between rearrangements and transposition: Rearrangement formation and *P* element transposition are both strictly under the control of the *P-M* cytotype system such that at any given time these processes are either both on or both off (ENGELS 1979a,b; ENGELS and PRESTON 1981a,b; KIDWELL 1983). We are, therefore, encouraged to consider the possibility that both are outcomes of a single molecular mechanism. Little is known about the mechanism of *P* element transposition, but preliminary transcription studies (R. KARESS, personal communication) suggest that no RNA intermediate is involved. Since both rearrangement formation and transposition necessarily involve chromosome breakage, one hypothesis consistent with our data is that rearrangements result from the random reunion of chromosome fragments generated by the transposition process.

Some preliminary results suggest that *P* element transposition is replicative as is the case with prokaryotic transposons (CALOS and MILLER 1980) as opposed to a conservative process as has been suggested for some maize elements (GREENBLATT and BRINK 1962; reviewed by FEDEROFF 1983). By *in situ* hybridization through a series of generations, we have observed at least three cases in which a new *P* element site appeared without simultaneous loss of the original one, and

no cases in which such loss did occur (W. R. ENGELS and C. R. PRESTON, unpublished results). However, these observations do not necessarily imply replicative transposition since the donor element and the recipient site might have been on different sister chromatids. Nonreplicative transposition could then result in two daughter nuclei with zero and two sites, only the latter of which was recovered. More convincing data on this point were recently obtained by W. BENZ (personal communication) whose large-scale measurements of *P* transposition and excision rates can be interpreted as indicating that overall frequency of creation of new *P* sites is much greater than the total frequency of excision. Nonreplicative transposition is still possible but only if most excisions are lost by lethality.

Our observation that *P* elements are not necessarily created or even conserved at the breakpoints of chromosome rearrangements seems to argue against element replication being an essential part of the rearrangement process. This result can be reconciled with the possibility of transposition being replicative if we assume that the chromosome breaks occurring at recipient sites are somehow different from those at donor sites such that only the latter normally participate in chromosome rearrangements. The loss of *P* elements in many rearrangements is easily accounted for in this model if the donor site breakage occurs at either or both termini of the element. This model also explains the previous observation (ENGELS and PRESTON 1981a) that rearrangement breakpoints at *P* factor sites greatly outnumber the sporadic sites.

The question of whether rearrangement formation is part of the *P* factor's transposition mechanism is of considerable interest for understanding the overall biology of these and other transposable elements. A working null hypothesis proposed by several authors to account for the presence of transposable elements is that they are merely parasitic DNA with no positive contribution to organismal fitness (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980). If so, then the only functions we would expect to find associated with such elements are those directly involved with transposition or its regulation. We would, therefore, expect rearrangment formation and other traits in the hybrid dysgenesis syndrome to be merely aberrant by-products of the machinery whose primary function is to carry out a regulated transposition process that best ensures the propagation of the element. This view can be retained in the face of the present data on rearrangement formation but only by postulating a transposition mechanism different from those that have been observed for other transposons.

DOUGLAS KELLOGG, ALICE JASCOT and SARA SMITH-SCHULL provided technical help. Valuable comments on the manuscript came from WENDY BENZ, JAMES CROW, MARGARET KIDWELL, JOHNG LIM, MUNEO MATSUDA, KEVIN O'HARE and MICHAEL SIMMONS. This is paper 2716 from the Laboratory of Genetics. It was supported by grants from the National Institutes of Health (GM 30948) and the National Science Foundation (PCM8104332).

LITERATURE CITED

BERG, R. L., W. R. ENGELS and R. A. KREBER, 1980 Site-specific X-chromosome rearrangements from hybrid dysgenesis in *Drosophila melanogaster*. Science **210:** 427–429.

BINGHAM, P. M., 1981 A novel dominant mutant allele at the *white* locus of *Drosophila melanogaster* is mutable. Cold Spring Harbor Symp. Quant. Biol. **45:** 519–525.

BINGHAM, P. M., M. G. KIDWELL and G. M. RUBIN, 1982 The molecular basis of P-M hybrid dysgenesis: the role of the P element, a P strain-specific transposon family. Cell **29:** 995–1004.

BINGHAM, P. M., R. LEVIS and G. M. RUBIN, 1981 The cloning of the DNA sequences from the *white* locus of *Drosophila melanogaster* using a novel and general method. Cell **25:** 693–704.

BREGLIANO, J. C. and M. G. KIDWELL, 1983 Hybrid dysgenesis determinants. pp. 363–410. In: *Mobile Genetic Elements,* Edited by J. A. SHAPIRO. Academic Press, New York.

CALOS, M. P. and M. J. MILLER, 1980 Transposable elements. Cell **20:** 579–595.

DOOLITTLE, W. F. and C. SAPIENZA, 1980 Selfish genes, the phenotype paradigm and genome evolution. Nature **284:** 601–603.

ENGELS, W. R., 1979a Extrachromosomal control of mutability in *Drosophila melanogaster* Proc. Natl. Acad. Sci. USA **76:** 4011–4015.

ENGELS, W. R., 1979b Hybrid dysgenesis in *Drosophila melanogaster*: rules of inheritance of female sterility. Genet. Res. **33:** 219–23.

ENGELS, W. R., 1981 Germline hypermutability in Drosophila and its relation to hybrid dysgenesis and cytotype. Genetics. **98:** 565–587.

ENGELS, W. R., 1983 The P family of transposable elements in Drosophila. Annu. Rev. Genet. **17:** 315–344.

ENGELS, W. R. and C. R. PRESTON, 1980 Components of hybrid dysgenesis in a wild population of *Drosophila melanogaster*. Genetics **95:** 111–128.

ENGELS, W. R. and C. R. PRESTON, 1981a Characteristics of a "neutral" strain in the P-M system of hybrid dysgenesis. Drosophila Inform. Serv. **56:** 35–37.

ENGELS, W. R. and C. R. PRESTON, 1981b Identifying P factors in *Drosophila* by means of chromosome breakage hotspots. Cell **26:** 421–428.

FANO, U., 1941 On the analysis and interpretation of chromosomal changes in Drosophila. Cold Spring Harbor Symp. Quant. Biol. **9:** 113–121.

FEDEROFF, N., 1983 Controlling elements in maize. pp. 1–65. In: *Mobile Genetic Elements,* Edited by J. A. SHAPIRO. Academic Press, New York.

FINCHAM, J. R. S. and G. R. K. SASTRY, 1974 Controlling elements in maize. Annu. Rev. Genet. **8:** 15–50.

GREENBLATT, I. M. and R. A. BRINK, 1962 Twin mutations in medium variegated pericarp maize. Genetics **47:** 489–501.

GRUNBERG, H., 1936 A case of complete reversion of a chromosomal rearrangement in *Drosophila melanogaster*. Nature **138:** 508.

HENDERSON, S. A., R. C. WOODRUFF and J. N. THOMPSON, 1978 Spontaneous chromosome breakage at male meiosis associated with male recombination in *Drosophila melanogaster*. Genetics **88:** 93–107.

HINTON, C. W., 1979 Two mutators and their suppressors in *Drosophila ananassae*. Genetics **92:** 1153–1171.

KALISCH, W., 1970 Über eine mutable *white*-inversion bei *Drosophila melanogaster* (English abstr.). Mol. Gen. Genet. **107:** 336–350.

KAUFMANN, B. P., 1941 Induced chromosome breaks in Drosophila. Cold Spring Harbor Symp. Quant. Biol. **9:** 82–92.

KIDWELL, M. G., 1983 Hybrid dysgenesis in *Drosophila melanogaster*: factors affecting chromosomal contamination in the *P-M* system. Genetics **104:** 317–341.

LEA, D. E., 1962 *Actions of Radiations on Living Cells,* Ed. 2. Cambridge University Press, London.

LEFEVRE, G., 1976 A photographic representation and interpretation of the polytene chromo-

somes of *D. melanogaster* salivary glands. pp. 31–66. In: *The Genetics and Biology of Drosophila*, Vol. 1a, Edited by M. ASHBURNER and E. NOVITSKI. Academic Press, London.

LEVIS, R., M. COLLINS and G. M. RUBIN, 1982 FB elements are the common basis for the instability of the w^{DZL} and w^{c} Drosophila mutations. Cell **30:** 551–565.

LEVITAN, M., 1963 A maternal factor which breaks paternal chromosomes. Nature **200:** 437–438.

LIM, J. K., 1979 Site-specific instability in *Drosophila melanogaster*: the origin of the mutation and cytogenetic evidence for site specificity. Genetics **93:** 681–701.

LIM, J. K., 1981 Site-specific intrachromosomal rearrangements in *Drosophila melanogaster*, cytogenetic evidence for transposable elements. Cold Spring Harbor Symp. Quant. Biol. **45:** 553–560.

LINDSLEY, D. L. and E. H. GRELL, 1968 *Genetic Variations of Drosophila melanogaster*. Carnegie Inst. Wash. Publ. 6.

LINDSLEY, D. L. and K. T. TOKUYASU, 1980 Spermatogenesis. pp. 225–294. In: *Genetics and Biology of Drosophila*, Vol. 2d, Edited by M. ASHBURNER and T. R. F. WRIGHT. Academic Press, London.

NOVITSKI, E., 1961 The regular reinversion of the roughest^{3} inversion. Genetics **46:** 711–718.

O'HARE, K. and G. M. RUBIN, 1983 Structures of P transposable elements of *Drosophila melanogaster* and their sites of insertion and excision. Cell **34:** 25–35.

ORGEL, L. E. and F. H. C. CRICK, 1980 Selfish DNA: the ultimate parasite. Nature **284:** 604–606.

PAINTER, T. S. and H. J. MULLER, 1929 Parallel cytology and Genetics of induced translocations and deletions in Drosophila. J. Hered. **20:** 287–298.

PARDUE, M. L. and J. GALL, 1975 Nucleic and hybridization to the DNA of cytological preparations. Methods Cell Biol. **1:** 1–17.

RUBIN, G. M., 1983 Dispersed repetitive DNAs in Drosophila. pp. 329–362. In: *Mobile Genetic Elements*, Edited by J. A. SHAPIRO, Academic Press, New York.

RUBIN, G. M., M. G. KIDWELL and P. M. BINGHAM, 1982 The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. Cell **29:** 987–994.

SIMMONS, M. J. and J. K. LIM, 1980 Site specificity of mutations arising in dysgenic hybrids of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **77:** 6042–6046.

SPRADLING, A. C. and G. M. RUBIN, 1982 Transposition of cloned P elements into Drosophila germ line chromosomes. Science **218:** 341–347.

YAMAGUCHI, O. and T. MUKAI, 1974 Variation of spontaneous occurrence rates of chromosomal aberrations in the second chromosomes of *Drosophila melanogaster*. Genetics **78:** 1209–1221.

YANNOPOULOS, G., 1978 Studies on male recombination in a Southern Greek *Drosophila melanogaster* population. (c) Chromosomal abnormalities at male meiosis. (d) Cytoplasmic factor responsible for the reciprocal cross effect. Genet. Res. **31:** 187–196.

YANNOPOULOS, G., A. ZACHAROPOULOU and N. STAMATIS, 1982 Unstable chromosome rearrangements associated with male recombination in *Drosophila melanogaster*. Mutat. Res. **96:** 41–51.

Corresponding editor: A. SPRADLING

APPENDIX: THE EXPECTED NUMBER OF REARRANGEMENT TYPES

If we assume that the number of nonselected breakpoints is Poisson distributed with parameter x and that the probability of the selected breakpoint is c, then, the frequency of k-point rearrangements in the sample will be

$$c\frac{e^{-x}x^{k-1}}{(k-1)!}$$

However, this includes the unobservable case of $k = 1$. The actual frequency will, therefore, be the conditional probability obtained by dividing the above quantity by $c(1 - e^{-x})$ which is the probability that k is at least 2. Therefore,

$$P_k = \frac{e^{-x}x^{k-1}}{(1-e^{x})(k-1)!} \tag{1}$$

Note that this conditional probability is independent of the parameter c, allowing us to obtain the maximum likelihood estimate of x simply by maximizing the quantity

$$L = \sum_{k=2}^{\infty} n^k \ln(P_k), \tag{2}$$

where n_k is the observed number of rearrangements with k breakpoints. This quantity is monotonic with the multinomial probability of observing exactly n_k chromosomes of the k-point type for all k.

Alternatively, if we assume that the actual breakage events themselves (not counting the selected site) are Poisson distributed, and that they reassemble at random, then a k-point rearrangement can come about through any number $j \geq k - 1$ of breaks. The unconditional probability of a k-point rearrangement including the selected site is therefore

$$Pr(k\text{-}point) = c\sum_{j=k-1}^{\infty}\frac{e^{-x}x^{j}}{j!}T_{jk} \tag{3}$$

where T_{jk} is the probability that a chromosome broken in the selected site plus j nonselected ones will assemble into a k-point rearrangement including the selected point. This probability can be written as

$$T_{jk} = \frac{\binom{j}{k-1}R_k}{j!\,2^{j}}. \tag{4}$$

The denominator is the total number of ways that the chromosome can reassemble, including the original chromosome, and the numerator is the number of the required type. R_k is the number of distinct rearrangements involving exactly k breakpoints. We are not aware of a closed expression for these quantities, but they can be readily calculated recursively as

$$R_k = (k-1)!\,2^{k-1} - \sum_{i=0}^{k-1} R_i\binom{k}{i}. \quad (R_0 = 1) \tag{5}$$

The first term is the total number of ways to assemble a chromosome broken in k points, and the second term subtracts all such assemblies that have fewer than k breakpoints and, therefore, should not be counted. The last term in this series, therefore, represents the original chromosome. The unconditional probability of a k-point rearrangement can now be computed by combining equations (3), (4) and (5).

Finally, we obtain the conditional probability as in (1) by dividing by $c(1 - e^{-x})$. The result is again independent of c and allows numerical maximization of the multinomial likelihood as in (2) to obtain the estimate of x. This value is then used to calculate the expectations such as in Table 1.