PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

On the Evolution and Population Genetics of Hybrid-Dysgenesis-Causing Transposable Elements in Drosophila

W. R. Engels

Phil. Trans. R. Soc. Lond. B 1986 **312**, 205-215
doi: 10.1098/rstb.1986.0002

Phil. Trans. R. Soc. Lond. B **312**, 205–215 (1986)
Printed in Great Britain

On the evolution and population genetics of hybrid-dysgenesis-causing transposable elements in *Drosophila*

By W. R. Engels

Genetics Department, University of Wisconsin, Madison, Wisconsin 53706, *U.S.A.*

Much has been learned about transposable genetic elements in *Drosophila*, but questions still remain, especially concerning their evolutionary significance. Three such questions are considered here. (i) Has the behaviour of transposable elements been most influenced by natural selection at the level of the organism, the population, or the elements themselves? (ii) How did the elements originate in the genome of the species? (iii) Why are laboratory stocks different from natural populations with respect to their transposable element composition? No final answers to these questions are yet available, but by focusing on the two families of hybrid dysgenesis-causing elements, the P and I factors, we can draw some tentative conclusions.

Transposable elements are now known to be a significant component of many eukaryotic genomes. There have been extensive studies on the genetics and molecular structure of these elements, and some work on their behaviour (see review articles in Shapiro (1983). However, the fundamental questions of how these elements evolved and what their significance might be for the biology of the organism remain unanswered.

In this article, I will discuss three of these questions and consider the available evidence in terms of the two families of hybrid-dysgenesis-causing transposable elements in *Drosophila melanogaster*.

Introduction: the hybrid-dysgenesis-causing elements

The P and I factors are two families of transposable elements in *Drosophila* noted for their ability to cause syndromes of germ-line abnormalities known as hybrid dysgenesis (Kidwell *et al.* 1977; Engels 1983; Bregliano & Kidwell 1983). These abnormalities, which include temperature-sensitive sterility and elevated rates of mutations and chromosome rearrangements, are thought to be by-products of very high transposition rates.

The distinctive feature of P and I elements is the regulation of their transpositional activity. When a genome has only one or a small number of autonomously transposing elements, transposition and excision events proceed rapidly. This condition, hybrid dysgenesis, is ultimately unstable. It sometimes results in the loss of all autonomous elements. Alternatively, the number might increase until there are sufficiently many to bring about a stable cellular state or cytotype. Transposition in this condition is greatly reduced but not completely absent (Preston & Engels 1984). This stable condition is known as the P cytotype for P elements (Engels 1979), and the inducer state for I elements (Bucheton & Picard 1978).

The physical nature of this regulation is not known, but it appears to be jointly determined by the number of elements in an individual's genome and by the cytotype of the individual's mother (Bucheton & Picard 1978; Engels 1979; Kidwell 1981; Simmons & Bucholz 1985).

Thus, part of the inheritance of cytotype is through the maternal line, suggesting an extrachromosomal component.

This type of regulation along with the long-term instability of the dysgenic condition means that a given strain or population will eventually reach one of the two equilibrium states: either the number of autonomous elements will be zero or it will be large enough for the stable cytotype to prevail. These are called M strains and P strains, respectively, for the P family, and R strains and I strains for the I family. Since the P and I cytotype systems function independently, strains can be doubly classified as IP, RM, etc. (Kidwell 1979). Note that the extrachromosomal component of cytotype determination results in dysgenesis in the hybrid progeny of M♀ × P♂ crosses and R♀ × I♂ crosses. Dysgenesis is absent or greatly reduced in hybrids from the reciprocal crosses.

In addition to regulation by cytotype, both these families of elements display tissue-specific regulation that results in transposition being much reduced in somatic cells relative to the germ-line.

Each of these two families can be further subdivided into autonomous and non-autonomous elements (O'Hare & Rubin 1983; Bucheton *et al.* 1984). In the case of the P family, the non-autonomous elements are mostly derived from the autonomous ones by internal deletions. They can transpose and excise, but only if at least one autonomous element is present (Spradling & Rubin 1982; Engels 1984).

Some strains behave as though they have only non-autonomous elements, and lack the stable cytotype (Bingham *et al.* 1982; Bucheton *et al.* 1984) as indicated by the dysgenesis observed in the progeny when these females are crossed to males with autonomous elements. All R strains appear to be of this type; they possess only partly deleted I elements which are present mainly in the heterochromatin (Bucheton *et al.* 1984). In the P–M system, such strains are sometimes called M′ or pseudo-M strains to distinguish them from true M strains which have no sequences homologous to P elements.

Despite these remarkable similarities, the P and I element families have little else in common. (i) There is no obvious sequence homology (Bingham *et al.* 1982; Bucheton *et al.* 1984), and the two elements differ in their structural organization: I elements have no repeats, whereas P elements have inverted 31-base repeats at their termini (O'Hare & Rubin 1983; Bucheton *et al.* 1984; Sang *et al.* 1985). (ii) Transpositional activity of I factors occurs only in females, whereas P elements transpose and excise in both sexes. (iii) Both elements cause sterility, but the details, such as developmental timing, are not the same (Engels & Preston 1979; Schaefer *et al.* 1979; Picard *et al.* 1977; Kidwell 1979). (iv) both P and I factors have partial site-specificity of insertional target sites (Simmons & Lim 1980; O'Hare & Rubin 1983; Engels 1983; Sang *et al.* 1985) but the favoured chromosomal positions differ. (v) Finally, and most importantly for this discussion, the two elements differ in their distributions among strains and species (see below), indicating different evolutionary histories.

Does natural selection act mainly on the organism, the population, or the transposable elements themselves?

For Mendelian genes, natural selection can be thought of as acting simultaneously on individuals and populations. Sometimes these two selective modes tend to move gene frequencies in opposite directions, in which case the long-term evolutionary consequences

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

depend on which selective mode predominates. The situation becomes even more complicated when we consider transposable elements. Since these elements undergo replication that is partly independent of the rest of the genome, it is necessary to consider a third mode of selection which acts at the level of the elements themselves. A major question, therefore, is to what extent each of these three modes of natural selection has been important in determining the current behaviour of transposable elements. To approach this question, we can consider the likely outcome of selection on each level, and compare it with the observed behaviour of P and I factors.

(*a*) *Selection on organisms*

Several observations suggest that selection at the level of the organism has not been the primary force. All traits associated with the hybrid dysgenesis syndrome appear to be deleterious to *Drosophila*. Sterility has the most obvious effect on fitness, although its temperature sensitivity makes it difficult for us to determine how important it is in natural populations. In the P–M system, gonadal sterility only occurs when the late embryonic and early larval stages are exposed to elevated temperatures (25 °C or greater). In the I–R system, temperature is important at all stages, but there is also an age effect such that fertility approaches normal levels in older adults. Mutations, some of which are caused by insertions and others by imprecise excisions, and chromosome rearrangements will also contribute to lowering the fitness of dysgenic flies.

Thus the only way these elements are likely to enjoy a selective advantage at the organismal level is by conferring some kind of benefit to the members of non-dysgenic P and I strains relative to their M and R counterparts. So far, no such advantage has been observed although no accurate quantitative comparisons have been made. They could not, for example, function as controlling elements on the level of gene expression, since their chromosomal positions vary widely even within a given P or I strain. At least qualitatively, RM strains appear very similar to IP strains. In fact, the only way to classify strains in the P–M and I–R systems is by examining the hybrids of crosses with standard reference strains, or by molecular methods using P or I element nucleic acid probes. Of course, it is still possible that there are more subtle fitness advantages to having P and I elements, especially if the advantage only appears under special conditions. For example, by analogy with bacterial transposons that code for antibiotic resistance, we might postulate that P and I elements could provide resistance to environmental hazards not present in laboratory tests. Until such an advantage is found, however, there is no evidence that organismal selection has been important in determining the behaviour of the P and I families of elements.

(*b*) *Selection on populations*

Another possibility is that selection at the population level is mainly responsible for these elements. A view often expressed (see, for example, Thompson & Woodruff 1978) is that since these elements cause mutations, they could give a population an advantage by speeding the rate of evolution. In support of this notion, one could cite the growing evidence (Rubin 1983) that a large fraction of spontaneous mutations in *Drosophila* are insertions of transposons. These insertional mutations can result in phenotypes different from chemical- or radiation-induced mutations. However, there are some serious weaknesses in this argument. First, it is by no means clear that increasing the mutation rate would cause evolution to proceed at a faster pace. Instead, standing variability in the population is likely to be sufficient for natural selection to occur at rates in excess of those needed for most evolutionary changes.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

A stronger argument against some populational advantage arises from the finding that all known transposable elements in *D. melanogaster* are highly polymorphic in their genomic positions. Although some heterochromatic I element sites appear to be fixed (Bucheton *et al.* 1984) these are unlikely to affect gene expression. For the euchromatic part of the genome, there is no case where a given element site is fixed, or even exists in appreciable frequencies in the species. The highest observed frequency is one case of the element *297* being present in the same *in situ* hybridization site in 5 out of 20 X chromosomes from a North Carolina population (Montgomery & Langley 1983). P element distributions show an even greater degree of polymorphism; every chromosomal site appears unique or at a very low frequency in a Madison population (W. Engels and C. Preston, unpublished). Since the cytotype of this population is P (Engels & Preston 1980) the observed variability is presumably the result of background insertions and excisions that occur even in the P cytotype (Preston & Engels 1984). This background activity, though low relative to rates in the M cytotype, is still significant. Therefore, from these observed levels of polymorphism, we can infer that mutations caused by transposable element insertions are not being fixed in natural populations, and are thus not important for evolution. Of course, this argument applies only to insertion mutations; it is still possible that mutations from imprecise excisions or dysgenesis-induced chromosome rearrangements do sometimes get fixed.

(*c*) *Selection on transposable elements*

Most evidence suggests that selection on the level of the elements themselves has had the greatest influence on the evolution of P and I factors. This type of selection would result in elements that behave as parasitic or 'selfish' DNA (Doolittle & Sapienza 1980; Orgel & Crick 1980). We would expect under this hypothesis that the elements would be well adapted for self-sufficient transposition with as little help from cellular genes as possible. Such self sufficiency would minimize the elements' vulnerability to the cell's defences. This expectation is borne out by the finding that P elements carry a gene for their own transposition (Spradling & Rubin 1982; Engels 1984). Moreover, the *in vitro* mutagenesis studies of Karess & Rubin (1984) strongly suggest that this transposase is the only cistron carried on the autonomous P factor. We can also infer that I factors carry at least one function necessary for their own transposition from the observation (Bucheton *et al.* 1984) that defective I elements are present in the genomes of R strains. Since no hybrid dysgenesis occurs in these genomes, the autonomous I elements must be providing a necessary function for transposition.

Another prediction from the parasitic DNA hypothesis is that any deleterious effects on the host organism are the necessary consequences of transposition. A well-adapted parasite will not damage its host more than necessary. This appears to be the case for dysgenesis-induced mutations, many of which are the direct result of insertion events. Other mutations come about by imprecise excisions or through chromosome rearrangements. There is some evidence that these two processes are also the necessary by-products of transposition, at least for P elements. Both processes are suppressed by the P cytotype, and neither occurs in the absence of autonomous (transposase-producing) P elements. Significantly, these processes are almost entirely limited to the germ-line, where transposition will be helpful for the spread of the element, but where minimum damage will be done to the host.

It is less clear that sterility is a necessary consequence of high transposition rates, but the very close correlation between the frequency of gonadal sterility caused by P factors and P

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

element excision as measured by mutability of the sn^w allele (Engels 1984) suggests that this is the case. One possibility is that gonadal sterility is an indirect consequence of the presence of transposase rather than a result of actual transposition events.

(*d*) *Conclusions*

The behaviour of P and I factors appears to be closest to that predicted for parasitic elements. Although all three modes of selection have undoubtedly been operating, the behaviour of these elements argues that the most important mode of selection has been on the level of the elements themselves rather than on the organismal or population levels.

The origin of P and I elements in *D. melanogaster*

(*a*) *Origin of I factors*

Recent work by A. Bucheton and D. Finnegan (cited by Simmons & Karess 1985) indicates that transposable elements sharing significant homology with I factors exist in the sibling species *D. simulans* and *D. mauritiana*. They were also found in other species of the melanogaster subgroup: *D. tessieri*, *D. erecta* and *D. yakuba*. By restriction mapping, the I elements in *D. simulans* appeared to be complete. Thus it seems likely that these elements have been present in *D. melanogaster* since before the divergence of these species. It is not yet known how much sequence divergence has occurred between the elements found in *D. melanogaster* and those of other species, and to what extent each species' elements are specifically adapted to their hosts.

(*b*) *Origin of P factors*

The story for P elements is very different. There is no trace of P sequence homology in any of the sibling species of *D. melanogaster*, or in any of the other most closely related species (Brookfield *et al.* 1984; Daniels *et al.* 1984). Yet P-like elements are common in four of five species examined in the more distantly related *willistoni* group. This species group, which is in the *Sophophora* subgenus as is the *melanogaster* group, is found mainly in South America and Mexico (Daniels *et al.* 1984; S. Daniels and L. Strausbaugh, personal communication). From restriction map data, Daniels *et al.* (1984) showed that the P elements in the *willistoni* group are much less diverged from the P elements in *D. melanogaster* than are other sequences in these two species, including such highly conserved genes as the histones. The best explanation seems to be that at some point since the divergence of the *melanogaster* sibling species (roughly, the last million years) there was an interspecific transfer of P element sequences from *willistoni* or another group to *D. melanogaster*. Transfer in the other direction is less likely since the presence of P-like elements in most of the *willistoni* group species, which are not interfertile, indicates a much older origin in these species. The transfer to *D. melanogaster* was apparently followed by the invasion and spread of the elements with some internal deletions occurring sometimes to yield non-autonomous elements.

In support of this hypothesis, there is recent evidence that when autonomous P elements are injected into the embryos of other species ranging from *D. simulans* to *D. hawaiiensis*, a very distant relative, these elements can function sufficiently well to transpose from the injected plasmid vector and integrate into the new host genome. Furthermore, once in the foreign genome, they have been seen to undergo transposition, excision and the production of chromosome rearrangements much as they do in *melanogaster* (Brennan *et al.* 1984; Scavarda

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

& Hartl 1984; Daniels *et al.* 1985). Thus, if P elements are introduced into a new species, we might expect them to be invasive. It is not yet known what the long-term outcome of these interspecific transfer experiments will be, or whether P elements will retain their capacity for regulation by cytotype in their new hosts.

If we accept the hypothesis of interspecific transfer, questions still remain concerning how the initial horizontal transfer into *D. melanogaster* might have occurred, and why the same thing apparently did not happen to *D. simulans*, since these two species coexist worldwide.

A clue to the nature of the initial introduction event is provided by the work of Miller & Miller (1982). While collecting mutations in the genome of a large (*ca.* 100 kilobases) DNA virus that infects certain lepidopteran species, they found a case in which a transposable element common in the host insect genome had inserted into the genome of the virus. The element was typical of the *copia*-like (sometimes called 'retrovirus-like' or 'retro-transposon') class of transposable elements which is very common in *Drosophila*. This integration did not prevent the DNA virus from carrying out its normal life cycle. Thus, we can postulate that by integrating into virus genomes, transposable elements can sometimes be transferred from one species to another provided both species are included in the virus' host range. To explain the situation in *Drosophila*, we must further postulate that transfer events of this kind are sufficiently rare so that no successful transfer of P elements has yet occurred into *D. simulans* from any P-element-bearing species.

Finally the above evidence favouring an invasion by P factors within the last million years should be distinguished from Kidwell's (1979, 1983) hypothesis that they invaded within the last 35 years. The latter will be discussed below.

The origin of the current distribution of P and I elements among *Drosophila melanogaster* strains and populations

One of the earliest and most dramatic findings related to hybrid dysgenesis was that very old laboratory stocks dating from the 1920s were invariably of the type RM. That is, their genomes were lacking in both I factors and P factors. By contrast, all strains recently taken from natural populations had I factors and P elements, at least of the non-autonomous kind. Recently derived strains from areas near the Mediterranean or from some populations in Asia or Australia had the M cytotype, and were therefore thought to lack autonomous P elements. Since these strains still had genomic P elements, they were classified as IM′. If they were derived from other areas, they had the P cytotype and were IP. Laboratory strains of intermediate age were of the type IM: they had complete I factors but no homology to P elements. The remaining classification, RP, has not been found in laboratory or wild populations (Picard *et al.* 1976; Kidwell *et al.* 1977; Kidwell 1983; Anxolabéhère *et al.* 1984).

The correlation between age of laboratory stocks and the presence or absence of these two transposon families poses an intriguing problem. It seems obvious that one of two things must have happened: either natural populations have recently acquired I factors and P factors (in that order), or else laboratory stocks have lost them (in the opposite order). However, as I will discuss below, there are powerful arguments against both of these possibilities.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

(*a*) *Recent loss of P and I factors from laboratory stocks*

Since the establishment of a laboratory stock results in drastic changes in all aspects of the captured organism's ecology, it would seem more reasonable that these stocks, rather than the global population, would undergo a rapid change in their transposable element composition. The primary difficulty with this view is the lack of any plausible mechanism for ridding the genome of all its P or I factors. The only such mechanism that has been proposed is that of random drift in small laboratory populations (Engels 1981*a*). This possibility, however, has been largely eliminated by molecular studies indicating that the number of P elements in the genome, 30–50 (Bingham *et al.* 1982), is larger than previously thought. Thus the expected time until stochastic loss of all elements even in small populations would be exceedingly long (see, for example, Charlesworth & Charlesworth (1983) and Kaplan *et al.* (1985) for theoretical analyses of stochastic models of transposable elements with some properties analogous to P and I factors, but with a regulation system different from cytotype). In addition, the majority of P elements were found to be of the non-autonomous type (O'Hare & Rubin 1983). This implied that even if sufficient time were allowed for stochastic loss, the likely result would be a pseudo-M population with a large number of inert, defective elements present in the genome. Once positions of defective elements were fixed, no further changes would occur, since the transposase function is required for excision (Karess & Rubin 1984). Thus, there seems no way under this hypothesis to account for the existence of true M strains among old laboratory stocks.

These difficulties are somewhat reduced in the I–R system, where there is nothing analogous to true M strains. However, a stochastic loss model would still predict the presence of some defective I elements in the euchromatic arms of R strains, rather than just in the heterochromatin, as has been observed (Bucheton *et al.* 1984).

(*b*) *Recent invasion of I and P factors in natural populations*

To overcome these difficulties with the recent loss hypothesis, Kidwell (1979, 1983) proposed that neither P nor I elements were present in *D. melanogaster* before 1920. According to this model 'populations are often invaded by (presumably parasitic) elements of which the P element is an example' (Bingham *et al.* 1982). The invasion by I factors is postulated to have taken place during the period of 1930–60, and P factors began their spread around 1950 (Kidwell 1983).

One immediate problem with this hypothesis is that it implies an unreasonably high rate of acquisition of new transposon families over evolutionary time. At the current rate of two invasions per century, *D. melanogaster* would have gained at least 20000 families of transposable elements during the estimated minimum of one million years since its divergence from *D. simulans*. This is much too large a number since the entire genome is estimated to contain only 20–50 transposable element families (Rubin 1983; Engels 1981*b*). One counter-argument is that many elements might be gained and subsequently lost from the genome. If so, however, then there is no longer an advantage to the recent invasion model over the loss model. A mechanism for ridding the genome of all its elements of a given family must still be found. In fact, it is much harder to imagine a mechanism applying to the entire species as opposed to one that merely removes transposons from laboratory strains. Another possible counter-argument is that the available sample size (two element invasions in one century) is too small to estimate accurately the true rate of such takeovers. However, if we assume that invasions follow a Poisson

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

distribution, we can still say with 95% confidence that the rate is at least 0.355 per century, and the number of elements gained by the species since divergence from *D. simulans* is at least 3550. Since this number is still unreasonably large, we must conclude that either the invasions of P and I elements are not typical events, possibly being an indirect result of human activity in the last century, or that the recent invasion model is incorrect.

Another objection to the recent invasion hypothesis, perhaps even more serious, comes from the finding (Bucheton *et al.* 1984) that all R strains have many defective I elements. Each such element represents a different subset of the complete 5.4 kilobase I factor found only in I strains. Together, the defective elements in a typical R strain are thought to have all the information contained in complete I factors. It is difficult to imagine any way to produce such a collection of defective elements unless complete I factors pre-existed in the R strain genomes. Furthermore, the finding that transposons closely related to I factors are present in the genomes of all other species in the *melanogaster* subgroup (A. Bucheton and D. Finnegan, cited by Simmons & Karess 1985) strongly argues that these elements have been present in the genome since long before the time postulated for the recent invasion event.

We might attempt to rescue the recent invasion hypothesis by supposing that complete I factors existed long ago, were subsequently lost except for some defective remnants, and now are in the last stages of a second invasion. This reinvasion might have begun by either an interspecific transfer of an exogenous I element, or by somehow reconstructing a complete I element from defective ones. This sequence of events, however, requires a loss *and* an invasion, and is therefore one step less parsimonious than a model that requires only a loss of elements. It suffers from many of the problems (above) of the recent loss model, plus the additional objections to recent invasion.

It appears, therefore, that both models for the I–R system require that R strains were derived by the loss of complete I factors which had been present in the species for a very long time. Since *D. simulans* appears to have complete I elements, this loss is most likely to have happened within the last million years. The difference between the two hypotheses is that in the recent loss model, we postulate that this loss occurred rapidly in very small populations, whereas in the recent invasion hypothesis, it is assumed to have happened gradually but in much larger populations. The rapidity of the hypothesized loss and the large size of natural populations (respectively) make random drift insufficient to account for the loss under either hypothesis. Instead, some new mechanism must be found to rid the genome of complete I elements, and this mechanism must function in very short times (65 years) or in large populations (of the order of 10^6) or both.

Finally, we can ask whether the recent invasion model, which now seems implausible for the I–R system, might still work for the P–M system. If so, however, it would seem a remarkable coincidence that both kinds of elements would, for two entirely different reasons, have the unexpected property of being missing from old laboratory strains but common in nature. Furthermore, the recent invasion model does not explain the existence of pseudo-M strains in many parts of the world without including an additional postulate of even more recent loss. As argued above in the case of defective I elements, the defective P elements founds in pseudo-M populations, assuming these are similar to the pseudo-M strains that have been examined in the laboratory, must have been derived from pre-existing complete P factors. Therefore, some mechanism must still be found to rid the genome of all complete P elements in a large population. Since the primary motivation for the recent invasion model is to avoid the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

difficulties of recent loss, it would not seem useful to propose a model that requires both recent invasion and recent loss.

(*c*) *Conclusions and speculations*

The above discussion, which seems to lead to the paradoxical conclusion that all possible explanations should be discarded, can be taken to imply that our understanding of the biology of these elements is critically deficient in some way. The various difficulties could be solved if, for example, the *Drosophila* genome were found to code for some enzymic mechanism which, under conditions not yet studied experimentally, could specifically excise P and I elements.

Models more complicated than the recent loss and recent invasion models discussed above can also be considered. One suggestion (Engels 1983) is that both P and I factors were in existence long before any laboratory stocks were established, but only in relatively isolated parts of the world. Early samples taken for laboratory use might well have excluded these areas purely by chance. With the greatly enhanced migration rates of *Drosophila* brought about by human activity in this century, these elements might have been better able to spread to previously uninfected populations. In the case of I factors, we can imagine the global population being at a quasi-equilibrium such as in the theoretical analysis by Kaplan *et al.* (1985) in which most elements are defective, but with a few complete I factors present in some populations. The large-scale demographic changes in *Drosophila* populations during the last 80 years might have shifted the position of the quasi-equilibrium such that complete I factors are now much more prevalent.

Other *Drosophila* transposons

To what extent do the preceding discussions apply to other families of *Drosophila* transposable elements? Two other important classes of elements are the *copia*-like (or retrovirus-like) families and the FB ('foldback') elements. These elements have repeat structures (Rubin 1983) that are very different from P and I elements. These differences might indicate different evolutionary histories. In the case of the *copia*-like elements, there is much evidence (Shiba & Saigo 1983; Kugimiya *et al.* 1983; Flavell 1984) that they share a recent common ancestor with retroviruses. There is no indication of such a relationship for P or I elements.

Some of the arguments concerning the mode of selection can be applied to other transposons. In particular, the absence of monomorphic chromosomal sites for any of the transposable elements studied to date strongly suggests that insertion mutations in general have not played an important role in evolution since such mutations are apparently not being fixed.

The postulated interspecific transfer of P factors from the *willistoni* species group to *D. melanogaster* is probably not typical of *Drosophila* transposons. In a genus-wide survey with six transposable elements, Martin *et al.* (1983) found that in most cases the distributions of these elements followed the accepted phylogeny reasonably well. There were, however, some cases that could best be explained by interspecific transfer events, or else by rapid evolution of sequences to the point where hybridization is no longer detectable.

The question of how laboratory stocks came to differ from natural populations does not arise for other transposons studied to date. Other than P and I factors, all *Drosophila* transposable elements appear to be ubiquitous in the species. The reason might be that P and I elements differ from other transposons in their behaviour (the recent loss hypothesis) or in their time of arrival in the species (the recent invasion hypothesis).

The characteristic property of a transposable element is its ability to create new copies of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

itself in the genome. Any DNA sequence with this property can potentially spread through a population without the aid of selection at the organismal level, or even despite mild selection against it. In this respect, transposable sequences are analogous to meiotic drive systems (Crow 1979). There might be many evolutionary routes to such a sequence, as indicated by the diversity of structural types among the various families of *Drosophila* transposons. In addition, the mechanisms of transposition itself appear to be variable; *copia*-like elements are thought to transpose by means of reverse transcription of an RNA intermediate (Flavell 1984), whereas P element transposition is probably by a different mechanism (Karess & Rubin 1984). There are good reasons to suspect convergent evolution for the regulatory mechanisms of P and I elements. As mentioned earlier, these two elements differ in every structural detail, yet they share a very unusual mode of inheritance of cytotype.

With the likelihood of a multiplicity of evolutionary pathways all leading to transposable elements, an equally diverse set of approaches to studying them will almost certainly be needed. No single family of elements and no single set of techniques, whether classical or molecular, will be sufficient. However, with a co-ordinated approach, an understanding of the evolution of transposable elements and their role in the biology of higher organisms might soon be within reach.

I thank Christine R. Preston, Wendy Benz and Michael Simmons for many useful comments on the manuscript. This work was supported by the National Institutes of Health (GM30948). It is paper number 2807 from the Laboratory of Genetics, University of Wisconsin.

References

Anxolabéhère, D., Kai, H., Nouaud, D., Périquet, G. & Ronsseray, S. 1984 The geographical distribution of P–M hybrid dysgenesis in *Drosophila melanogaster*. *Génét. Sél. Evol.* **16**, 15–26.

Bingham, P. M., Kidwell, M. G. & Rubin, G. M. 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.

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

Brennan, M. D., Rowan, R. G. & Dickinson, W. J. 1984 Introduction of a functional P element into the germ line of *Drosophila hawaiiensis*. *Cell* **38**, 147–151.

Brookfield, J. F. Y., Montgomery, E. & Langley, C. H. 1984 Apparent absence of transposable elements related to the P elements of *D. melanogaster* in other species of *Drosophila*. *Nature, Lond.* **310**, 330–332.

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

Bucheton, A. & Picard, G. 1978 Non-Mendelian female sterility in *Drosophila melanogaster*: hereditary transmission of reactivity levels. *Heredity* **40**, 207–223.

Charlesworth, B. & Charlesworth, D. 1983 The population dynamics of transposable elements. *Genet. Res., Camb.* **42**, 1–27.

Crow, J. F. 1979 Genes that violate Mendel's rules. *Scient. Am.* **240**, 134–146.

Daniels, S. B., Strausbaugh, L. D. & Armstrong, R. A. 1985 Behavior of a P transposable element in *Drosophila simulans* transformants. *Molec. Gen. Genet.* (In the press.)

Daniels, S., Strausbaugh, L. D., Ehrman, L. & Armstrong, R. 1984 Sequences homologous to P elements occur in *Drosophila paulistorum*. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6794–6797.

Doolittle, W. F. & Sapienza, C. 1980 Selfish genes, the phenotype paradigm and genome evolution. *Nature, Lond.* **284**, 601–603.

Engels, W. R. 1979 Hybrid dysgenesis in *Drosophila melanogaster*. Rules of inheritance of female sterility. *Genet. Res., Camb.* **33**, 219–236.

Engels, W. R. 1981*a* Hybrid dysgenesis in *Drosophila* and the stochastic loss hypothesis. *Cold Spring Harbor Symp. quant. Biol.* **45**, 561–565.

Engels, W. R. 1981*b* On estimating the total number of genes in the genome and similar problems. *J. math. Biol.* **11**, 45–50.

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

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Engels, W. R. 1984 A trans-acting product needed for P factor transposition in *Drosophila*. *Science, Wash.* **226**, 1194–1196.

Engels, W. R. & Preston, C. R. 1979 Hybrid dysgenesis in *Drosophila melanogaster*: the biology of male and female sterility. *Genetics* **92**, 161–175.

Engels, W. R. & Preston, C. R. 1980 Components of hybrid dysgenesis in a wild population of *Drosophila melanogaster*. *Genetics* **95**, 111–128.

Flavell, A. J. 1984 Role of reverse transcription in the generation of extrachromosomal *copia* mobile genetic elements. *Nature, Lond.* **310**, 514–516.

Kaplan, N., Darden, T. & Langley, C. 1985 Evolution and extinction of transposable elements in Mendelian populations. *Genetics* **109**, 459–480.

Karess, R. E. & Rubin, G. M. 1984 Analysis of P transposable element functions in *Drosophila*. *Cell* **38**, 135–146.

Kidwell, M. G. 1979 Hybrid dysgenesis in *Drosophila melanogaster*: the relationship between the P–M and I–R interaction systems. *Genet. Res., Camb.* **33**, 205–217.

Kidwell, M. G. 1981 Hybrid dysgenesis in *Drosophila melanogaster*: the genetics of cytotype determination in a neutral strain. *Genetics* **98**, 275–290.

Kidwell, M. G. 1983 Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1655–1659.

Kidwell, M. G., Kidwell, J. F. & Sved, J. A. 1977 Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility, and male recombination. *Genetics* **36**, 813–83.

Kugimiya, W., Ikenaga, H. & Saigo, K. 1983 A close relationship between the long terminal repeats of avian leukosis–sarkoma virus and *Drosophila copia*-like movable genetic elements. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3193–3197.

Martin, G., Wiernasz, D. & Schedl, P. 1983 Evolution of *Drosophila* repetitive-dispersed DNA. *J. molec. Evol.* **19**, 203–213.

Miller, D. W. & Miller, L. K. 1982 A virus mutant with an insertion of a *copia*-like transposable element. *Nature* **299**, 562–564.

Montgomery, E. A. & Langley, C. H. 1983 Transposable elements in Mendelian populations. II. Distribution of three *copia*-like elements in a natural population of *Drosophila melanogaster*. *Genetics* **104**, 473–483.

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

Orgel, L. E. & Crick, F. H. C. 1980 Selfish DNA: the ultimate parasite. *Nature, Lond.* **284**, 604–6.

Picard, G., Bucheton, A., Lavige, J-M. & Pelisson, A. 1976 Répartition géographique des trois types de souches impliquées dans un phénomène de stérilité à déterminisme non mendélien ches *Drosophila melanogaster*. *C.r. hebd. Séanc. Acad. Sci., Paris* **282**, 1813–1816.

Picard, G., Lavige, J. M., Bucheton, A. & Bregliano, J. C. 1977 Non-Mendelian female sterility in *Drosophila melanogaster*: physiological pattern of embryo lethality. *Biol. Cellulaire* **29**, 89–98.

Preston, C. R. & Engels, W. R. 1984 Movement of P elements within a P strain. *Dros. Inform. Ser.* **60**, 169–170.

Rubin, G. M. 1983 Dispersed repetitive DNAs in *Drosophila*. In *Mobile genetic elements* (ed. J. A. Shapiro, pp. 329–362). New York, London: Academic Press.

Sang, H. M., Pélisson, A., Bucheton, A. & Finnegan, D. J. 1985 Molecular lesions associated with white gene mutations induced by I–R hybrid dysgenesis in *Drosophila melanogaster*. *EMBO J.* **3**, 3079–3085.

Scavarda, N. J. & Hartl, D. L. 1984 Interspecific DNA transformation in *Drosophila*. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7615–7619.

Schaefer, R. E., Kidwell, M. G. & Fausto-Sterling, A. 1979 Hybrid dysgenesis in *Drosophila melanogaster*: morphological and cytological studies of ovarian dysgenesis. *Genetics* **92**, 1141–1152.

Shapiro, J. 1983 (ed.) *Mobile genetic elements*. 688 pages. New York, London: Academic Press.

Shiba, T. & Saigo, K. 1983 Retrovirus-like particles containing RNA homologous to the transposable element *copia* in *Drosophila melanogaster*. *Nature, Lond.* **302**, 119–124.

Simmons, M. J. & Karess, R. E. 1985 A report on molecular and population biology of hybrid dysgenesis. 24–26 September 1984, Cambridge, England. *Dros. Inf. Serv.* **61**. (In the press.)

Simmons, M. J. & Lim, J. K. 1980 Site specificity of mutations arising in dysgenic hybrids of *Drosophila melanogaster*. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6042–6046.

Simmons, M. J. & Bucholz, L. M. 1985 Transposase titration in *Drosophila melanogaster*: a model for cytotype in the P–M system of hybrid dysgenesis. *Proc. natn. Acad. Sci. U.S.A.* (In the press.)

Spradling, A. C. & Rubin, G. M. 1982 Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science, Wash.* **218**, 341–347.

Thompson, J. N. & Woodruff, R. C. 1978 Mutator genes – pacemakers of evolution. *Nature, Lond.* **274**, 317–321.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES