SELFISH DNA: A SEXUALLY-TRANSMITTED NUCLEAR PARASITE

DONAL A. HICKEY

Department of Biology, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5

Manuscript received November 30, 1981
Revised copy accepted April 10, 1982

ABSTRACT

A quantitative population genetics model for the evolution of transposable genetic elements is developed. This model shows that "selfish" DNA sequences do not have to be selectively neutral at the organismic level; indeed, such DNA can produce major deleterious effects in the host organism and still spread through the population. The model can be used to explain the evolution of introns within eukaryotic genes; this explanation does not invoke a long-term evolutionary advantage for introns, nor does it depend on the hypothesis that eukaryotic gene structure may be an evolutionary relic. Transposable genes that carried information specifying sexual reproduction in the host organism would favor their own spread. Consequently, it is tempting to speculate that some of the genes controlling sex were originally selected as transposable elements.

INTRODUCTION

THE evolutionary origin and maintenance of noncoding DNA sequences in the genomes of eukaryotic organisms has been the subject of much recent debate (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980; CAVALIER-SMITH 1980; DOVER 1980; DOVER and DOOLITTLE 1980; ORGEL, CRICK and SAPIENZA 1980). Generally, it had been assumed that, if these sequences exist within eukaryotic genomes, they must perform functions that are beneficial to the organism that contains them. However, the type of functional explanations that are currently favored, such as the generation of new gene functions through exon shuffling (GILBERT 1979), involve a long-term evolutionary advantage and, as pointed out by DOOLITTLE (1978), this might explain the persistence of the phenomenon within a species, but would not explain its origin and initial spread within a population of organisms. Therefore, one is left to ask what immediate advantage does the acquisition of noncoding DNA confer on an individual. One suggestion is that extra DNA contributes to genome size, which in turn contributes to nuclear volume (CAVALIER-SMITH 1978, 1980). Alternatively, it has been suggested that those noncoding sequences which occur within cistrons, the introns, are, in fact, not an evolved trait but rather an evolutionary relic (DOOLITTLE 1978; DARNELL 1978). More recently an explanation for the origin of noncoding DNA, other than introns, has been proposed that does not impute a benefit to the organism (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980). This latter explanation has been labelled the concept of

"selfish DNA"—specifically, it has been argued that sequences whose only "function" is rapid self-replication, will tend to persist within genomes. However, as pointed out by CAVALIER-SMITH (1980) rapid intra-genome spread may explain intra-genomic persistence but it cannot explain inter-genomic spread of an element. Therefore, he concludes that the spread of any element to all genomes within a population must be due to inter-genomic rather than intra-genomic selection. The following definitions may help to clarify the arguments concerning the inter- and intra-genomic spread of genetic elements.

The average copy number per cell (f) of a transposable repetitive genetic element in a population may be expressed as follows:

$$f = \frac{a \cdot b}{N}$$

where a is the average number of copies per genome, among those genomes which contain the repetitive sequence; b is the number of genomes containing at least one copy of the element; N is the total number of genomes in the population. In any given population, the value of f can increase due to an increase in the value of either a or b. However, in the event that a increases and b does not increase, the element cannot be said to spread in the population; its overall "frequency" increases, but this is only because those genomes that already contain some copies acquire more copies. Therefore, in an asexual population, selfish genes can increase in number within certain genomes, but they cannot colonize new genomes. Their evolutionary fate is dependent on the survival of their host genomes; and if they confer no selective advantage on these host genomes, there will be no systematic tendency towards an increase in the frequency of genomes containing the selfish elements. Thus, in agreement with the argument of CAVALIER-SMITH (1980), selfish genes cannot spread in any deterministic fashion among asexual organisms with strictly clonal patterns of reproduction. The evolutionary fate of transposable or "selfish" genes in a sexual population will, however, be quite different. Indeed, such genes may spread quite rapidly through sexual populations due to their ability to colonize new genomes during zygote formation.

The model presented here demonstrates that it is possible to define genetic elements whose only function is to replicate and transpose within a genome. These elements may actually have serious deleterious effects on the organism that harbors them but will nevertheless spread deterministically to fixation in a population of diploid sexually-reproducing organisms. Such an element would be characteristic not of all eukaryotes, but only of all diploid, obligately sexual eukaryotes.

THE MODEL

Assume, for simplicity, a single transposable element that replicates in the process of transposition. This "copy out" mechanism is characteristic of observed transposons (HARSHEY and BUKHARI 1981). Assume, also for simplicity,

that there is only a single recognition site per haploid genome for insertion of this element. (The more commonly observed case of multiple insertion sites will be discussed later.) A further assumption is that this element has no effect on the fitness of the organism; again, this assumption will be relaxed later.

This genetic element can transpose only from its single insertion site to the homologous site on the homologous chromosome. The number of genomes in which both homologous sites are occupied will depend on the efficiency of transposition; this may vary widely between different transposable elements. Thus, if the transposition process is very efficient, virtually all of the gametes from a "heterozygous" zygote will contain a copy. In other words, an individual who receives one copy of this element from a parent could pass along a copy in every gamete it produces. This contrasts with the well known case of nontransposable (Mendelian) genes where, the heterozygous offspring having received a single copy from a parent, then passes a copy on the average, to 50% of his/her offspring. Stated in this way, it would appear that a transposable element, even one with only a single insertion site per haploid genome, could spread within a population at twice the rate of nontransposable genes. I now present a more precise and quantitative description of this process.

Let p equal the frequency of gametes in a diploid sexual population that contain a transposable element of the type defined above. Initially, the value of p is close to zero. If gametes combine randomly to form zygotes, zygotic types and frequencies will be as described in Table 1.

If we assume that gametes containing a copy of the element occur with frequency p, then with random fusion of gametes, p^2 of the zygotes will contain two copies ("homozygotes") and $2p(1-p)$ of the zygotes will receive a single copy ("heterozygotes"). However, by its nature, the element will tend to transform these "heterozygotes" into "homozygotes." Therefore, after a number of mitotic cell divisions have occurred, we expect to get only "homozygotes" carrying a copy of the element on each of two homologous chromosomes, and the alternative homozygotes which contain no copy of the element.

In the course of a single generation the frequency of gametes carrying a copy of the element changes from p to $(p^2+2p(1-p))$. Initially, when the value of p is close to zero and the value of $(1-p)$ is close to unity, the change in p per generation will equal approximately p. That is, the frequency of gametes carrying the transposable element will double every generation. Of course, if the fre-

TABLE 1

Frequency of genomes carrying one, two or no copies of a transposable genetic element, assuming that the element has no effect on fitness

	2 copies	1 copy	no copy
gametes	0	p	$(1-p)$
zygotes	p^2	$2p(1-p)$	$(1-p)^2$
adults	$p^2+2p(1-p)$	0	$(1-p)^2$
F_1 gametes	0	$p^2+2p(1-p)$	$(1-p)^2$

quency of transposition is low, the increase in frequency of the element will be correspondingly slow. Given that there is a nonzero probability of transposition, the frequency of the element in the population will continue to increase deterministically. It is interesting to note, at this point, that if the population were diploid and sexual, but obligately self-fertilizing, the change in the value of p would equal zero. The spread of this element depends not only on sexuality and diploidy, but also on out-breeding.

As the frequency of the transposable element increases in the population, its rate of increase will no longer approximately equal its current frequency. For instance, when $p = 0.5$, the rate of increase still equals $(p)(1-p)$, which in this case will be 0.25, or a 50% increase in frequency in a single generation. Thus as p increases, the relative rate of increase diminishes, but the absolute rate of increase of the element increases to reach a maximum when $p = 0.5$ (Table 2). However, the change in frequency is always positive and, in the absence of a countervailing selective force, this element will spread rapidly through the population.

I will expand the model to allow for the possibility that the element in question does in some way affect the fitness of the organism which carries it. Obviously, if this effect is positive, (*i.e.*, the fitness of its host organism increases), it will simply hasten its spread through the population. The more interesting case to consider is where the fitness of the host organism is lowered relative to that of conspecifics lacking this element. Intuitively, we predict that a genetic element which lowers the fitness of the organism that contains it will itself be eliminated from the population. That prediction is not true in this case.

Assume that the fitness, in terms of reproductive output, of organisms carrying the element in question is lowered by some fraction that I shall define as s,

TABLE 2

Absolute and relative rates of increase in the frequency of a transposable element in a diploid, sexual, panmictic population. The element is assumed to produce no phenotypic effect in this case

Frequency of element (p)	Absolute rate of Change in frequency: $p(1-p)$	Relative rate of increase in frequency $\frac{p(1-p)}{p}$	Relative rate of loss of alternative genotype $\frac{p(1-p)}{1-p}$
0.001	0.000999	0.999	0.001
0.01	0.0099	0.99	0.01
0.1	0.09	0.90	0.1
0.2	0.16	0.80	0.2
0.3	0.21	0.70	0.3
0.4	0.24	0.60	0.4
0.5	0.25	0.50	0.5
0.6	0.24	0.40	0.6
0.7	0.21	0.30	0.7
0.8	0.16	0.20	0.8
0.9	0.09	0.10	0.9
0.99	0.0099	0.01	0.99

the selection coefficient. Now, the element still has a tendency to spread in the population because of its ability to transpose onto the homologous chromosome but it is also at a disadvantage since organisms containing it leave less offspring. The frequency of gametes and zygotes containing the element are shown in Table 3.

With selection, the change in frequency becomes:

$$\begin{aligned}&(\text{Frequency in } F_1 \text{ gametes}) - (\text{Frequency in "parental" gametes})\\ &= [\{p^2 + 2p(1-p)\}(1-s)/\bar{W}] - p\\ &= \frac{p(1-p)[1-s\{1+(1-p)\}]}{\bar{W}}\end{aligned}$$

where $\bar{W}$ is the mean fitness of individuals in the population. In summary, whereas the frequency change in the absence of selection equaled $p(1-p)$, the change when selection acts against individuals carrying the transposable element is: $(p)\ (1\text{-}p)$ multiplied by a factor $\frac{1\text{-}s\ \{1 + (1\text{-}p)\}}{\bar{W}}$. This frequency change can be positive or negative depending on the value of s. The critical question is how large a value of s (selection against organisms containing the element) will still allow for the spread of this element. Mathematically, we ask what values of s will yield a positive value for the expression $(p)\ (1\text{-}p).\ \frac{1\text{-}s\ \{1 + (1\text{-}p)\}}{\bar{W}}$, *i.e.*, the expression that describes the change in frequency from one generation to the next. The above expression will be positive provided the numerator $[1-s\{1 + (1-p)\}]$ is positive.

i.e.,
$$s < \frac{1}{1 + (1\text{-}p)}$$

This means that the maximum value of the selection coefficient (s) that still allows for the spread of the transposable element varies with the frequency of the element. A sample of values for this relationship is shown in Table 4.

TABLE 3

Frequency of gametes and zygotes containing a transposable genetic element, given that the element has a deleterious effect on its host organism

	Number of transposon copies per individual		
	2	1	0
gametes	0	p	$(1-p)$
zygotes	p^2	$2p(1-p)$	$(1-p)^2$
adults	$p^2 + 2p(1-p)$	0	$(1-p)^2$
F_1 gametes	0	$\frac{[p^2 + 2p(1-p)](1-s)}{\bar{W}}$	$\frac{(1-p)^2}{\bar{W}}$

$\bar{W} = [p^2 + 2p(1-p)]\ (1-s) + (1-p)^2 \equiv [1-(1-p)^2]\ (1-s) + (1-p)^2$
$(1-s) \leqslant \bar{W} \leqslant 1$ in this case.

TABLE 4

Relationship between the frequency of a transposable element and the maximum level of negative organismic selection which allows for its continued spread

Frequency of element (p)	Maximum value of selection coefficient (s)
0	0.50
0.1	0.53
0.2	0.55
0.5	0.67
0.8	0.83
0.99	0.99

Initially, when the value of p is close to zero, the maximum value of s is approximately 0.5. This is consistent with the observation that, in the absence of selection, the element approximately doubles in frequency every generation (see Table 2). Naturally, this tendency would be counteracted if each organism containing the element left only half as many offspring as those organisms which do not contain it. In other words, when the element is rare, most zygotes receive a single copy. However they pass along twice as many copies per gamete as would be expected for a nontransposable gene. However, if they pass along only half as many gametes as conspecifics, then they will pass along as many copies of this element as would be expected for a nonselected, nontransposable gene.

The important conclusion to be drawn from these calculations is that if a transposable element is introduced into a population of randomly mating sexual diploid organisms, even if it reduces the fitness of those organisms that contain it by any fraction that is less than 50% ($s < 0.50$), it will still spread quite rapidly through the population provided that the efficiency of transposition is high. The reason is simply that its initial rate of spread equals approximately twice the reproductive rate of its host genome; therefore, if the host genes continue to replicate at greater than half the rate of allelic genes in other individuals, then the frequency of the transposable element will increase.

Even more surprising is the fact that, once the element reaches an appreciable frequency in the population, the selection against individuals can be much greater without impeding its continued spread to fixation in the population (see Table 4). Indeed, if 90% of the population contained the element, it would colonize the remaining 10% of the genomes, unless it were virtually lethal ($s = 0.91$) to its host organism. This frequency-dependent effect is most easily understood by noting that, in the absence of selection, the relative rate of loss of the "transposon-free" genotype is positively correlated with the frequency of the transposon (see Table 2, column 4).

The end result of the process will be two-fold: the transposable element will become fixed in the population and, secondly, the average fitness of the whole population will decrease from a value of 1 to a value of $(1-s)$.

As noted earlier, the intrinsic rate of spread of the element depends on the efficiency of the transposition. Therefore, for those elements with a low effi-

ciency of transposition, the selection coefficient that is necessary to prevent their spread will be correspondingly lower.

Some transposable elements have many potential insertion sites within the genome. In this case we need to consider "pairs of genomes" rather than "pairs of insertion sites." There may be transposition within genomes as well as between genomes and gametes may contain not one but several copies of the element. The mathematical analysis of this "multiple site" model is more complex than for the more restrictive situation described here (B. CHARLESWORTH, personal communication), but the main conclusion remains unchanged, *i.e.*, the rate of spread of the element depends on the rate of colonization of new genomes after zygote formation.

Introns as Selfish DNA: For a plausible mechanism for the evolution of introns we could reason as follows. Assume, in this case, an element that has several potential sites of insertion in the genome. These insertion sites may be within the coding sequences of essential genes, and the inserted sequence may destroy the gene function. Obviously, such a sequence would be strongly selected against in a haploid organism. In diploids, however, we would effectively have a zygote that inherited a "recessive lethal" at one locus and, by the act of transposition, was rendered "recessive lethal" at another locus. The initial spread of the element would not be stopped because of the destruction of host genes at the point of insertion. As the frequency of the element increased in the population, the probability of a cell that contained the element at the same locus in each of two homologous chromosomes, would increase—we might think of these individuals as being "homozygous lethal." Such genotypes would not survive and would not foster the further spread of the transposon. The situation would be equivalent to a mutation-selection balance for recessive lethal genes. However, in this case, the transposition process provides a strong directional "mutation pressure." Any other mutations that lessened the deleterious effects of these insertions would favor the survival of host genes and also the continued spread of the transposon. Therefore selection would act on host genes, or transposon genes, or both, to lessen the deleterious effects. Indeed, the RNA splicing mechanism that appear to counteract the potential negative effect of introns on their hosts may be a variant form of the original transposition mechanism (see DISCUSSION).

DISCUSSION

The model presented here describes the evolutionary fate of a specific kind of transposable genetic element. The type of element described is meant to be illustrative of certain basic features of such genes and not an accurate description of any real transposable element. Now, we may ask how the model might apply to real instances of transposable genetic elements and to "selfish DNA" in general.

Previous discussion of "selfish" and "parasitic" DNA sequences (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980) have focused on a transposon's ability to occupy several sites within a genome rather than their ability to colonize new haploid genomes at zygote formation. Most real transposons prob-

ably possess both of these abilities. Intragenomic spread will contribute to the persistence of transposon within a cell line but it cannot bring about a systematic increase in the frequency of cells or individuals that contain the element (CAVALIER-SMITH 1980; see INTRODUCTION). However, this does not mean that there is no such thing as "selfish" DNA. Given the possibility of colonizing new genomes at zygote formation, these selfreplicating elements can spread rapidly in a "selfish" manner.

The evolutionary dynamics of the system described is similar in many respects to the dynamics of other non-Mendelian systems, such as segregation distortion and meiotic drive (SANDLER and NOVITSKI 1957; HIRAIZUMI, SANDLER and CROW 1960; CROW 1979) or gene conversion (GUTZ and LESLIE 1976). Indeed, the mathematical model described here is quite similar to that developed by PROUT (1953) for meiotic drive. Biologically, however, gene transposition is a very different process; it is one of "gene addition" rather than gene conversion, and there is no distortion of the segregation ratios of linked genes. It is interesting to note that HOLLIDAY (1981) has suggested gene conversion as a possible mechanism for deleting selfish genes.

The major difference between this model and previous models of the population genetics of selfish DNA (OHTA and KIMURA 1981; OHTA 1981) is that it allows for the deterministic spread of a mobile genetic element that is actually harmful (in terms of reduced fitness) to its host. The possibility of the spread of DNA sequences that reduce the fitness of their hosts has generally been discounted, and this problem has been outlined very effectively by DOOLITTLE (1978) with reference to the evolutionary origin of introns. In contrast, ORGEL and CRICK (1980) have considered the possibility that some selfish genes which have a slight negative effect on host fitness might spread through a population. However, they stressed that the reduction in fitness would have to be small and they do not give a quantitative estimate of what its magnitude might be. Since the present model predicts a rapid and systematic increase in the frequency of genomes containing a transposable element, the assumptions regarding the lack of significant negative selection on the host organism can be relaxed. This allows for a much wider potential application of the model to different types of noncoding DNA sequences, including a transposable precursor of existing introns.

There have been several suggestions in the recent literature (*e.g.*, CALOS and MILLER 1980; CAVALIER-SMITH 1978; CHAMBON 1981; FYRBERG *et al.* 1981; WAHLI *et al.* 1981) that introns within eukaryotic genes may be classifiable as selfish DNA sequences that are transposable within the genome and that are evolutionarily related to prokaryotic insertion sequences (KLECKNER 1977). As stated by CHAMBON (1981), "it might be argued that introns are really mobile genetic elements similar to "transposons" of prokaryotic cells, inserted in the course of evolution into genes that were once whole." Here I have presented an evolutionary mechanism whereby this could come about despite the obvious potential for the initial disruption of host genes. CRICK (1979) suggests that RNA splicing may have evolved as a defense by the cell against an insertion element it was harboring. Alternatively, it is possible that RNA splicing is an

intrinsic property of the elements themselves. For instance, they may be similar to retroviral proviruses in that their replication and transposition involves an RNA intermediate. Thus, splicing could be viewed as a mechanism that evolved initially for separating the insertion element genome from flanking sequences.

There is growing evidence for an evolutionary link between retroviruses and transposable genetic elements (GREEN 1980; TEMIN 1980; FLAVELL 1981; YOUNG and SCHWARTZ 1981; FLAVELL and ISH-HOROWICZ 1981 and others). Both types of element are flanked by relatively long terminal repeats and both cause a small duplication of host DNA at the point of integration. According to the protovirus theory proposed by TEMIN (1980), retroviruses may have evolved from transposable genetic elements; prokaryotic insertion sequences might have been the evolutionary ancestors of the long terminal repeats and were essential for transposition. TEMIN (1980) further suggests that the protoviral element could move either by direct transposition at the DNA level or by transcription, reverse transcription and integration. There is no conclusive evidence that excludes this latter indirect method of transposition for existing transposable genes. GREEN (1980) asks: "could it be that viral RNA serves as a template for a reverse transcriptase resulting in a DNA which serves as an insertion sequence?" There is evidence that several mobile genetic elements are transcribed with a high efficiency (FINNEGAN 1981) and that the transcripts are processed and transported to the cytoplasm (TASHIMA *et al.* 1981). Recent experiments (FLAVELL and ISH-HOROWICZ 1981) show that these transcripts may be converted into circular DNA molecules. Thus, transposable genetic elements and endogenous retroviral proviruses may not only be related, but may be indistinguishable other than by differences in their potential infectivity and physiological effects.

If introns are derived from prokaryotic insertion sequences and are equivalent to the long terminal repeats of transposable elements and retroviruses (or to parts of those sequences), then exon shuffling (GILBERT 1979) and oncogene mobilization (HAYWARD, NEEL and ASTRIN 1981) may both be caused by the same underlying mechanism and both be "accidental" consequences of insertion element behavior.

Regardless of the mechanism of intragenomic replication, the reason that transposable genetic elements can spread through a population of cells or organisms is their ability to integrate at new chromosomal sites after zygote formation. Their spread does not depend on sex *per se* but, more specifically, on a biparental pattern of host genome reproduction. Therefore, the model would predict that those classes of organisms in which biparental reproduction is rare or absent would be encumbered by less nonfunctional or parasitic DNA than most vertebrates. Although the data are rather sparse, this prediction is generally borne out in a comparison of the genomes of fungi and vertebrates (CHAMBON 1981; DONS and WESSELS 1980). Perhaps the best test of the model available to date is the structure of nuclear and mitochondrial genomes in yeast and humans, and particularly with regard to the presence or absence of introns. Among vertebrates the inheritance of nuclear genes is biparental but the inheri-

tance of mitochondrial genomes is uniparental (DAWID and BLACKLER 1972; HUTCHINSON *et al.* 1974). In contrast to this, nuclear genes of yeast are inherited uniparentally during asexual reproduction and biparentally during sexual reproduction; yeast mitochondrial genes are generally inherited biparentally (BIRKY 1978). Given these facts, we would predict that vertebrate nuclear genes and yeast mitochondrial genes might contain significant amounts of nonfunctional DNA: yeast nuclear genes would contain less but possibly some nonfunctional DNA; whereas vertebrate mitochondria, like haploid asexual prokaryotes, should contain virtually no "parasitic" DNA. Recent data (BORST and GRIVELL 1981; ANDERSON *et al.* 1981) on the nucleotide composition of yeast and human mitochondrial genomes fits this prediction exactly. Observations on the nuclear genomes of the two organisms indicate that although yeast nuclear genes may contain noncoding sequences, they are not as common as in the nuclear genes of vertebrates (F. SHERMAN, personal communication). In other words, the correlation between the frequency of biparental reproduction of these genomes and the relative amounts of "selfish" DNA is surprisingly good. However, a rigorous test of this correlation must await the availability of much more data on genome structure from a wide variety of organisms.

The claim that the spread of eukaryotic transposable genes is linked to biparental reproduction immediately raises the question of why transposons are found within the genomes of predominantly asexual prokaryotes. Bacterial transposons can move from the main chromosome to a plasmid genome (KLECKNER 1977). If this is a conjugative plasmid, it will carry the transposon to other cells where it can then transpose back onto the main chromosome. In this way bacterial transposons are indirectly infectious. The same may be true of eukaryotic transposable elements which integrate into viral genomes.

Finally, it should be noted that the spread of transposable genetic elements depends, according to this model, on sexual reproduction. Therefore, a transposon that caused sex in the host would favor its own spread. Consequently, an intriguing possibility is that sex itself, and especially outbreeding, is a product of parasitic genes. Certainly, bacterial conjugation would appear to have derived from a mechanism for the transfer of F-element genes rather than host genes (WILLETTS and SKURRAY 1980). Among eukaryotes, it is known that mating type in yeast is controlled by a transposable genetic element (KLAR *et al.* 1981). Moreover, maleness in many dipteran insects is controlled by a localized chromosomal element, called the M-factor (GREEN 1980), and there is evidence that, in some cases, this factor is transposable (MAINX 1964). One would expect that if the same were true of the initial stages in the evolution of sex in higher eukaryotes these genes are, in general, no longer transposable. This suggestion concerning the origin of sex does not deny a long-term evolutionary advantage for sex due to its function in reshuffling allelic combinations. However, it has been difficult to develop a satisfactory model for the origin and short-term maintenance of sexual reproduction (see MAYNARD-SMITH, 1978 for a review). It is clear, from the calculations presented above, that a transposable gene

which favored sexual reproduction would be strongly selected for in the short term.

Generally, the existence of transposable genes may explain many of those evolutionary problems that have proven notoriously difficult to understand when one uses population genetics models which are based solely on the principles of Mendelian inheritance. It may be that many genetically-controlled characteristics whose advantage is long-term owe their origin and initial spread within populations to their behavior as short-term genetic pathologies. A relevant analogous example might be the bacterial transducing phages. Clearly, the evolutionary origin of these phages is not due to their transducing capability, yet this property can have a major impact on the evolution of their bacterial host genomes.

In summary, we can make the following generalization about any DNA sequence that has the ability to self-replicate and transpose within the genome and can thus behave as a sexually-transmitted nuclear parasite. Initially, their rate of spread in a population depends on the reproductive success of their hosts but is not equal to it; rather they can spread at twice the rate of the host genes. This is why they may continue to spread provided the reduction in host fitness is less than 50%. However, once they become fixed in a population, the fitness of these transposable elements becomes identical to the fitness of the host; thus their fitness can then be increased only by increasing the fitness of the host. We might consider them to be *semi-parasitic* genes.

I thank Drs. W. H. Cade, F. H. Crick, W. F. Doolittle, G. Dover, T. Holliday, R. McGowan, K. Paigen, J. Rossant and R. S. Singh for comments on an early draft of the manuscript. I am especially grateful to Drs. J. Barrett, B. Charlesworth, M. Kimura and an anonymous reviewer for pointing out the parallels with previous work on segregation distorter and meiotic drive systems. Financial support was provided by a grant from NSERC, Canada.

LITERATURE CITED

Anderson, S., A. T. Bankier, B. C. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. H. Smith, R. Staden and I. G. Young, 1981 Sequence and organization of the human mitochondrial genome. Nature **290**: 457–465.

Birky, C. W., 1978 Transmission genetics of mitochondria and chloroplasts. Ann. Rev. Genet. **12**: 471–512.

Borst, P. and L. A. Grivell, 1981 Small is beautiful—portrait of a mitochondrial genome. Nature **290**: 443–444.

Calos, M. P. and J. H. Miller, 1980 Transposable elements. Cell **20**: 579–595.

Cavalier-Smith, T., 1978 Nuclear volume control by nucleoskeletal DNA, selection for cell volume and cell growth rate, and the solution of the DNA c-value paradox. J. Cell Sci. **34**: 247–278. ——, 1980 How Selfish is DNA? Nature **285**: 617–618.

Chambon, P., 1981 Split Genes. Sci. Amer. **244**(5): 60–71.

Crick, F., 1979 Split genes and RNA splicing. Science **204**: 264–271.

Crow, J. F., 1979 Genes that violate Mendel's rules. Sci. Amer. **240**(2): 134–146.

DARNELL, J. E., 1978 Implications of RNA-RNA splicing in evolution of eukaryotic cells. Science **202:** 1257–1260.

DAWID, I. and A. W. BLACKLER, 1972 Maternal and cytoplasmic inheritance of mitochondrial DNA in *Xenopus*. Dev. Biol. **29:** 152–161.

DONS, J. J. M. and J. G. H. WESSELS, 1980 Sequence organization of the nuclear DNA of *Schizophyllum commune*. Biochim. et Biophys. Acta **607:** 385–396.

DOOLITTLE, W. F., 1978 Genes in pieces: were they ever together? Nature **272:** 581–582.

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

DOVER, G., 1980 Ignorant DNA? Nature **285:** 618–620.

DOVER, G. and W. F. DOOLITTLE, 1980 Modes of genome evolution. Nature **288:** 646–647.

FINNEGAN, D. J., 1981 Transposable elements and proviruses. Nature **292:** 799.

FLAVELL, A., 1981 Did retroviruses evolve from transposable elements? Nature **289:** 10–11.

FLAVELL, A. J. and D. ISH-HOROWICZ, 1981 Extrachromosomal circular copies of the eukaryotic transposable element *copia* in cultured *Drosophila* cells. Nature **292:** 591–595.

FYRBERG, E. A., B. J. BOND, N. D. HERSHEY, K. S. MIXTER and N. DAVIDSON, 1981 The actin genes of Drosophila: protein coding regions are highly conserved but intron positions are not. Cell **24:** 107–116.

GILBERT, W., 1979 Introns and Exons: Playgrounds of Evolution. pp. 1–12. In: *Eukaryotic Gene Regulation ICN-UCLA Symposia on Molecular and Cellular Biology*. Vol. 14. Edited by R. AXEL, T. MANIATIS, C. F. FOX. Academic Press, New York.

GREEN, M. M., 1980 Transposable elements in *Drosophila* and other diptera. Ann. Rev. Genet. **14:** 109–120.

GUTZ, H. and J. R. LESLIE, 1976 Gene conversion: a hitherto overlooked parameter in population genetics. Genetics **83:** 861–866.

HARSHEY, R. M. and A. I. BUKHARI, 1981 A mechanism for DNA transposition. Proc. Natl. Acad. Sci. U.S.A. **78:** 1090–1094.

HAYWARD, W. S., B. G. NEEL and S. M. ASTRIN, 1981 Activation of a cellular *Onc* gene by promotor insertion in ALV-induced lymphoid leukosis. Nature **290:** 475–480.

HIRAIZUMI, Y., L. SANDLER and J. F. CROW, 1960 Meiotic drive in natural populations of *Drosophila melanogaster* III. Populational implications of the segregation-distorter locus. Evolution **14:** 433–444.

HOLLIDAY, R., 1981 Gene conversion: a possible mechanism for eliminating selfish DNA. In: *Proceeding of the Symposium on Molecular and Cellular Mechanisms of Mutagenesis*. Gatlinburg, Tennessee, 1981.

HUTCHINSON, C. A., J. E. NEWBOLD, S. S. POTTER and M. H. EDGELL, 1974 Maternal inheritance of mammalian mitochondrial DNA. Nature **251:** 536–538.

KLAR, A. J. S., J. N. STRATHERN, J. R. BROACH and J. B. HICKS, 1981 Regulation of transcription in expressed and unexpressed mating type cassettes of yeast. Nature **289:** 239–244.

KLECKNER, N., 1977 Translocatable elements in procaryotes. Cell **11:** 11–23.

MAINX, F., 1964 The genetics of *Megalselia scalaris* Loew (Phoridae): a new type of sex determination in Diptera. Amer. Nat. **98:** 415–430.

MAYNARD-SMITH, J., 1978 *The Evolution of Sex*. Cambridge Univ. Press, London.

OHTA, T., 1981 Population genetics of selfish DNA. Nature **292:** 648–649.

OHTA, T. and M. KIMURA, 1981 Some calculations on the amount of selfish DNA. Proc. Natl. Acad. Sci. U.S.A. **78:** 1129–1132.

ORGEL, L. E. and F. H. C. CRICK, 1980 Selfish DNA: the ultimate parasite. Nature **284:** 604–607.

ORGEL, L. E., F. H. C. CRICK and C. SAPIENZA, 1980 Selfish DNA. Nature **288:** 645–646.

PROUT, T., 1953 Some effects of variations in the segregation ratio and of selection on the frequency of alleles under random mating. Acta Genetica **4:** 148–151.

SANDLER, L. and E. NOVITSKI, 1957 Meiotic drive as an evolutionary force. Amer. Natur. **91:** 105–110.

TASHIMA, M., B. CALABRETTA, G. TORELLI, M. SCHOFIELD, A. MAIZEL and G. F. SAUNDERS, 1981 Presence of a highly repetitive and widely dispersed DNA sequence in the human genome. Proc. Natl. Acad. Sci. U.S.A. **78:** 1508–1512.

TEMIN, H. M., 1980 Origin of retroviruses from cellular moveable genetic elements. Cell **21:** 599–600.

WAHLI, W., I. B. DAWID, G. U. RYFFEL and R. WEBER, 1981 Vitellogenesis and the vitellogenin gene family. Science **212:** 298–304.

WILLETTS, N. and R. SKURRAY, 1980 The conjugation system of F-like plasmids. Ann. Rev. Genet. **14:** 41–76.

YOUNG, M. W. and H. E. SCHWARTZ, 1981 Nomadic Gene families in Drosophila. Cold Spr. Harb. Symp. Quant. Biol. Vol. XLV: 629–640.

Corresponding editor: W. J. EWENS

Note added in proof

The fact that transposable elements can spread more rapidly in sexual than in asexual populations has also been discovered by J. F. CROW (personal communication); the analogy between the spread of transposable elements in sexual populations and "meiotic drive" systems has been noticed independently by J. BARRETT, Cambridge University (personal communication).