THE ROYAL SOCIETY

Received 14 October 2002
Accepted 12 December 2002
Published online 19 March 2003

Site-specific selfish genes as tools for the control and genetic engineering of natural populations

Austin Burt

Department of Biological Sciences and Centre for Population Biology, Imperial College, Silwood Park, Ascot, Berkshire SL5 7PY, UK (a.burt@ic.ac.uk)

Site-specific selfish genes exploit host functions to copy themselves into a defined target DNA sequence, and include homing endonuclease genes, group II introns and some LINE-like transposable elements. If such genes can be engineered to target new host sequences, then they can be used to manipulate natural populations, even if the number of individuals released is a small fraction of the entire population. For example, a genetic load sufficient to eradicate a population can be imposed in fewer than 20 generations, if the target is an essential host gene, the knockout is recessive and the selfish gene has an appropriate promoter. There will be selection for resistance, but several strategies are available for reducing the likelihood of it evolving. These genes may also be used to genetically engineer natural populations, by means of population-wide gene knockouts, gene replacements and genetic transformations. By targeting sex-linked loci just prior to meiosis one may skew the population sex ratio, and by changing the promoter one may limit the spread of the gene to neighbouring populations. The proposed constructs are evolutionarily stable in the face of the mutations most likely to arise during their spread, and strategies are also available for reversing the manipulations.

Keywords: population eradication; population genetic engineering; homing endonuclease genes; vector-borne diseases

1. INTRODUCTION

Some species—a relatively small number—cause substantial harm to the human condition; most prominent are those that cause disease, transmit disease or reduce agricultural output. Many such species have long been targets of population control, with varying degrees of success, but some species are still beyond control by current methods, and new approaches are required. Genetic methods of engineering or eradicating natural populations have been much discussed (Knipling 1979; Curtis 1985; Hastings 1994), most recently in the context of using transposable elements or bacterial symbionts to drive novel genes of interest into a population (Ribeiro & Kidwell 1994; Beerntsen *et al.* 2000; Braig & Yan 2002). However, there are inherent difficulties with these proposals, in particular relating to the stability of the proposed constructs, and good reasons to think they may not work (Turelli & Hoffmann 1999; Braig & Yan 2002; Spielman *et al.* 2002; and § 6, below). In this paper I explore a series of alternative genetic approaches based on the use of site-specific selfish genes—genes that exploit host functions to copy themselves into a particular target sequence. These alternative approaches appear to have a number of desirable features, including evolutionary stability and reversibility.

Naturally occurring examples of site-specific selfish genes include homing endonuclease genes (HEGs), group II introns and some site-specific LINE-like transposable elements (Chevalier & Stoddard 2001; Belfort *et al.* 2002; Eickbush 2002). Out of the three types, HEGs have the simplest mechanism of action (described further in § 2), while the other two spread via an RNA intermediate and reverse transcription. For simplicity of exposition, HEGs will be used as exemplars throughout the paper, though this is not meant to imply that the other two types will not be useful. Whichever class of site-specific selfish gene is used, all the various proposals presuppose the ability to engineer such genes to recognize a new target sequence. Work on designing enzymes to recognize a specified DNA sequence is ongoing, motivated in part by their potential use in functional genomics and gene therapy (Chandrasegaran & Smith 1999; Segal *et al.* 1999; Bibikova *et al.* 2001; Buchholz & Stewart 2001; Chevalier *et al.* 2002; Santoro & Schultz 2002; Seligman *et al.* 2002; Takahashi & Fujiwara 2002). Engineering group II introns to target new sequences is particularly simple, as recognition largely depends upon RNA–DNA basepairing (Guo *et al.* 2000). Despite this activity, the uses of engineered selfish genes for manipulating natural populations appear not to be recognized, and an exploration of the possibilities seems warranted on three grounds: to motivate more rapid development of the technology; to warn of containment issues that ought to be addressed during development; and to stimulate discussions on the desirability of eradicating or genetically modifying particular species.

2. THE BASIC CONSTRUCT

HEGs are selfish or parasitic genes that can spread through populations owing to their biased 'super-Mendelian' inheritance (Chevalier & Stoddard 2001; Goddard *et al.* 2001). They encode an enzyme that recognizes and cleaves a 20–30 bp sequence found on chromosomes not containing a copy of the HEG. The HEG itself is inserted in the middle of its own recognition sequence, and so chromosomes carrying the HEG are protected from being cut. The broken HEG^- chromosome will typically be repaired by the cell's recombinational repair system, which uses the intact HEG^+ homologue as a template. After

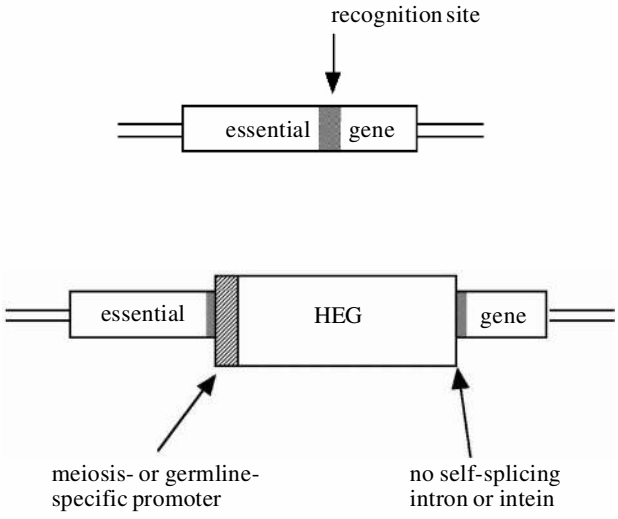

Figure 1. A construct for biological control: a HEG engineered to recognize a sequence in an essential gene for which the knockout phenotype is recessive. Note that the HEG is inserted into the middle of its own recognition sequence.

repair, both chromosomes will contain a copy of the HEG, and a heterozygote will have been converted into a homozygote. Thus, the biased inheritance arises from a combination of a sequence-specific endonuclease inserted in the middle of its own recognition sequence and the cell's own recombinational repair pathway.

The simplicity of this mechanism suggests that it may be open to human artifice. The proposed construct is illustrated in figure 1, and the essential features are as follows.

(i) A HEG is engineered to recognize and cut a sequence in the middle of an essential gene, and the HEG is inserted into the middle of its own recognition sequence, simultaneously disrupting the gene and protecting the chromosome from being cut. Naturally occurring HEGs do not usually disrupt the function of the host gene because they are associated with self-splicing group I introns or inteins (Chevalier & Stoddard 2001), but the engineered element would not have these.
(ii) The target gene is chosen such that the knockout mutation has little phenotypic effect in the heterozygous state, but is severely deleterious when homozygous (i.e. the knockout is recessive).
(iii) Finally, the HEG is under the control of a meiosis-specific promoter, so that heterozygous zygotes develop normally, but transmit the HEG to a disproportionate fraction of their gametes.

This last condition may be relaxed, depending upon when the target gene is expressed: if the target gene is expressed only in larvae, or in somatic tissues, then the promoter can be adult-specific, or germline-specific.

If such a construct is introduced at low frequency into a population, then initially it will appear mostly in the heterozygous state, and so it will show transmission-ratio distortion (TRD) but few harmful effects. It will therefore increase in frequency, until it reaches an equilibrium frequency at which the harmful effects balance the TRD. If

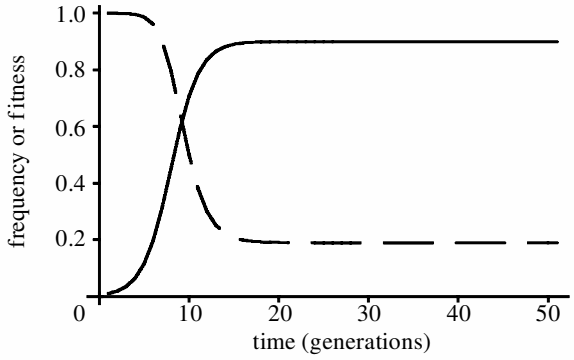

Figure 2. Frequency of the HEG (solid curve) and population mean fitness (dashed curve) assuming $e = 0.9$ and an initial release frequency of 1%. These results, and all others in the paper, are for an idealized population, from which all real populations will deviate in some way. They should, therefore, be taken as rough indications, not precise predictions.

we assume that the population is large and mates randomly, that the knockout is a recessive lethal and that TRD occurs equally in males and females, then the equilibrium frequency of the HEG ($\hat{q}$) can be shown to be $\hat{q} = e$, where e is the probability that the HEG$^-$ allele in a heterozygote is converted to a HEG$^+$ allele; $e = 0$ for Mendelian inheritance. The load imposed upon the population (i.e. the fraction of the reproductive effort that is rendered unproductive) is then equal to the frequency of homozygotes, $L = \hat{q}^2$, and the mean fitness of the population is 1 minus this, or $\hat{\bar{w}} = 1 - e^2$. For example, HEGs of yeasts can show extreme TRD, with $e \approx 0.99$ (Jacquier & Dujon 1985; Wenzlau *et al.* 1989). If one considers, conservatively, an engineered HEG with a TRD of $e = 0.9$ (this same assumption will be made in all numerical examples in this paper), then the equilibrium mean fitness of the population will be $\hat{\bar{w}} = 0.19$. That is, four-fifths of zygotes produced will die, and only one-fifth will survive to reproduce. Moreover, this load will arise relatively quickly. If the HEG is introduced into 1% of the population, then it will take only $t_{1,90} = 12$ generations for the load to reach 90% of its equilibrium value. If one can manage to release an initial frequency of only 0.01%, then it will take 19 generations. The progression to equilibrium is shown in figure 2.

These results are fairly robust to changes in the fitness scheme. For example, if the homozygote has some residual viability (i.e. the knockout is sub-lethal), then this can actually increase the genetic load. Load is highest just at the point at which the HEG can go to fixation. The results are also robust to a certain level of heterozygote impairment (i.e. the knockout is incompletely recessive). For example, if the homozygote is lethal and the heterozygote has a fitness level 90% of the wild-type, then $\hat{\bar{w}} = 0.192$ (instead of 0.19) and $t_{1,90} = 14$ generations (instead of 12 generations; calculations not shown).

(a) *Evolutionary stability*

A key feature of this construct is that it is evolutionarily stable, in the sense that the mutant forms most likely to arise as it spreads through a population will be selected against and lost. For example, a mutant HEG that loses the ability to recognize or cut the target DNA will have

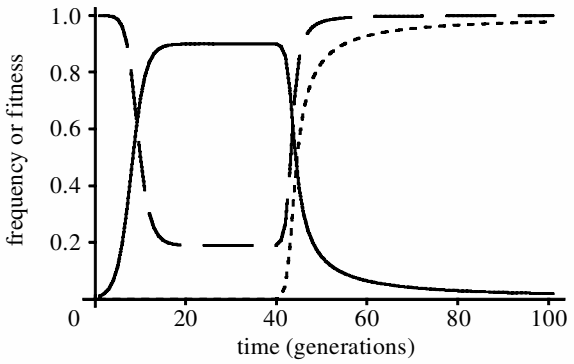

Figure 3. Recalling a HEG. A resistant allele is introduced at 1% in generation 40 (dotted curve). All other parameters are as in figure 2.

reduced TRD, without reducing the harm done to the host, and so will be lost from the population. With a mutation rate of 1%, there will be little effect on the dynamics: $\hat{\bar{w}} = 0.22$ (instead of 0.19) and $t_{1,90} = 12$ generations (as before). Even if the mutation rate is as high as 10%, the load will still be substantial ($\hat{\bar{w}} = 0.5$, $t_{1,90} = 13$ generations). Similarly, mutant HEGs that are active in somatic as well as germline tissues, or that show less sequence specificity, will be more harmful to the host without increasing the TRD, and so will also be selected out. Simulations show that such mutations will have even less effect on the invasion dynamics than mutations to non-functional HEGs (not shown).

This evolutionary stability of the proposed construct gives it a marked advantage over the alternative strategy of using a non-Mendelian genetic element to drive a toxic gene into a population. In the latter case, knockout mutations destroying the function of the toxic gene are sure to arise, and these mutations will spread at the expense of the toxic gene. Unless one can achieve quite high release frequencies, such an approach is probably limited to modifying a population, rather than eradicating it (discussed further in § 6).

(b) *Reversibility*

A further attractive feature of the proposed construct is that it is fully reversible. If one targets a gene that, when knocked out, is strongly deleterious, then there will be strong selection in favour of resistant alleles—sequences that are functional, but are not recognized and cut by the HEG. One could engineer resistant alleles by, for example, using the degenerate property of the genetic code to create a DNA sequence that coded for the same amino acid sequence but differed in nucleotide sequence from the target (e.g. by changing many third-position sites). Releasing such resistant alleles could be used to effectively 'recall' a HEG, as the resistant allele would spread through the population, driving the HEG extinct (figure 3).

(c) *Wide applicability*

The logic of the approach requires that the knockouts be largely recessive, and that homing occurs either at meiosis, or in such a way that the fitness of heterozygous zygotes is not impaired but they produce a predominance of HEG^+ meiotic products. These criteria should be achievable in most outcrossed eukaryotes. Even in species that are predominantly haploid, with only a brief diploid phase, one could target a gene encoding a protein needed for the entry into meiosis, and then have homing occur during meiosis. Alternatively, some predominantly haploid taxa (including malarial *Plasmodium*) have an extended post-meiotic syncytial phase, and so in these species one might also target a protein needed during meiosis (e.g. a synaptonemal complex protein) and have homing occur during the syncytial phase. Highly inbred and wholly asexual species will be less amenable to control by engineered selfish genes.

3. INCREASING THE LOAD

Having described the logic of the construct, I now consider how to increase the load imposed upon a population. The extent to which this is necessary will depend upon the efficacy of the construct in showing TRD. If one can engineer a highly effective HEG with $e = 0.999$, then targeting a single recessive lethal gene will give an equilibrium mean fitness of $\hat{\bar{w}} = 0.002$ (i.e. 99.8% of reproductive output is wasted, and only 0.2% is viable), with $t_{1,90} = 11$, probably sufficient to eradicate many populations. However, such high levels of TRD may not be achievable. If, as assumed here, one can achieve a TRD of only $e = 0.9$, then only four-fifths of the population will die, and for some pests this may not have a substantial effect on their population dynamics, as it may merely relax density-dependent pressures on survival and reproduction. It is therefore worthwhile investigating how one might increase the load further. Two possibilities will be considered: targeting alternative loci and targeting multiple loci.

(a) *Load as a function of the gene targeted*

In the model analysed in § 2, the frequency of the HEG increases due to TRD, and decreases due to the homozygous lethality. For many species that one might want to control, killing males is worse than useless because it reduces the frequency of the HEG but will do little to reduce population growth rates or equilibrium density, which will largely be determined by female productivity. One way to avoid the 'wastage' of killing males would be to target a gene that, when knocked out, kills only females. For the TRD assumed here, targeting such a gene gives a three-fold reduction in mean fitness compared with targeting a recessive lethal gene (table 1). Knockouts causing females to be sterile will be equally effective. Indeed, knockouts causing male sterility will have the same effect, if they have no effect on the male's fertilization success (e.g. the only effect is to make sperm that are defective after karyogamy).

Targeting other classes of loci can be even more effective, though finding suitable candidate loci may be more difficult. If the knockout causes both sexes to be sterile, then mean fitness will be the square of what it is under lethality, because both parents have to be fertile in order for a zygote to be formed. For the TRD assumed here, this gives a five-fold reduction in mean fitness compared with targeting a lethal gene (table 1). If one targets a 'grandchildless' mutation (i.e. homozygous females produce viable but sterile sons and daughters; Ashburner

Table 1. Equilibrium mean fitness as a function of the class of gene targeted.

	$\hat{\bar{w}}$		
knockout phenotype	general	$e = 0.9$	$t_{1,90}$
lethal	$1 - e^2$	0.19	12
female lethality; unisexual sterility	$(1 - 4e^2)/(1 + 3e^2)$	0.055	11
bisexual sterility	$(1 - e^2)^2$	0.0361	11
maternal effect bisexual sterility ('grandchildless')	$[(1 - 4e^2)/(1 + 3e^2)]^2$	0.0031	12
conditional lethal[a]		0.16	

[a] Assumes five (summer) generations of no selection on the gene, followed by one (winter) generation of selection. Mean fitness in the overwintering generation will be $\hat{\bar{w}} = 1.6 \times 10^{-5}$.

1989), then mean fitness can be 60-fold lower than for a lethal gene (table 1). Finally, in some species one might be able to target loci that are essential only in some generations, creating a conditional lethal. This can give a modest increase in mean load compared with an unconditional lethal (table 1); more importantly, with density dependence, the effect on the population may be substantially greater (Knipling 1979).

(b) *Multiple loci*

Another means of increasing the load is to target multiple loci simultaneously. In the simplest case where the TRD and phenotypic effects at one locus are independent of genotype at the other locus, the equilibrium mean fitness will be the product of the fitnesses of the two loci separately. For example, if one engineers HEGs to target n different loci essential for female fertility and $e = 0.9$ at each, then mean fitness will be 0.055^n. If one targets five loci, then mean fitness will be $\hat{\bar{w}} = 5 \times 10^{-7}$, enough to drive any population extinct. Simulations show that it makes little difference whether recombination between the loci is 0 or 1/2, and, even if it is 0, it makes little difference whether the HEGs are introduced in coupling or repulsion (not shown). This is because the gene conversion events act analogously to recombination to break up correlations between loci, and so the alleles end up in linkage equilibrium. In *Drosophila melanogaster*, at least, there are thought to be some 3000 essential genes, and more than 100 required for fertility (Ashburner 1989, p. 435; Miklos & Rubin 1996; Ashburner *et al.* 1999). Most prospective target species are likely to have an abundance of suitable target loci.

4. DEALING WITH NATURAL RESISTANCE

The simulation shown in figure 3 demonstrates that if a functional host gene exists that is resistant to the HEG, then it will increase rapidly in frequency and drive the HEG extinct. Care must therefore be taken to minimize the likelihood that such sequences exist, or that they can arise before the population is eradicated. The first step would be to use mutagenesis experiments and structural studies to choose target genes, and sites within genes, that seem unlikely to be able to change (at the amino acid level) without seriously compromising the function. One would also want to choose sites that show little sequence variation in the target population. One could plausibly sequence 10^3–10^4 alleles, and if one engineered a HEG with some redundancy (as naturally occurring HEGs have) to recognize all sequence variants detected, one could thereby ensure that the initial frequency of resistant alleles was less than 1 in 10^3–10^4. Indeed, one might be able to go further, and engineer a HEG that could recognize all sequence variants actually detected, plus, say, all possible single-nucleotide variants of the observed sequences. Insertion or deletion mutations in the target site may be particularly difficult for the HEG to recognize, and so one will want to choose regions that are well conserved for length across species, or for which structural information suggests that any length variant is likely to be non-functional.

This is not to say that recognition-site redundancy should be maximized (or, put another way, that sequence specificity should be minimized). If the endonuclease cleaves non-homologous sites, then it will reduce fitness even when heterozygous, slowing or preventing its spread. Also, for safety, one will want to be able to release a resistant allele. Finally, it may also be safer if the endonuclease does not recognize the homologous sequence in closely related non-target species, so as to reduce the risk of horizontal transfer. Ideally, one wants to target a site that shows little variation within species, but considerable divergence between species (at least at the nucleotide level).

'Combination therapy'—the use of multiple drugs simultaneously—has recently been acknowledged to slow the evolution of resistance in human pathogens (White *et al.* 1999; Palumbi 2001), and the same approach can be used with engineered HEGs, virtually without limit. First, one could release, say, 10 different HEGs, attacking 10 different sites along the length of a single gene. The more HEGs that are released, the lower the likelihood that resistant sequences will exist for all of them, and, even if they do exist, the longer it will take for a multiply resistant sequence to be stitched together by recombination. The genetic load imposed in the meantime can be substantial, and may well be enough to drive the population extinct. One could also engineer two HEGs with adjacent recognition sites, such that if either one was able to cut, then both would be transmitted. Since alleles resistant to only one of the HEGs would have little selective advantage (arising only from non-functional mutant HEGs), the evolution of the doubly resistant allele would be substantially retarded. The only limit on the number of adjacent HEGs one could use would be in the length of sequence

that can be copied from one chromosome to another during recombinational repair.

The second form of 'combination therapy' would be to target multiple loci simultaneously. Even if resistance could evolve at each locus separately, in a combined attack the genetic load can be sufficient to drive the population extinct. As noted in § 3b, there is unlikely to be a shortage of suitable target loci. Suppose one is targeting n recessive female sterility genes and that at every one there is a resistant allele at a frequency of 10^{-6}. If release frequencies are 1% and all HEGs have $e = 0.9$, then there is a five-generation window (from generations 11 to 15 inclusive) in which the mean fitness of the population is about 0.08^n. With 10 HEGs, mean fitness will be about 10^{-11}, enough to drive any population extinct.

Two other classes of resistant genotypes are also possible, though whether either one is likely to arise in any real population is unclear. First, a mutation might arise that compensates for the knockout of the target gene but is otherwise neutral. Were such a mutation to arise, it too would increase rapidly in frequency, and population mean fitness would return to normal. A simple duplication of the target locus is unlikely to be sufficient, as the HEG will readily transfer over to the new locus. If such resistance does turn out to be a problem, then it will be one more reason to attack multiple genes simultaneously.

Finally, a mutation might arise that somehow reduces or eliminates the homing activity, but is otherwise neutral. Several points here seem relevant.

(i) Homing depends upon very basic cellular processes (transcription, translation, nuclear transport, recombinational repair), and so it is not clear how such a mutation might arise.
(ii) Many prospective target species do not appear to have HEGs naturally, and so are unlikely to have evolved general defences against them.
(iii) Naturally occurring HEGs fall into three or four distinct protein families (Chevalier & Stoddard 2001) and artificial HEGs can be different again (e.g. fusions of a sequence-specific zinc finger protein with a non-specific endonuclease domain; Bibikova *et al.* 2001), so if resistance evolves against one of them, there may not be cross-resistance to others.
(iv) Site-specific selfish genes are also available that use reverse transcription rather than recombinational repair to propagate, including group II introns and some LINE-like transposable elements, and cross-resistance to these is highly unlikely.

In conclusion, the unthinking use of HEGs for population control and eradication may lead to the evolution of resistance, as for any other method of pest management. However, numerous strategies exist for minimizing this likelihood. At the very least, it is not obvious *a priori* that resistance will inevitably evolve.

5. PREVENTING HORIZONTAL TRANSMISSION

Horizontal transmission between species has been demonstrated for HEGs in plant mitochondria and in yeast mitochondria and nuclei, and probably occurs in all taxa in which HEGs exist (Cho *et al.* 1998; Goddard & Burt 1999; Koufopanou *et al.* 2002). However, demonstration that horizontal transmission occurs regularly on an evolutionary time-scale of millions of years does not mean it is a substantial risk during a 10-year population control programme. Transposable elements have been demonstrated to transfer readily between *Drosophila* species (Jordan *et al.* 1999), but the likelihood of a HEG transferring is probably substantially lower, as they have no extrachromosomal part of the life cycle and no time when the protein is bound to the gene. For horizontal transmission to occur, one needs the DNA containing the HEG somehow to get into a germline nucleus of another species, and be sufficiently intact that it can be transcribed and then used as a template for repair. This is unlikely to occur in many prospective target species, and, indeed, may be the main reason why HEGs appear to be absent from animals with segregated germlines. As suggested in § 4, one way to reduce further the probability of horizontal transfer is to engineer the HEG not to recognize the homologous sequence in related species. Indirect evidence from yeasts suggests this can be an effective way to limit transfer (Koufopanou *et al.* 2002). In addition, one could target a region with low overall nucleotide similarity between species (to reduce the likelihood of homologous recombination), and have the HEG regulated by a species-specific promoter.

6. POPULATION GENETIC ENGINEERING

I have focused thus far on the problem of imposing a genetic load so as to control or eradicate a population. More subtle approaches may often be desirable. In particular, one may not want to eradicate a population, but rather to transform it genetically such that it is less noxious. For example, one might use engineered HEGs to perform a population-wide knockout of, say, a gene necessary for mosquitoes to transmit malaria. If the knockout is not too harmful, then the HEG will spread to fixation, and this may have little or no effect on population numbers. Indeed, if the homozygous knockout has fitness greater than $1/(e + 1)$, then the HEG is expected to spread to fixation even if its expression is not limited to the germline (calculations not shown). Many (perhaps most) genes could be knocked out in this way. Such an approach also has the advantage that resistant genotypes will be less strongly selected.

Rather than knocking out a gene, one might instead want to change some specific aspect of it (e.g. its promoter). If the gene was essential, so that there was strong selection against the knockout, then one could perform a population-wide gene replacement by engineering a HEG that attacked the unwanted sequence but not the desired one, and then simultaneously releasing individuals carrying the HEG and the resistant allele into the population. As shown in figure 4, the resistant allele comes to predominate relatively quickly after introduction, with only a relatively small and temporary reduction in mean fitness.

Finally, one might also want to drive a novel gene into a population—to perform a population-wide transformation. To achieve this, the novel gene could be linked to a HEG targeting a neutral region of the genome, and the whole construct inserted into the recognition site. As the

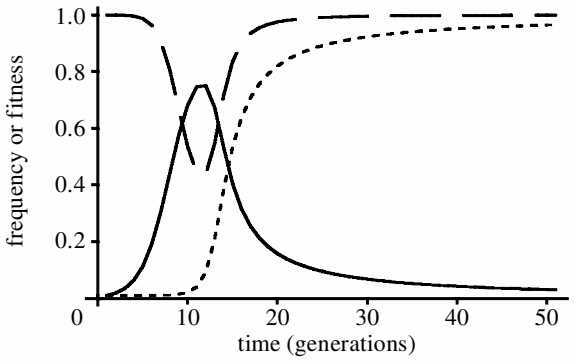

Figure 4. Population-wide gene replacement. The HEG and the resistant allele are introduced simultaneously at a 1% frequency. All other parameters are as in figure 3.

HEG spread through the population, it would bring the novel gene with it. Other strategies are also possible. For example, a HEG might be engineered to target an essential gene, and the novel gene linked to a resistant allele, possibly using an inversion. This would greatly expand the size of the gene one could introduce, even allowing multiple genes to be introduced simultaneously. It would also reduce the rate at which non-functional mutant genes arise, if DNA replication associated with cell division has a lower error rate than that associated with DNA repair and gene conversion.

Regardless of how exactly the novel gene is driven into the population, a key limitation of this approach is that, if it is harmful, the construct will not be evolutionarily stable. Mutant constructs in which the novel gene is deleted or otherwise defective will be selected, because they cause less harm to the host but still have the transmission advantage. However, if the novel gene is not too harmful, and the mutation rate is not too high, then it can persist for a considerable length of time before going extinct. Suppose, for example, the resistant allele in figure 4 has linked to it a novel gene that is fully dominant, reduces fitness by 10% and mutates to a non-functional form with a frequency of 10^{-6}. This novel gene would have a frequency greater than 50% within 15 generations and greater than 95% within 40 generations, and would remain above 95% for about 4000 generations (not shown). Full dominance is critical here, for then heterozygous mutants have no selective advantage. Introducing multiple copies of the gene, linked to one or more resistant alleles, should also help to reduce selection in favour of the non-functional mutants and prolong the transformation.

Some authors have suggested using transposable elements or cytoplasmic incompatibility agents as vectors to drive novel genes through a population (Ribeiro & Kidwell 1994; Turelli & Hoffmann 1999), but the use of HEGs is likely to have a number of advantages. Constructs using transposable elements are likely to be less stable, owing to the high mutation rate during transposition, and give less precise control over genomic location and copy number. Constructs using cytoplasmic incompatibility agents are likely to spread more slowly and require larger introduction frequencies and do not allow the gene to be in the nucleus. Both alternatives are perhaps more likely to transfer the novel gene to another species than is a HEG construct, because at no time is the latter separate from the host chromosome. Introducing the novel gene linked to a resistance locus should be even safer. Finally, transformations using a HEG should be fully reversible, by releasing a HEG engineered to target the novel gene.

7. OTHER USES

Two other potential uses for engineered HEGs can briefly be mentioned. First, it has long been recognized that if a Y chromosome were to show TRD and spread to fixation in a population, then the sex ratio would become male biased, and if the TRD was extreme, then the population could be driven extinct for want of females (Hickey & Craig 1966; Hamilton 1967). In *Aedes* and *Culex* mosquitoes, there are Y chromosomes that somehow cause the X chromosome to break during the first meiotic division, and thus show TRD (Newton *et al.* 1976; Sweeny & Barr 1978). Though the molecular mechanism is not yet known (the breaks may represent failed attempts at crossing-over), these observations suggest the following strategy. Insert onto a Y chromosome one or more endonuclease genes that recognize and cut sequences specific to the X chromosome, and have them under the control of pre-meiotic-specific promoters. Then, during spermatogenesis, the X chromosome would be cut, and, as there would not be an appropriate template for repair, the Y chromosome would show TRD. It would spread through the population, and if the TRD was sufficiently extreme, the population could be eradicated. Such an approach would not rely on recombinational repair or homing. In some species the sex chromosomes are inactivated prior to meiosis, which could complicate the design of the construct, but this is not so in all species, including many dipterans (McKee & Handel 1993).

Second, all the manipulations discussed thus far are 'inoculative', in that the release of relatively few engineered individuals will drive the population manipulation. Often this will be an advantage, but not always: if, for example, one wanted to eradicate only one population and leave others in the rest of the species range undisturbed, then inoculative methods may not be appropriate. 'Inundative' strategies such as the release of sterile males (Knipling 1979) are inherently self-limiting, and so more appropriate for such population-specific targeting. Engineered HEGs could be used in an inundative strategy if they caused dominant female lethality or sterility. Knockouts causing dominant female-specific effects are rare, but if the HEG was engineered to be constitutively active in all tissues, then even if a zygote started heterozygous, the organism would be converted to a homozygote. Thus, one could still target a recessive female-specific locus. Females inheriting the HEG would be dead or sterile, and males would pass on the HEG to the next generation. As long as the HEG was not perfectly efficient ($e < 1$), it would slowly disappear from the population, but could cause a substantial load before doing so. Thus, simply by changing the promoter, the threat posed by rare emigrants to neighbouring populations can be avoided. The use of such engineered HEGs would be more efficient than the release of sterile males, allowing either fewer individuals to be released, or larger populations to be targeted (Thomas *et al.* 2000; A. Burt, unpublished data).

8. PROSPECTS

The idea of driving a foreign gene into a natural population using a non-Mendelian genetic element has been much discussed, particularly in the context of rendering mosquitoes unable to transmit malaria (e.g. Ribeiro & Kidwell 1994; Turelli & Hoffmann 1999; Beerntsen *et al.* 2000; Ito *et al.* 2002). As we have seen, engineered HEGs may be useful tools for such population genetic engineering. However, this remains a relatively complex manipulation. The alternative approach, of using HEGs to perform a population-wide gene knockout, is much simpler, whether the goal is to modify the population or to eradicate it. The requisite construct consists of only one gene instead of two, and it is evolutionarily stable in the face of the most obvious mutations that will arise as it spreads through the population. Knockouts can also be more easily reversed, by releasing alleles resistant to the HEG. These advantages suggest that gene knockouts may be a more suitable approach to population genetic engineering than gene introductions, particularly in the first attempts.

Before any field trial can begin, further progress is needed in the basic molecular biology of site-specific selfish genes. First, further developments are needed in the re-targeting of these selfish genes to novel host sequences, though important progress has been made, both by design and by selection (Segal *et al.* 1999; Guo *et al.* 2000; Buchholz & Stewart 2001; Wilson *et al.* 2001; Chevalier *et al.* 2002; Santoro & Schultz 2002; Seligman *et al.* 2002; Takahashi & Fujiwara 2002). Second, it must be demonstrated that these elements can be made to work in the prospective target species: insects, for example, are natural hosts only for the LINE-like elements. Encouragingly, a yeast HEG has recently been shown to be active in stimulating recombinational repair in *D. melanogaster* (Bellaiche *et al.* 1999; Rong & Golic 2000; Rong *et al.* 2002).

In carrying out this work, it should be noted that the ease and rapidity with which these selfish genes can invade a population applies not just to planned releases, but also to unintentional releases of laboratory escapees. Proper attention to containment issues is needed to prevent natural populations of *D. melanogaster*, to take an obvious example, from being infected (and possibly endangered) by engineered selfish genes that were meant to stay in the laboratory. Finally, wide-ranging discussions are needed on the criteria for deciding whether to eradicate or genetically engineer an entire species. Clearly, the technology described here is not to be used lightly. Given the suffering caused by some species, neither is it obviously one to be ignored.

M. Ashburner, C. Godfray, B. Sinden and B. Trivers provided useful discussions and/or comments, and the Biosocial Research Foundation and the Wellcome Trust supplied financial support.

REFERENCES

Ashburner, M. 1989 *Drosophila: a laboratory handbook*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Ashburner, M. (and 25 others) 1999 An exploration of the sequence of a 2.9-Mb region of the genome of *Drosophila melanogaster*: the *Adh* region. *Genetics* **153**, 179–219.

Beerntsen, B. T., James, A. A. & Christensen, B. M. 2000 Genetics of mosquito vector competence. *Microbiol. Mol. Biol. Rev.* **64**, 115–137.

Belfort, M., Derbyshire, V., Parker, M. M., Cousineau, B. & Lambowitz, A. M. 2002 Mobile introns: pathways and proteins. In *Mobile DNA II* (ed. N. L. Craig, R. Craigie, M. Gellert & A. M. Lambowitz), pp. 761–783. Washington, DC: ASM Press.

Bellaiche, Y., Mogila, V. & Perrimon, N. 1999 I-*Sce*I endonuclease, a new tool for studying DNA double-strand break repair mechanisms in *Drosophila*. *Genetics* **152**, 1037–1044.

Bibikova, M., Carroll, D., Segal, D. J., Trautman, J. K., Smith, J., Kim, Y.-G. & Chandrasegaran, S. 2001 Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* **21**, 289–297.

Braig, H. R. & Yan, G. 2002 The spread of genetic constructs in natural insect populations. In *Genetically engineered organisms: assessing environmental and human health effects* (ed. D. K. Letourneau & B. E. Burrows), pp. 251–314. Boca Raton, FL: CRC Press.

Buchholz, F. & Stewart, A. F. 2001 Alteration of Cre recombinase site specificity by substrate-linked protein evolution. *Nature Biotechnol.* **19**, 1047–1052.

Chandrasegaran, S. & Smith, J. 1999 Chimeric restriction enzymes: what is next? *Biol. Chem.* **380**, 841–848.

Chevalier, B. S. & Stoddard, B. L. 2001 Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.* **29**, 3757–3774.

Chevalier, B. S., Kortemme, T., Chadsey, M. S., Baker, D., Monnat Jr, R. J. & Stoddard, B. L. 2002 Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell.* **10**, 895–905.

Cho, Y., Qiu, Y.-L., Kuhlman, P. & Palmer, J. 1998 Explosive invasion of plant mitochondria by a group I intron. *Proc. Natl Acad. Sci. USA* **95**, 14 244–14 249.

Curtis, C. F. 1985 Genetic control of insect pests: growth industry or lead balloon? *Biol. J. Linn. Soc.* **26**, 359–374.

Eickbush, T. H. 2002 R2 and related site-specific non-long terminal repeat retrotransposons. In *Mobile DNA II* (ed. N. L. Craig, R. Craigie, M. Gellert & A. M. Lambowitz), pp. 813–835. Washington, DC: ASM Press.

Goddard, M. R. & Burt, A. 1999 Recurrent invasion and extinction of a selfish gene. *Proc. Natl Acad. Sci. USA* **96**, 13 880–13 885.

Goddard, M. R., Greig, D. & Burt, A. 2001 Outcrossed sex allows a selfish gene to invade yeast populations. *Proc. R. Soc. Lond* B **268**, 2537–2542. (DOI 10.1098/rspb.2001.1830.)

Guo, H., Karberg, M., Long, M., Jones III, J. P., Sullenger, B. & Lambowitz, A. M. 2000 Group II introns designed to insert into therapeutically relevant DNA target sites in human cells. *Science* **289**, 452–457.

Hamilton, W. D. 1967 Extraordinary sex ratios. *Science* **156**, 477–488.

Hastings, I. M. 1994 Selfish DNA as a method of pest control. *Phil. Trans. R. Soc. Lond.* B **344**, 313–324.

Hickey, W. A. & Craig Jr, G. B. 1966 Distortion of sex ratio in populations of *Aedes aegypti*. *Can. J. Genet. Cytol.* **8**, 260–278.

Ito, J., Ghosh, A., Moreira, L. A., Wimmer, E. A. & Jacobs-Lorena, M. 2002 Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature* **417**, 452–455.

Jacquier, A. & Dujon, B. 1985 An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* **41**, 383–394.

Jordan, I., Matyunina, L. & McDonald, J. 1999 Evidence for the recent horizontal transfer of long terminal repeat retrotransposon. *Proc. Natl Acad. Sci. USA* **96**, 12 621–12 625.

Knipling, E. P. 1979 *The basic principles of insect population suppression and management*. Washington, DC: US Department of Agriculture.

Koufopanou, V., Goddard, M. R. & Burt, A. 2002 Adaptation for horizontal transfer in a homing endonuclease. *Mol. Biol. Evol.* **19**, 239–246.

McKee, B. D. & Handel, M. A. 1993 Sex chromosomes, recombination, and chromatin conformation. *Chromosoma* **102**, 71–80.

Miklos, G. L. G. & Rubin, G. M. 1996 The role of the genome project in determining gene function: insights from model organisms. *Cell* **86**, 521–529.

Newton, M. E., Wood, R. J. & Southern, D. I. 1976 A cytogenetic analysis of meiotic drive in the mosquito, *Aedes aegypti* (L.). *Genetica* **46**, 297–318.

Palumbi, S. 2001 Humans as the world's greatest evolutionary force. *Science* **293**, 1786–1790.

Ribeiro, J. M. C. & Kidwell, M. G. 1994 Transposable elements as population drive mechanisms: specification of critical parameter values. *J. Med. Entomol.* **31**, 10–16.

Rong, Y. S. & Golic, K. G. 2000 Gene targeting by homologous recombination in *Drosophila*. *Science* **288**, 2013–2018.

Rong, Y. S., Titen, S. W., Xie, H. B., Golic, M. M., Bastiani, M., Bandyopadhyay, P., Olivera, B. M., Brodsky, M., Rubin, G. M. & Golic, K. G. 2002 Targeted mutagenesis by homologous recombination in *D. melanogaster*. *Genes Dev.* **16**, 1568–1581.

Santoro, S. W. & Schultz, P. G. 2002 Directed evolution of the site specificity of Cre recombinase. *Proc. Natl Acad. Sci. USA* **99**, 4185–4190.

Segal, D. J., Dreier, B., Beerli, R. R. & Barbas III, C. F. 1999 Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5′-GNN-3′ DNA target sequences. *Proc. Natl Acad. Sci. USA* **96**, 2758–2763.

Seligman, L. M., Chisholm, K. M., Chevalier, B. S., Chadsey, M. S., Edwards, S. T., Savage, J. H. & Veillet, A. L. 2002 Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res.* **30**, 3870–3879.

Spielman, A., Beier, J. C. & Kiszewski, A. E. 2002 Ecological and community considerations in engineering arthropods to suppress vector-borne disease. In *Genetically engineered organisms: assessing environmental and human health effects* (ed. D. K. Letourneau & B. E. Burrows), pp. 315–329. Boca Raton, FL: CRC Press.

Sweeny, T. L. & Barr, A. R. 1978 Sex ratio distortion caused by meiotic drive in a mosquito, *Culex pipiens*. L. *Genetics* **88**, 427–446.

Takahashi, H. & Fujiwara, H. 2002 Transplantation of target site specificity by swapping the endonuclease domains of two LINEs. *EMBO J.* **21**, 408–417.

Thomas, D. D., Donnelly, C. A., Wood, R. J. & Alphey, L. S. 2000 Insect population control using a dominant, repressible, lethal genetic system. *Science* **287**, 2474–2476.

Turelli, M. & Hoffmann, A. A. 1999 Microbe-induced cytoplasmic incompatibility as a mechanism for introducing transgenes into arthropod populations. *Insect Mol. Biol.* **8**, 243–255.

Wenzlau, J., Saldanha, R., Butow, R. & Perlman, P. 1989 A latent intron-encoded maturase is also an endonuclease needed for intron mobility. *Cell* **56**, 421–430.

White, N. (and 16 others) 1999 Averting a malaria disaster. *Lancet* **353**, 1965–1967.

Wilson, D. S., Keefe, A. D. & Szostak, J. W. 2001 The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl Acad. Sci. USA* **98**, 3750–3755.

As this paper exceeds the maximum length normally permitted, the author has agreed to contribute to production costs.