Trends in Ecology & Evolution 3: S18-S20 (1988)

STABILITY OF RECOMBINANT DNA AND ITS EFFECTS ON FITNESS

RICHARD E. LENSKI AND TOAI T. NGUYEN

The stability of recombinant DNA is influenced by the fidelity of its genetic transmission and by its effects on fitness of the engineered organism. According to the 'excess baggage' hypothesis, environmental applications of engineered organisms are inherently safe because these organisms will disappear in the absence of selection for their intended functions, owing to the costs of carriage and expression of the recombinant DNA. There are many examples that support this hypothesis, but there are also some interesting and important exceptions.

According to many proponents, the release of genetically engineered organisms into the environment for biotechnological purposes is inherently safe. This safety is presumed to derive from the 'costs' of carriage and expression of recombinant DNA, which reduce the fitness of engineered organisms relative to their wild-type counterparts, thereby preventing the unintended spread of engineered organisms[1,2]. We shall refer to this argument as the 'excess baggage' hypothesis.

The excess baggage hypothesis has antecedents in evolutionary biology extending at least to Charles Darwin. In *The Origin of Species*, Darwin states: 'I think there can be little doubt that use . . . strengthens and enlarges certain parts, and disuse diminishes them; and that such modifications are inherited . . . '. Darwin proceeds to discuss flightless birds and wingless beetles, burrowing rodents with rudimentary eyes and blind cave-dwelling creatures. His discussion is remarkable, in that he considers not only the effects of selection acting against disused characters, but also the effects of the *absence* of selection on these characters. Although Darwin was not aware of the genetic mechanisms leading to the loss of a character, we now know that mutations which destroy a character will accumulate in the absence of selection for that character's function. In our discussion of the stability of recombinant DNA, we must consider not only its effects on fitness, but also the fidelity of its replication and transmission, as this too will influence its fate.

Richard Lenski and Toai Nguyen are at the Department of Ecology & Evolutionary Biology, University of California, Irvine, CA 92717, USA.

Segregation and selection

Imagine that some engineered character is so unstable as to defeat its intended application. If it is assumed naively that instability arises because of genetic infidelity, then one might reasonably increase the number of copies of the recombinant DNA, thereby increasing the likelihood that each offspring receives a functional copy of the engineered gene. But if the instability is due largely to intense selection against expression of the recombinant gene product, then an increase in copy number may amplify this selection, further aggravating the instability. Models derived from the field of population genetics can be used to distinguish the effects of segregation (i.e. genetic infidelity) and selection (i.e. differences in fitness) on the stability of recombinant DNA (Box 1).

Segregation is likely to be especially important when recombinant DNA is carried on extrachromosomal elements, like plasmids, whose replication and transmission is more or less independent of the rest of the genome. The likelihood of segregation of extrachromosomal elements depends on several factors, including their rate of replication, their copy number, and the presence or absence of *par* functions, which are presumed to increase stability by equalizing the partitioning of copies to daughter cells[7–10]. Even when there is selection against carriage of an extrachromosomal element, *par* functions can be used to reduce the rate at which segregants arise *de novo*[11].

Energetic burden and physiological disruption

There are two distinct problems that a genetically engineered organism may face. First, there is the energetic burden of synthesizing additional macromolecules, including nucleic acids and proteins. Secondly, there may be disruption of normal physiological processes caused by the expression of novel gene products. (In fact, these same two effects may be manifest as the result of any genetic change, whether engineered or not.)

DaSilva and Bailey[12] have estimated that an *E. coli* cell harboring 50 copies of plasmid pBR322 (4.4 kb) and expressing the cloned gene product as 20% of its total cell protein would require about 0.1% and 13% more ATP for plasmid DNA and protein syntheses, respectively.

© 1988 Elsevier Publications Cambridge 0169-5347/88/$02.00

From these energetic considerations alone, one would expect selection to favor plasmid-free segregants over plasmid-bearing cells, at least in the absence of selection for the cloned gene product. Zund and Lebek[13] surveyed naturally occurring antibiotic resistance plasmids for their effects on the generation length of an *E. coli* host, and they found that a substantial proportion of these plasmids increased the doubling time by more than 15%. However, the increased generation length was not strongly correlated with a plasmid's size, nor with its copy number. Thus, the energetic burden associated with the additional DNA appears to be of only minor importance, supporting the calculations of DaSilva and Bailey. The longer doubling times presumably reflect the effects of expression of plasmid-encoded proteins; however, it is not clear whether the energetic burden of protein synthesis or the associated disruption of physiological processes is the most important.

In certain cases, it is clear that energetic calculations can greatly *underestimate* the disadvantage associated with expression of unnecessary proteins. Lwoff[14] suggested that selection favors auxotrophic mutants (which cannot synthesize some amino acid) over their prototrophic progenitors in medium supplemented with an ample supply of the amino acid, owing to the energetic savings that accrue to the auxotrophs. Several experiments have demonstrated that auxotrophs do, in fact, have a selective advantage under such conditions[15]. However, Dykhuizen[16] has shown that the magnitude of the observed advantage for tryptophan auxotrophs is far greater than can be explained solely on the basis of energetic efficiency. Dykhuizen hypothesizes that there is disruption of some physiological process caused by metabolic intermediates that are present only in the prototrophic strain, but the precise nature of the discrepancy between the observed selection differential and that calculated from energetic savings is not known.

The physiological disruption associated with the expression of an unnecessary gene product can sometimes be quite extreme. Bassford *et al.*[17] found that *E. coli* cells induced for synthesis of a malE-lacZ hybrid protein were severely impaired. This impairment is thought to occur because the hybrid molecules become physically 'stuck' in the cytoplasmic membrane, thereby occupying sites that are essential for the transport of periplasmic and outer membrane proteins[18]. Moyed *et al.*[19] observed that induction of the Tn*10* tetracycline resistance determinant, when carried on a multicopy plasmid, severely inhibits the growth of its *E. coli* host. This inhibition apparently occurs because the resistance protein has some disruptive effect on cell physiology; this effect may be associated with the protein's efflux activity (with respect to tetracycline and possibly other metabolites), or it may result from interference by the protein with the functioning of the cytoplasmic membrane. Interestingly, Moyed *et al.* also demonstrated that the level of phenotypic resistance to tetracycline actually decreases when there are many copies of the resistance gene. This work therefore illustrates the potential complexity of the relationship between gene expression and phenotypic characters.

Box 1. Mathematics of segregation and selection

For simplicity, we will consider an engineered organism that is haploid and reproduces clonally, so we need consider only two genotypes. Two distinct processes may contribute to the instability of an engineered gene. Segregation (including mutation) occurs when there are failures in gene replication or transmission. Selection may occur when there are differences between the genotypes in survival or reproduction. Segregation is a decay process, and therefore its effect depends on the frequency of the engineered genotype. In contrast, the rate of change due to selection depends upon the product of the two genotype frequencies. As shown by Fisher[3], the response to selection is proportional to the genetic variation in fitness, which is greatest at intermediate frequencies of the two genotypes. From these basic principles, one can derive an equation that describes the population dynamics of an engineered gene that is subject to both segregation and selection.

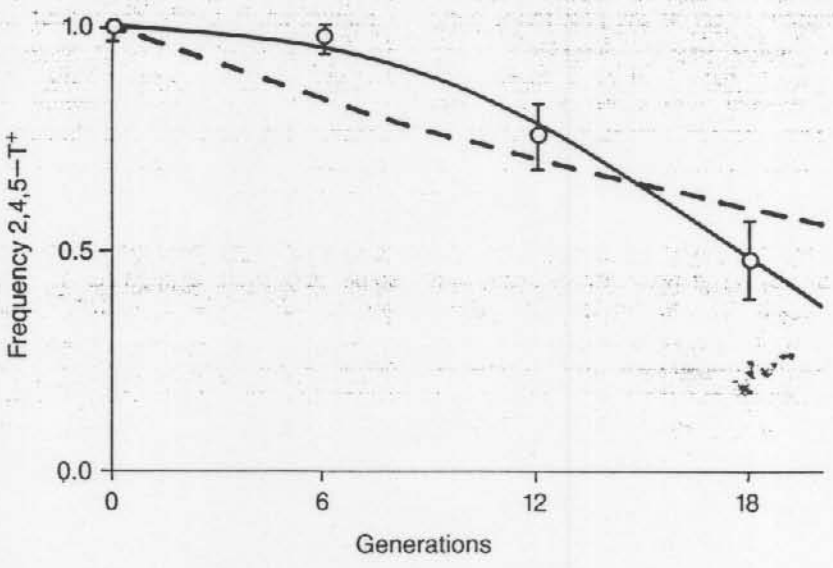
Let p be the frequency of the engineered genotype and $q = 1 - p$ be the frequency of segregants, which have lost the engineered gene. Let u be the segregation rate, and let s be the selection coefficient. The rate of change in the frequency of the engineered genotype is given by:

$$dp / dt = -up - spq \qquad (1)$$

where segregants are more fit than the engineered genotype when $s > 0$.

How can one determine whether the instability of some engineered genotype is due to the effect of segregation alone or to the combined effects of segregation plus selection? Appropriate statistical procedures can be used to estimate the segregation rate and the selection coefficient that minimize the sum of the squared deviations of the fitted model about a series of sample frequencies taken at several points in time[4,5]. We will briefly illustrate these procedures using data on the instability of plasmid-borne genes encoding metabolic functions that permit the bacterium *Pseudomonas cepacia* to degrade 2,4,5-Trichlorophenoxyacetic acid (2,4,5-T), a toxic and environmentally persistent compound in the herbicide Agent Orange.

Kilbane *et al.*[6] scored the frequency of 2,4,5-T$^+$ individuals during 18 generations of growth in the absence of 2,4,5-T (see Figure). In each sample, 150 individuals were scored; we have calculated the 95% confidence interval for each sample frequency from the binomial distribution. If one assumes that the instability of the 2,4,5-T$^+$ genotype is due entirely to segregation, then the dashed line ($u = 0.0292$ and $s = 0$) provides the best fit of equation 1 to the experimental data. However, there are clearly too few segregants early in the experiment, and too many segregants late in the experiment, to be explained by segregation alone. The solid line gives the best fit when both segregation and selection parameters are allowed to vary ($u = 0.0047$ and $s = 0.2155$); it provides a much better fit to the data. Therefore, one can conclude that intense selection greatly amplified the frequency of segregants arising *de novo*.



Evolution and coevolution

The selective disadvantage associated with a gene, whether due to an

energetic burden or to the disruption of some physiological process, may be diminished by subsequent evolution. This could result from a mutation reducing the level of expression of the costly gene product, or it could result from a mutation, elsewhere in the genome, that compensates for the disruption of a critical physiological process[20]. Lenski and Levin[21] found that *E. coli* mutants resistant to the virus T4 are less fit than their sensitive progenitors when the two genotypes are allowed to compete in the absence of the virus. However, the selective disadvantage associated with T4-resistance can be overcome, at least in part, by additional mutations that increase competitive fitness without diminishing resistance[22]. (These additional mutations are as yet undefined, and the physiological basis of their function is unknown.)

Similarly, McKenzie *et al.*[23] have shown that a selective disadvantage in the sheep blowfly (*Lucilia cuprina*) associated with a mutation conferring resistance to the insecticide diazinon was reduced or eliminated by subsequent evolution of the blowfly. The genes causing this effect are not linked to the resistance gene, and their products' mechanisms of action are unknown, but they mitigate the disruptive effect of the resistance gene on the blowfly's development[24].

There have also been several experimental demonstrations of unexpected fitness advantages associated with the carriage of foreign DNA in bacteria. Hartl *et al.*[25,26] found that the transposon Tn*5* (and its component IS*50* insertion sequences) enhances the growth rate of some (but not all) *E. coli* hosts. This effect requires a functional gene product (either a transposase or an associated inhibitor), but it is not mediated by transposition events *per se*[25]. Several researchers have also shown that carriage of integrated viral genes can unexpectedly enhance the growth rate of infected bacterial hosts, at least under certain culture conditions[27–29]. Edlin *et al.*[30] demonstrated that one region of the Lambda virus genome, when cloned into a plasmid, enhanced the fitness of these plasmid-bearing cells relative to cells carrying plasmids that lacked the viral DNA.

It has been suggested that these fitness-enhancing functions represent vestiges of mutualistic coevolution between accessory genetic elements and their hosts[25,31]. On theoretical grounds, accessory elements that have limited opportunities for infectious transfer between hosts are especially likely to be selected for functions that enhance host fitness[31]. Bouma and Lenski (unpublished) examined the evolution of an association between a non-transmissible plasmid and its *E. coli* host. Initially, this association was mutualistic in the presence of antibiotic, but antagonistic in the absence of antibiotic. After 500 generations, the association had evolved into one that was mutualistic in both of these environments. Interestingly, the genetic change responsible for this transition occurred in the genome of the bacterial host, and not in the plasmid.

These results indicate that the selective disadvantage associated with novel gene functions may be quite labile. In particular, the magnitude of a disadvantage may depend on the genetic background in which the novel function is expressed, and that background may change with further evolution. Natural selection continues to act on genetically engineered organisms, and this process can be expected to increase the fitness of an engineered organism whether or not such an increase is in the best interest of the genetic engineer.

Conclusions

Mathematical models indicate that the stability of an engineered gene in a population is highly dependent on the fitness of individuals possessing that gene relative to segregants that have lost the gene. In the absence of selection for some specific character encoded by the engineered genetic material, recombinant DNA can be expected to impose a reduction in fitness, although there are apparently some important exceptions to this generalization. The magnitude of the reduction in fitness, if any, depends primarily on the energetic burden associated with synthesis of the recombinant gene product (and not of the recombinant DNA itself), and on any disruptive effects of the gene product on important physiological processes. In general, we expect that the selective disadvantage associated with some engineered gene will be greatest when recombinant DNA is newly introduced, and may be diminished by subsequent evolution.

Elucidating the factors that influence the stability and fitness of genetically engineered organisms will help to guard against their inadvertent spread. Such research may also indicate new ways in which to enhance the stability and fitness of engineered organisms that might otherwise not persist long enough to perform their intended functions[30]. Finally, this research promises to bring about the further integration of molecular approaches into the study of evolutionary mechanisms; in particular, we can look forward to many new results concerning the genetic basis of fitness in laboratory populations and in nature.

Acknowledgments

This review was prepared while the authors were supported by a research grant (CR-813401) from the Gulf Breeze Laboratory of the US Environmental Protection Agency. Michael Rose provided valuable comments and discussion.

References

1 Brill, W.J. (1985) *Science* 227, 381–384
2 Regal, P.J. (1986) in *Ecology of Biological Invasions of North America and Hawaii* (Mooney, H.A and Drake, J.A., eds), pp. 111–129, Springer-Verlag
3 Fisher, R.A. (1930) *The Genetical Theory of Natural Selection*, Oxford University Press
4 Lenski, R.E. and Bouma, J.E (1987) *J. Bacteriol.* 169, 5314–5316
5 Cooper, N.S., Brown, M.E. and Caulcott, C.A. (1987) *J. Gen. Microbiol.* 133, 1871–1880
6 Kilbane, J.J., Chatterjee, D.K., Karns, J.S., Kellogg, S.T. and Chakrabarty, A.M. (1982) *Appl. Environ. Microbiol.* 44, 72–78
7 Tucker, W.T., Miller, C.A. and Cohen, S.N. (1984) *Cell* 38, 191–201
8 Meacock, P.A. and Cohen, S.N. (1980) *Cell*, 20, 529–542
9 Miki, T., Easton, A.M. and Rownd, R.H. (1980) *J. Bacteriol.* 141, 87–99
10 Nordstrom, K., Molin, S. and Aargaard-Hansen, H. (1980) *Plasmid* 4, 215–227
11 Lee, S.W. and Edlin, G. (1985) *Gene* 39, 173–180
12 DaSilva, N.A. and Bailey, J.E. (1986) *Biotech. Bioeng.* 28, 741–746
13 Zund, P. and Lebek, G. (1980) *Plasmid* 3, 65–69
14 Lwoff, A. (1944) *L'Evolution Physiologique*, Herman & Cie
15 Diamond, J.M. (1986) *Nature* 321, 565–566
16 Dykhuizen, D. (1978) *Evolution* 32, 125–150
17 Bassford, P.J., Jr., Silhavy, T.J. and Beckwith, J.R. (1979) *J. Bacteriol.* 139, 19–31
18 Ito, K., Bassford, P.J., Jr. and Beckwith, J.R. (1981) *Cell* 24, 707–717
19 Moyed, H.S., Nguyen, T.T. and Bertrand K.P. (1983) *J. Bacteriol.* 155, 549–556
20 Uyenoyama, M.K. (1986) in *Pesticide Resistance: Strategies and Tactics for Management*, pp. 207–221, National Academy Press
21 Lenski, R.E. and Levin, B.R. (1985) *Am. Nat.* 125, 585–602
22 Lenski, R.E. *Evolution* (in press)
23 McKenzie, J.A., Whitten, M.J. and Adena, M.A. (1982) *Heredity* 49, 1–9
24 Clarke, G.M. and McKenzie, J.A. (1987) *Nature* 325, 345–346
25 Hartl, D.L., Dykhuizen, D.E., Miller, R.D., Green, L. and DeFramond, J. (1983) *Cell* 35, 503–510
26 Biel, S.W. and Hartl, D.L. (1983) *Genetics* 103, 581–592
27 Lin, L., Bitner, R. and Edlin, G. (1977) *J. Virol.* 21, 554–559
28 Edlin, G., Lin, L. and Bitner, R. (1977) *J. Virol.* 21, 560–569
29 Dykhuizen, D.E., Campbell, J.H. and Rolfe, B.G. (1978) *Microbios* 23, 99–113
30 Edlin, G., Tait, R.C. and Rodriguez, R.L. (1984) *Biotechnology* 2, 251–254
31 Levin, B.R. and Lenski, R.E. (1983) in *Coevolution* (Futuyma, D.J. and Slatkin, M., eds), pp. 99–127, Sinauer