Cryptic genes are silenced genes that can still be reactivated by mutation. Since they can make no positive contribution to the fitness of their carriers, it is not clear why many cryptic genes in microbial populations have not degenerated into useless DNA sequences. Hall et al. (1983) have suggested that cryptic genes have persisted because of occasional strong environmental selection for reactivated genes. The present mathematical study supports their suggestion. It shows that a cryptic gene can be retained without having any selective advantage over a useless DNA sequence, if selection for the reactivated gene occasionally occurs for a substantially long time.

Introduction

A cryptic gene is a silenced gene that is a single nucleotide substitution, that is present at a high frequency in a population, and that can still be reactivated by mutation, recombination, insertion, deletion, or other genetic mechanisms (Hall et al. 1983). A nonfunctional gene is defined as a silenced gene that cannot be reactivated by a single mutational event. A cryptic gene cannot make a positive contribution to the fitness of its carriers, and it may become nonfunctional by the pressure of mutation causing further degeneration or deletion. Yet cryptic genes appear to be widespread in microbial populations (Hall et al. 1983). To explain this phenomenon, Hall et al. (1983) have proposed a model in which the cryptic state is advantageous under one set of environmental conditions, whereas the functional state is favored under another set of environmental conditions. In their mathematical treatment, however, they have not taken into account the environmental change but have assumed constant fitnesses for all three states. Under this assumption, the cryptic gene (and its functional allele) can be retained in the population only if the fitness of the functional allele is higher than those of the cryptic and nonfunctional alleles, or if the fitness of the cryptic allele is higher.
than that of the nonfunctional allele. Under the former condition, the frequency of the functional allele would usually be high—a situation contrary to the observation that the frequency is usually very low. The latter condition does not seem realistic because the only difference between a cryptic and a nonfunctional allele is that the latter cannot be reactivated by a single mutation. In the following, I shall remove the assumption of constant fitnesses and show that neither of the above two conditions is necessary for the retention of cryptic genes in a population.

**Mathematical Theory**

**Absence of Mutation**

Hall et al. (1983) used $A_1$, $A_2$, and $A_3$ to denote the cryptic, the functional, and the nonfunctional genes, respectively. It is more natural to denote the functional gene by $A_1$ and the cryptic gene by $A_2$ because the functional gene was probably the original allele and because the cryptic state is the intermediate state between the other two states. Let $m_1(t)$, $m_2(t)$, and $m_3(t)$ be the Malthusian fitnesses of $A_1$, $A_2$, and $A_3$ at time $t$, respectively. The first problem to be understood is how fast the gene frequencies can change under different environments. For simplicity, let us assume that $q(t) = m_1(t)$. Let $x(t)$ be the frequency of $A_1$ at time $t$ and a selective difference $s(t) = m_1(t) - q(t)$. It can then be shown that

$$
\frac{dx}{dt} = s(t) x(1-x)
$$

or

$$
\ln \frac{x(t)}{1-x(t)} - \ln \frac{x(0)}{1-x(0)} = \int_0^t s(\tau) d\tau.
$$

(see Crow and Kimura 1970, pp. 190–192; Nagylaki 1975). As $\ln[x/(1-x)]$ is an increasing function of $x$, $x(t) \geq x(0)$ if the integral is zero or positive, and $x(t) < x(0)$ if otherwise.

Let us assume that there are two types of environments; in environment 1, $s(t) = s_1$, and in environment 2, $s(t) = -s_2$. We first consider the increase in $x$ in environment 1. In this case,

$$
\ln \frac{x(t)}{1-x(t)} - \ln \frac{x(0)}{1-x(0)} = s_1 t.
$$

(2)

It will be seen from table 1 that, if $s_1$ is large, the increase in $x$ is extremely rapid. Indeed, if $s_1 = 0.1$, it takes only 92 generations for $x$ to increase from 0.01 to 0.99. If $s_1 = 0.001$, the time required is 100 times longer. This is, however, still a relatively short period, for the generation time in microorganisms is very short.

### Table 1

**Number of Generations Required for the Frequency of $A_1$ to Increase from $y$ to $z$**

<table>
<thead>
<tr>
<th>$s$</th>
<th>$y = 10^{-6}$, $z = .01$</th>
<th>$y = .01$, $z = .99$</th>
<th>$y = .99$, $z = .999999$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.1</td>
<td>92.2</td>
<td>91.9</td>
<td>92.2</td>
</tr>
<tr>
<td>.001</td>
<td>9,220</td>
<td>9,190</td>
<td>9,220</td>
</tr>
</tbody>
</table>

**Note.**—Assuming that $A_1$ has an advantage of $s$ over $A_2$ and $A_3$. 
In environment 2 the frequency of $A_1$ will decrease. If $s_2 = 0.001$, it takes only 9,190 generations for $x$ to decrease from 0.99 to 0.01.

When the environment varies with time, we can write $t = t_1 + t_2$, where $t_i$ is the number of generations the environment is of the $i$th type. We can then write equation (1) as

$$\ln \frac{x(t)}{1-x(t)} - \ln \frac{x(0)}{1-x(0)} = s_1 t_1 - s_2 t_2. \tag{3}$$

One simple situation is that the environment varies cyclically. Let $T$ be the period, and assume that in each cycle $T$, generations are in environment $i$ so that $T = T_1 + T_2$. If $s_1 T_1 - s_2 T_2$ is zero, $x$ will vary cyclically with period $T$, but if the difference is positive (negative) $x$ will increase (decrease) with the number of cycles. As an example, assume $s_1 = 0.1$, $s_2 = 0.001$, $T_1 = 10$, $T_2 = 990$, and $T = 1,000$. Then $s_1 T_1 - s_2 T_2 = 0.01$, and it will take 919 cycles or 919,000 generations for $x$ to increase from 0.01 to 0.99.

Presence of Mutation

To understand the retention of cryptic genes in a population, we must also consider mutation. Let $u_{ij}$ be the mutation rate per generation from $A_i$ to $A_j$. It can be shown that

$$\frac{dx_1}{dt} = [m_1(t) - m_2(t)]x_1 x_2 + [m_1(t) - m_3(t)]x_1 x_3$$

$$- (u_{12} + u_{13})x_1 + u_{21} x_2 + u_{31} x_3,$$  \tag{4}

$$\frac{dx_2}{dt} = - [m_2(t) - m_1(t)]x_1 x_2 + [m_2(t) - m_3(t)]x_2 x_3$$

$$- (u_{21} + u_{23})x_2 + u_{12} x_1 + u_{32} x_3,$$  \tag{5}

$$\frac{dx_3}{dt} = - [m_3(t) - m_1(t)]x_1 x_3 - [m_3(t) - m_2(t)]x_2 x_3$$

$$+ u_{12} x_1 + u_{32} x_2,$$  \tag{6}

where $x_1$, $x_2$, and $x_3$ are the frequencies of $A_1$, $A_2$, and $A_3$ at time $t$. In the above equations, I assumed that reversible mutation cannot occur from $A_3$ to $A_1$ or $A_2$, that is, $u_{31} = u_{32} = 0$. Since $x_1 + x_2 + x_3 = 1$, it suffices to consider only two of the three equations. These equations are equivalent to equation (1) of Hall et al. (1983), except that I used different notations and Malthusian instead of Wrightian fitnesses.

We are particularly interested in the following question: Can the cryptic allele be retained without having a selective advantage over the nonfunctional allele? To answer this question, we assume that $m_1(t) = m_2(t)$. The second term on the right-hand side of equations (5) and (6) then disappears. It seems from equation (6) that, unless $m_3$ is on the average larger than $m_1$ ($m_2$), $x_3$ will eventually increase to 1, and $A_1$ and $A_2$ will be eliminated. It turns out that this condition is not necessary for the retention of $A_1$ and $A_2$. (As mutation can occur between $A_1$ and $A_2$, if one of them is retained, the other is also retained.)

Table 2 shows some numerical examples that were obtained by iteration on a computer, replacing equations (5) and (6) by difference equations (see Discussion for the parameter values used). In all examples I assume that the environment...
Table 2
Changes in Gene Frequency under Selection and Mutation

<table>
<thead>
<tr>
<th>Time</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_1 = 200$, $u_{12} = u_{13} = u_{23} = 10^{-5}$</td>
<td>$T_1 = 200$, $u_{12} = u_{23} = 10^{-6}$, $u_{13} = 10^{-7}$</td>
<td>$T_1 = 200$, $u_{12} = u_{23} = 10^{-6}$, $u_{13} = 10^{-7}$</td>
<td>$T_1 = 100$, $u_{12} = u_{23} = 10^{-6}$, $u_{13} = 10^{-7}$</td>
<td>$T_1 = 150$, $u_{12} = u_{23} = 10^{-6}$, $u_{13} = 10^{-7}$</td>
</tr>
<tr>
<td>$x_1$</td>
<td>0.99975</td>
<td>0.99986</td>
<td>0.99743</td>
<td>0.99743</td>
<td>0.99111</td>
</tr>
<tr>
<td>$x_2$</td>
<td>0.00012</td>
<td>0.00012</td>
<td>0.00004</td>
<td>0.00004</td>
<td>0.00111</td>
</tr>
</tbody>
</table>

Note.—The environment is assumed to vary cyclically, with period $T = T_1 + T_2$; in each cycle, the first $T_1$ generations are in environment 1 and the last $T_2$ generations are in environment 2. In all cases, $s_1 = 1$, $s_2 = 0.001$, $T_3 = 25,000$, $u_{12} = 10^{-5}$, and the initial frequencies are $x_1 = 0$, $x_2 = 10^{-6}$, and $x_3 = 0.999999$.

In the first three cases, $T_1 = 200$. This $T_1$ was chosen because $A_1$ can increase from a very low frequency to a high frequency in 200 generations if $s_1 = 0.1$ (table 1) and because $s_1T_1 - s_2T_2 = -5$, so that $m_1$ is on the average smaller than $m_2$. In all of these three cases, $A_1$ and $A_2$ can be retained. In case 1, the mutation rate ($u_{13}$) from $A_1$ to $A_2$ is equal to that ($u_{23}$) from $A_2$ to $A_3$. Therefore, the frequency ($x_2$) of $A_2$ is expected to be lower than that ($x_3$) of $A_3$, for mutation occurs irreversibly from $A_2$ to $A_3$. Indeed, $x_2$ oscillates between 0.00012 and 0.39, whereas $x_3$ is between 0.00013 and 0.61. In case 2, $u_{13} = 10^{-6}$ is one order smaller than $u_{12} = 10^{-5}$, and $x_2$ can reach a value as high as 0.71; the maximum value of $x_3$ is only 0.29. In case 3, $u_{12} = u_{23} = 10^{-6}$ and $u_{13} = 10^{-7}$, and $x_2$ can become as large as 0.89. This value is higher than that in case 2 mainly because the mutation rate ($u_{23}$) from $A_2$ to $A_3$ in case 3 is one order lower than that in case 2 ($10^{-6}$ vs. $10^{-5}$), so that the decrease in $x_2$ resulting from irreversible mutation is slower in case 3 than in case 2.

In case 4, $T_1$ is only 100, so that $A_1$ cannot reach a substantially high frequency from a very low frequency (table 1). It is seen that $x_1$ and $x_2$ decrease with the number of cycles. Therefore, $A_1$ and $A_2$ will eventually be eliminated from the population. In case 5, $T_1$ is 150. Since $x_1$ and $x_2$ increase with the number of cycles, $A_1$ and $A_2$ can be retained in the population.
The last example suggests that, if \( T_1 = 200 \), \( A_1 \) and \( A_2 \) can be retained even if \( T_2 \) is considerably larger than 25,000. Indeed, if \( T_2 = 250,000 \) and the other parameter values are the same as those in case 3, \( x_2 \) increases from 0.000001 at \( t = 0 \) to 0.00014 at the end of the second cycle. That is, \( A_1 \) and \( A_2 \) can be retained in the population. Note that in this case \( s_1T_1 - s_2T_2 = 230 \), that is, \( A_1 \) is on the average much less fit than \( A_2 \) and \( A_3 \).

The constant-fitness model predicts that \( A_1 \) and \( A_2 \) cannot be retained in any of the above cases because \( s_1T_1 - s_2T_2 < 0 \) in every case. This prediction holds only for case 4. I shall explain later why the difference occurs between the two models.

Discussion

The biological rationale for the mutation rates used in the above examples is as follows. I assumed that the rate \( u_{12} \) of mutation from the functional allele \( (A_i) \) to the cryptic allele \( (A_j) \) is at least as high as that \( u_{13} \) from the functional allele to the nonfunctional allele \( (A_k) \). This assumption seems reasonable because the former type of mutation would include defects in regulatory sequences, nonsense mutations, and frameshifts owing to minor insertions or deletions, while the latter type would be more drastically destructive changes such as large deletions. Limited data suggest that, in bacteria, \( u_{12} \) is of the order of \( 10^{-5} \) or lower (table 23-1 in Strickberger [1976]). Thus, the assumption of \( u_{12} = 10^{-5} \) or \( 10^{-6} \) seems reasonable. I also assumed that back mutation from \( A_j \) to \( A_i \) occurs at the rate of \( u_{21} = 10^{-7} \). Available data suggest that this rate ranges from \( 10^{-5} \) to \( 10^{-8} \) (Strickberger [1976], table 23-1; Hall et al. 1983). Obviously a higher rate of back mutation is more favorable for the retention of \( A_1 \) and \( A_2 \). In addition, I assumed that the rate \( u_{13} \) of mutation from \( A_i \) to \( A_k \) is the same as that \( u_{33} \) from \( A_j \) to \( A_k \). Reasoning that most mutations of the former type would be similar to mutations of the latter type, that is, regulatory defects, nonsense mutations, minor insertions, and minor deletions. A higher \( u_{33} \) value is obviously less favorable for the retention of \( A_1 \) and \( A_2 \). However, unless \( u_{33} \) is actually considerably higher than assumed, the conclusion drawn from the above examples will still hold.

The rationale for the selection coefficients used is as follows. I assumed that \( s_1 = 0.1 \) and \( s_2 = 0.001 \). In practice, the \( s_1 \) value may be larger than 0.1 because, as a result of their inability to utilize a resource available to \( A_i \) individuals, individuals without the functional allele \( (A_i) \) may grow much more slowly or may not even be able to survive in environment 1. If \( s_1 \) is indeed larger, the number of generations required for a substantial increase in \( x_i \) would be smaller than those assumed in table 2. Of course, the contrary would be true if \( s_1 \) were smaller than 0.1. A larger \( s_2 \) value is less favorable for the retention of \( A_1 \) and \( A_2 \) because in environment 2 the decrease in \( x_i \) will become faster. It can, however, be shown that the retention of \( A_1 \) and \( A_2 \) is not strongly dependent on the magnitude of \( s_2 \). For example, if \( s_2 = 0.01 \) instead of 0.001 and the other parameter values are the same as those in case 2, \( A_1 \) and \( A_2 \) can still be retained.

In all the above examples, I have assumed that the environment changes cyclically. This assumption was made to simplify numerical computations. Actually, many general conclusions can be drawn from these examples. First, if \( x_1 \) or \( x_2 \) becomes high, then the retention of \( A_1 \) and \( A_2 \) is assured for a long period of time, say at least one million generations, regardless of how the environment changes. For instance, in case 2, \( x_2 \) was high (0.71) at the end of the fifth cycle,
so that $A_1$ and $A_2$ will still be retained even if the environment continues to be in type 2 for one million generations before switching to type 1. Second, if $x_1$ becomes very high before the environment switches to type 2, $x_2$ may increase to a high value (see cases 2 and 3). Third, $A_1$ and $A_2$ may persist in low frequencies for a long period of time without being eliminated from the population (case 5). Fourth, a long run of environment 1 is more effective for the retention of $A_1$ and $A_2$ than many short runs with a total duration equal to that of the long run. This follows from the observation that $A_1$ and $A_2$ cannot be retained if $T_1 = 100$ and $T_2 = 25,000$ (case 4) but can be retained if $T_1 = 200$ and $T_2 = 250,000$; the ratios of $T_1$ to $T_2$ for the two cases are 0.004 and 0.0008, respectively. In the extreme case where $T_1 = 1$, $A_1$ and $A_2$ cannot be retained if $T_2 \geq s_i/s_j$, because the selection for $A_1$ is too ineffective to compensate the effect of irreversible mutation. In this extreme case, the conditions for the retention of $A_1$ and $A_2$ are similar to those required under the constant-fitness model.

In essence, the above examples show that $A_2$ can be retained without having any selective advantage over $A_1$, if selection for $A_1$ occasionally occurs for a substantially long time. The reason for this is quite simple. Although in environment 2 the frequency of $A_1$ will decrease fairly rapidly, $A_2$ can persist for a long time because $x_2$ will first increase at a fast rate and then decrease at the rate of mutation. If the environment switches to type 1 for a substantially long period before $x_1$ becomes extremely small, a substantial increase in $x_1$—and, consequently, a substantial reduction in $x_2$—will occur. When the environment switches to type 2, $x_2$ will start to increase. In particular, should $x_1$ become almost 1, $x_2$ may increase to a high value, as was seen in cases 2 and 3 above. (A similar argument has been made by Hall et al. [1983], though no mathematical treatment was given.) By contrast, under the constant-fitness model, if $A_1$ is less fit than $A_2$ and $A_3$, and $A_2$ and $A_3$ are equally fit, then the frequency of $A_1$ will always decrease, eventually to zero, because of selection and irreversible mutation, and the frequency of $A_2$ will also eventually decrease to zero because of irreversible mutation. In conclusion, the constant-fitness model requires more stringent conditions than actually needed for the retention of cryptic genes.

So far I have not considered the effect of population subdivision. To see this effect, let us consider a simple example. Suppose that there are two populations which exchange 10% of their genes in each generation. We assume that the environment varies cyclically with $T = 25,200$ generations and that in each cycle population 1 is in environment 2, except for the first 100 generations, and population 2 is in environment 2, except for the second 100 generations. The other parameter values are the same as those in case 4 in table 2. Numerical computations show that at the end of the fifth cycle $x_1$ increases to 0.0002 in both populations. Thus, in contrast to case 4, the cryptic gene can be retained. This example indicates that the conditions required for the retention of cryptic genes are weaker in a subdivided population than in a population without subdivision.

We may conclude from these results that loss of a cryptic gene from a microbial population is an extremely rare event. This conclusion, however, does not hold if a population goes through a severe bottleneck, a situation that apparently often occurs at the time of speciation. This might have been the case in *Shigella dysenteriae*, which lost the gene for a lactose permease (*lac Y* gene) after its separation from *Escherichia coli* (Hall et al. 1983). The loss appears to have occurred rather quickly, because *Shigella* and *Escherichia* are close relatives and,
in fact, are considered by some authors to be one species (Whittam et al. 1983). The fact that the gene diversity is considerably lower in _S. dysenteriae_ than in _E. coli_ (Whittam et al. 1983) suggests that the former has gone through at least one bottleneck. Let us assume that a bottleneck has indeed occurred and see how quickly the nonfunctional allele can become fixed in the population, assuming fixation will eventually occur. Let \( N_e \) be the effective population size and \( p \) the frequency of \( A_1 \) right after the bottleneck. Neglecting the effect of mutation and the existence of \( A_2 \), we can treat the problem as the conditional fixation of a neutral allele in a population. Using formula (14) of Kimura and Ohta (1969), we can show that the mean conditional fixation time \( (\bar{t}_f) \) is between 2.8\( N_e \) and 4\( N_e \), if \( p \) is between .5 and 1/(2\( N_e \)). If \( N_e \) is 10,000, \( \bar{t}_f \) is less than 40,000 generations, which is a short period, because the generation time in bacteria is very short. Therefore, the _lac Y_ gene could have become lost quickly in _S. dysenteriae_ if a fairly severe bottleneck had indeed occurred after its separation from _E. coli_.

Acknowledgments

This study was stimulated by Hall, Yokoyama, and Calhoun (1983). Thanks to B. G. Hall and M. Kricker for many suggestions. This study was supported by research grants from NSF and NIH.

LITERATURE CITED


Robert K. Selander, reviewing editor

Received July 22, 1983; revision received September 21, 1983.