Lessons from the large literature on plant reproductive biology have already shown surprising parallels in marine systems, and will be a productive guide to future research.

Acknowledgments
I thank Ed Metz, Vic Vasquez and my parents for discussion about genetics, and Andrew Martin, Cary Morris, Hampton Carson, Richard Schusterman and Mike Hulley for comments on the manuscript. Supported by plants from NSF and the Whiffen Founda-

References


What is a Quiescence?
By Martin A. Novak

A quiescence is a well-defined distribution of mutants that is generated by a mutation-selection process. Selection does not act on a single mutant but on the quiescence as a whole. Experimental systems have been designed to study quiescence evolution under laboratory conditions. More recently, several populations have been used as clonal cultures to elucidate their extensive genetic heterogeneity. The most prominent examples are probably the human immunodeficiency virus HIV-1 and HIV-2. The quiescence nature of HIV has formed the basis of a model that provides a mechanism for the quiescence of acquired immunodeficiency syndrome (AIDS) in humans. This article focuses on the nature of the quiescence concept and its implications for evolutionary biology and virology.

The term 'quiescence' was introduced by Eigen and Schuster in 1971. In the context of their work on the origin of life, they described the cluster of closely related molecular species produced by errors in the self replication of macromolecules (nucleic acids). This followed Eigen's first theoretical model of molecular evolution based on chemical kinetics. In the original fiction of Eigen and Schuster, a quiescence is defined as the equilibrium mutant distribution that is generated by a specific mutation-selection process. The quiescence represents a specific nucleic acid sequence. This sequence number may determine the replication rate of a given variant. The replication rates of two variants, i and j, may be denoted by a, b, c, d, e, f. These quantities represent the selective value of the individual mutants. In the absence of mutation, the variant with the highest replication rate will grow fastest and most frequent.

The result of selection in this world without error is a distribution of the most replicating variant. But replication is not error-free. Hence, it is necessary to define the probabilities qt of (true) replication of template i, results in the production of the sequence j. The quantities qt for i = 1, 2, ... and j = 1, 2, ... form the so-called mutation matrices. A system of ordinary differential equations describes the time evolution of the population of these nucleic acid sequences. The growth rate of a specific variant, e.g. j, can be written as:
New particles of variant i, can be formed by error-free replication of 1, the probability of the rate a, and the probability of 2, the overall rate is therefore given by a Q k. Erroneous replication of any other mutants I, " I, can also lead to new 1, particles. This is represented by the growth terms a Q k, l + l - a Q k, Q k in Eqn 1. In the same way, we can write the rate of production of any of the other variants to obtain the whole system of differential equations

\[ \frac{dx_{i}}{dt} = a_{Q_{k}} x_{i} - a_{Q_{k}} x_{i} \quad i = 1, \ldots , n \]

In this context, the population will no longer consist only of the fastest growing sequence, but of a whole ensemble of mutants with different replication rates. This ensemble of mutants is the quasisspecies (see Box 1).

The frequency of a given variant within the quasisspecies does not depend on its replicative value alone but also on the likelihood with which it is produced by erroneous replication of other templates and their frequencies in the population. This is important to the understanding of the structural organization of a quasisspecies. The consequence of this effect is that the individual sequence i, with its replicative value \( a \) now serves as the unit (or target) of selection. The quasisspecies itself is the target of selection in a mutation-selection process. This fact has important implications. Evolution is normally thought of as the interaction between mutation and selection. Selection is a factor that favors advantageous mutants that have been generated by pure chance. Indeed, it is normally considered a mistake to think of mutations as being guided otherwise than by chance. A quasisspecies, however, can guide mutations. This does not mean that there is any correlation between the intrinsically stochastic act of mutation and the selective advantage of the mutant. But selection operates on the structure of the whole quasisspecies, which is adapted to its fitness landscape. This term is originally from Sewall Wright. Therefore, evolution can be guided towards the peaks of this fitness landscape. This happens because more-successful mutants (which may be closer to the peaks of the landscape) will provide more offspring than less-successful mutants (which may be further away from the peaks). Evolutionary computation can be viewed as a hill-climbing process of the quasisspecies that occurs along certain trajectories in sequence space (see Box 2).

Error thresholds

Another important concept in quasisspecies theory is the error threshold of replication \( q^{*} \). If replication were error free, no mutants would arise and evolution would stop. Evolution would, however, be also impossible if the error rate of replication were too high only some mutations may lead to an improvement in adaptation, but most will lead to deterioration. The quasisspecies concept allows us to quantify the resulting minimal replication accuracy that maintains adaptation.

Let us assume that a population consists of (1) a fast replicating variant \( i \), the wild-type sequence - with replication rate \( a_{i} \) and (2) its mutant distribution \( \epsilon_{j} \), with a lower average replication rate \( a_{i} (\epsilon_{i} > \epsilon_{j}) \). Let \( x_{i} \) denote the mean distribution (per \( a_{i} \) allele \( \epsilon_{i} \), with a lower average replication rate \( a_{i} (\epsilon_{i} > \epsilon_{j}) \). Let \( q^{*} \) denote the probability that a single base is accurately replicated. Thus, the probability that the whole sequence of length \( m \) is replicated without errors is given by \( Q^{*} \). Neglecting the small probability that erroneous replication of a mutant gives rise to a wild-type sequence leads to the equations

\[ \frac{dx_{i}}{dt} = a_{Q_{k}} x_{i} - a_{Q_{k}} x_{i} \]

Here the ratio of wild type to mutants converges to

\[ x_{i} \rightarrow \frac{Q_{i}}{1 + Q_{i}} \]

The wild type can only be maintained in the population if \( Q_{i} > a_{i} \). This means that the single-digit replication accuracy, \( a_{i} \), must be larger than a certain critical value. This error threshold relation is obtained as

\[ Q_{i} > a_{i} \Rightarrow \frac{Q_{i}}{a_{i}} > 1 \]

For replication accuracies lower than \( a_{i} \), the wild-type sequence will be lost from the population although it has the higher replication rate. This leads to an important relationship between the replication accuracy and the sequence length

\[ m < \frac{1}{1 + Q_{i}} \]

To evaluate the approximation that the logarithm of \( a_{i} \), is about 1. This represents an approximation.
Nucleic acids such as DNA or RNA consist of specific sequences or ‘strings’ of four individual nucleotides (G, A, C, and T). For a given chain length these are 4n different variants. This means that even for moderate chain lengths a ‘random (optimized)’ large number of different variants can be generated with a probability of length 300, which is just large enough to encode line of the smallest proteins. There are more than 1013 different variants. The genome length of HIV is about 10,000 bases. A particular HIV sequence is one choice out of 10^10^ different nucleotide sequences of the same length. For comparison, there are likely 10^10^ stars in our galaxy, or about 10^20^ proteins in our universe. Avogadro’s number is only 6.02 x 10^23.

Let us imagine all possible nucleotide se-
quence be arranged in a ‘sequence space’ such that two sequences are neighbors if one can be converted into another by a single point mutation. Thus the sequence space is formed by a set of sequences (of uniform length) together with a definition of a distance between sequences. An app-
propriate definition is given by the hamming distance. Let us specify two sequences ω and ω the Hamming distance counts the number of different positions in the two sequences. It represents the total number of point mutations that are necessary to change one sequence into the other. A cor-
nect ordering of all the mutants according to their mutual Hamming distances leads to an multidimensional space. The important features of this arrangement are: (1) the high dimensionality, (2) the large number of neighbors, (3) the possibility to generate a dist-\ninct mutant frequency (for two sequences approximately the square root of the total number of nearest mutants found), and (3) that many sequences are confined to a close neighborhood of each other. The genome of a sequence that contains 10^8 points in only 132 length unit, i.e., point mutations. The only feasible approach the point mu-
utations can leap from one region in the sequence space to a completely different region. In order to find this maximum, a guiding gradient to avoid going in ‘wrong directions’ is provided by natural selection.

for the upper genome length n that can be maintained by a given single
digit replication accuracy without losing fidelity.

Experimental studies

The quasispecies concept be-
comes important whenever muta-
tation rates are high. This is often the case in viral and bacterial popu-
lations. The first in vitro selection evolution experiment was per-
formed by Sewing (1980) and contin-
ed in parallel by the Elledge group. The traditional viral transfer experiment. adaptation of the E. coli phage Qβ to an artificial selection pressure was demonstrated. The artificial selec-
tion pressure was exerted by ethidium bromide, which inhibits

slow down replication. After a c
ertain number of passages, a multi-
ple sequence evolved that was growing twice as fast as the original wild type in the presence of ethidium bromide. The difference between wild type and mutant was lost three point mutations. Some serologic machines have been con-
structed to perform these trials by transferring automatically. Such evol-
utionary reactions allow one to slowly adapt the viral selection pressure (H. Oden, Thesis, Max Plank Institut, Göttingen, 1990).

Bevins et al. (1989) studied the quasispecies distribution of short-
chained RNA templates in a cell-free medium with Qβ replicase. Se-
quence analysis indicated a broad distribution of mutants around a master sequence. In the system, mutation rates and selective values of individual sequences can be measured.

Another interesting observation was the de novo synthesis of RNA templates by the Qβ replicase. Initially the system contains only Qβ replicase and nucleotides triphos-
phates, but no RNA molecules. Some templates are then formed by channe. These short templates conti-
act for replication and replication by the Qβ enzyme. This leads to the growth of the best adapted sequence.

Mutation rates can be deter-
mined experimentally. Mutation frequencies (i.e., the probability that a replication enzyme makes a mistake in the catalytic position) have to be distinguished from the replication frequencies (i.e., the population). For example, a host spot in a region of the genome with very high mutation frequency. A spot can be generated in two ways: (1) the mutation rate at this position is very high or (2) selection favors or tolerates variation in this region. In the second case, mutation and mutation frequencies are completely different. Thus, to count mutants in a quasispecies distribution is not a way to deter-
mine mutation rates. An elegant method to obtain mutation fre-
quencies is to measure the evolu-
tion rate of conditionally lethal mutants produced by site-directed mutagenesis.

Viral quasispecies

Virus populations in gen-

eral consist of a wide diversity of mutator distribution rather than homogenous population of a single mutation. This diversity is expressed in different ways. It is most easily seen in electron micrographs, but also in plaque size, plaque morphology, and virus morphology. The diversity has been detected in, for example, influenza or avian influenza. Viruses possess a great many different mutants. The results show a correlation between error rate and sequence length.

In a classical experiment by Domingo et al. (1988) single particles of the phage Φ29 were cloned. None of the clones had the same sequence. The results indicate that the viral quasispecies is present at a level of less than 1%. The mutational diversity of a virus is an important example of a viral quasispecies. Patients infected with HIV harbor a highly diverse virus population with many different mutants. Mutations are gener-
ated by the virus-encoded reverse transcriptase, which has an error rate of about 10^-9 to 10^-10 per base. This implies that during each replication of the whole genome 1-10 errors are produced. This high error rate is very close to its error threshold.

Differences in reading frame and protein sequence of different residues of the viral quasispecies. This is a major factor driving disease progression. The essential idea is that the virus escapes immune pressure by continuously pro-
cution of new mutants resistant to current immunological attack. This results in the accumulation of antigenic diversity during the asymptomatic period. The exis-
tence of an antigenic diversity (residual is derived from the virus quasispecies and the immune system; strain-specific virus response are directed against specific HIV antigens, but each virus strain can impair all immune responses regardless of their specificity. Therefore, in-
creasing HIV diversity enables the virus to escape from control by the immune system. In this context, the observed genetic variability is responsible for the fact that the virus establishes a persistent infection without being cleared by the immune response.
and induces immunodeficiency dis-
case after a long and variable in-
fection period.
HIV can also evolve drug re-
istance after about six months of treat-
ment with zidovudine, the drug most widely used to treat HIV infections, is me-
sured by three or four amino
acid substitutions in the virus-
encoded reverse transcriptase[16].

Conclusions
Quasispecies theory, which is based on chemical kinetics, provides a mathematical framework in order to understand molecular evolution. Selection and mutation form a distribution of mutants that is called a quasispecies. The target of selection is not an individual mutant sequence but the whole quasispecies. Therefore, fitness is a property of the quasispecies and not of individual mutants.

The fitness of a quasispecies is mathematically defined as the largest eigenvalue of the mutation-selection matrix. Selection stabil-
izes a quasispecies distribution in sequence space, and evolution can be viewed as a desaturation of an existing quasispecies upon arrival of a new advantageous mutant that establishes a new quasispecies[16]. A quasispecies may be centered around a master sequence with high efficiency of reproduction. The consensus sequence of the quasispecies need not be identical with the master sequence. For larger mutation rates, the frequency of the master sequence in the population can be very low. Evolution seems to work fastest close to the error threshold. The quasispecies has changed the classical view of evolution from the picture of a single wild type moving through sequence space by random walk into the picture of a quasispecies with its mutant distribution migrating through sequence space in an internally self-organized manner and yielding itself to the peaks of the fitness landscape.

References
1 Eigen, M. and Schuster, P. (1977) Naturwissenschaften 64, 541-565
15 Domingo, E. et al. (1986) Gene 40, 1-37
27 Mapoukhos, A. et al. (1989) Cell 58, 301-310
28 Simmonds, P. et al. (1990) J. Virol. 64, 540-541

TREE for colleagues abroad

In some countries, the currency needed to pay for a personal subscription (US dol-
ars or pounds sterling) is not available. If you wish to help a colleague abroad who
is not able to benefit from TREE for this reason, we will accept your payment for
another person's subscription. Simply com-
plete the subscription order card bound into
any issue, giving the recipient's name and
address labelled 'send to'. After signature,
give your own name and address and mark
this 'bill to'. Renewal notices will be sent
to your address and the recipient will receive the
monthly copy of the journal. Please inform
the recipient of your action.

Letters to the Editor

TREE welcomes correspondence. Letters to the Editor may address issues raised in the pages of TREE, or other matters of
general interest to ecologists and evolu-
tionary biologists. Letters should be no more than 300 words long, and should be sent to The Editor, Trends in Ecology and Evolution, Elsevier Trends Journals, 68 Hills Road, Cambridge CB2 1LA, UK. The decision to publish rests with the Editor, and the author(s) of any TREE
article criticized in a Letter will normally be invited to reply.