STADLER SYMP. Vol. 10 (1978) University of Missouri, Columbia 105

NON-DARWINIAN EVOLUTION AND THE BEARD OF LIFE

(genetic drift, repetitive DNA, philosophy of science)

GEORGE P. SMITH

Division of Biological Sciences
Tucker Hall
University of Missouri
Columbia, Missouri 65211

SUMMARY

This paper explores what would happen to a chromosomal segment that evolves randomly, without the surveillance of natural selection. In these circumstances a single segment present in some individual organism at any given starting time will eventually spread through the population until it becomes fixed in the entire gene-pool - or more properly the "segment pool". At the same time, the continual deletion, duplication and insertion of nucleotides in a randomly-evolving segment of DNA will lead to fixation of individual nucleotides in the total pool of nucleotides in all the homologous segments in the species. When this genetic drift at the nucleotide level is examined in detail, it is seen to provide a simple, natural solution to a contemporary problem in molecular evolution: the prevalence of apparently functionless segments consisting of simple nucleotide sequences repeated tandemly thousands to millions of times. The solution exemplifies a class of scientific theories that are difficult or impossible to corroborate experimentally, even though they may be intellectually satisfying and of considerable practical importance in the progress of science.

INTRODUCTION

Somewhere I have read of a museum that has on display a magnificent beard of stupendous length - let us say six feet. Perhaps it was like the beard shown *in vivo* in Fig. 1. But its length is by no means its most remarkable quality. Scientific analysis of its architecture revealed that no individual hair in it measured more than twelve inches! It is apparent that some of us would get on much better without a comb and brush.

In certain respects such a beard may be a more fitting image of evolutionary descent than the more usual image of an evolutionary tree. Just as the beard's overall continuity is

Figure 1. A magnificent beard of stupendous length.

not reflected in the continuity of its individual hairs, so also the continuous lines of descent leading from remote ancestors to contemporary living things may not be reflected in continuous lines of descent of their individual DNA nucleotides. Nucleotidyl lineages are continually becoming extinct, even in flourishing organisms. This extinction is balanced partly by over-expansion of other nucleotidyl lineages, and partly by the *de novo* creation of new nucleotidyl lineages.

This rather unconventional picture of the evolutionary process comes from a consideration of how DNA would evolve if it were not subject to natural selection - if heritable changes accumulated at random. This type of evolution is often called non-darwinian. I hope you will not think that I am foolishly going to try to argue against the role of natural selection in evolutionary change. What I do maintain, however, is that it is plausible to suppose that at some epochs, some segments of the genomes of most species are largely relieved from the surveillance of natural selection. The accumulation of changes in the nucleotide sequences of such genomic segments becomes a *stochastic* process - that is, a process governed by chance. I will argue that certain striking characteristics of the genomes of higher organisms find a much simpler and more elegant explanation in non-darwinian theory than in terms of natural selection.

GENETIC DRIFT OF SEGMENTS IN A SEGMENT-POOL

At the heart of non-darwinian theory is the concept of genetic drift, as first propounded by Sewall Wright (WRIGHT 1931). To illustrate drift, I will focus on one particular segment of the nucleotide sequence in one particular chromosome of a hypothetical species. The set of all allelic segments in the species constitutes what we can call a "segment pool". A segment

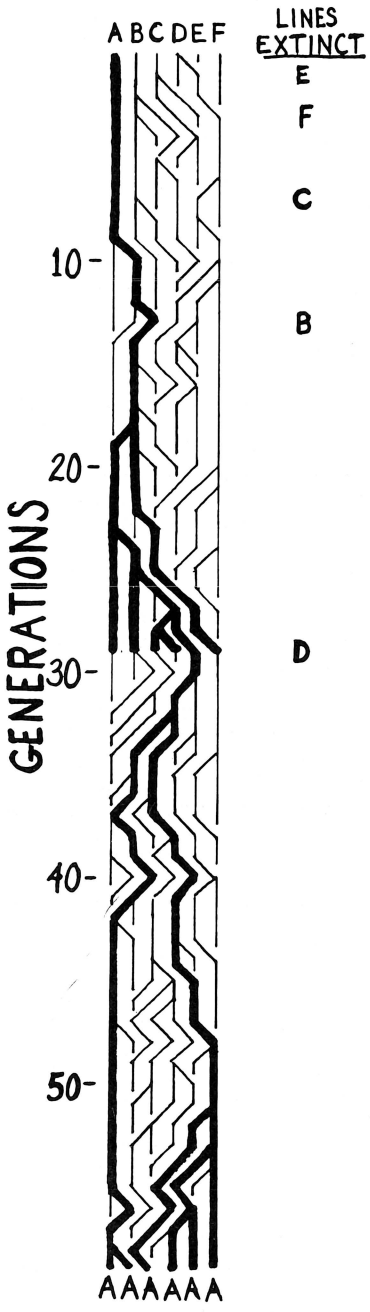

Figure 2. Fixation of DNA segments by random sampling of gametes. See the text.

pool is analogous to the more familiar gene pool, but is intended to be noncommittal about the functional role (if any) of the segment under consideration; if in fact the segment happens to be a functional gene, the segment pool can also be called a gene pool. But here we are particularly interested in segments of the genome that are not functional - at least for the time being. Under these assumptions, the probability of a given segment in the contemporary segment pool contributing to the segment pool of the next generation does not depend on the state of that segment - that is, the mutations it carries. All segments in the pool are intrinsically neutral as far as natural selection is concerned. Let us make the additional assumption that there are no extrinsic factors contributing to nonuniformity of the transmission of segments from one generation to the next. Thus I assume that the organismal population breeds at random and that the segment is not linked closely enough to any non-neutral segment of the genome to affect its probability of survival. Even under these assumptions, which maximize the uniformity of transmission of the segments from one generation to the next, there will be a certain random variance from perfect uniformity. Each generation can be seen as a random sample of the gametes produced by the previous generation, and sampling error will lead to a certain irreducible variance in transmission of segments from one generation to the next. I have illustrated this variance schematically in Fig. 2, which follows the descent of the segments in a hypothetical segment pool consisting of six segments in each generation - as would be the case, for example, in a hypothetical population of 3 diploid organisms. To simulate random sampling of gametes, one segment in each generation was chosen at random to leave no descendant segments in the next generation, and another to be represented by two descendants in the next generation. The former are indicated by terminations of lines of descent, the latter by branchings into two lines of descent. To see the effect of random sampling, let us imagine that we label all the segments in the segment pool at some starting time (as I have labeled A through F the segments present at generation 0 in Fig. 2) and keep track of those segments in later generations that descend from each of the zero-time segments. Variance in transmission will lead to fluctuation in the number of descendants representing the various zero-time segments, some original segments being underrepresented, others being overrepresented. At each generation, there is a finite chance that all remaining descendants of a given one of the original segments will be completely eliminated. In Fig. 2, for example, the lineage descending from original segment E happens to become extinct by generation 1, that from F at generation 3, and so forth, as shown in the list at the right-hand edge of Fig. 2. Hence, as time goes on, more and more of the original segments will become extinct. These extinctions can only be compensated by overrepresentation of the remaining original segments. Eventually all but one of the original segments must become extinct. In the parlance of population genetics, that one surviving segment has become *fixed* in the segment pool. In Fig. 2 it happens to be original segment A which becomes fixed in this sense by generation 29.

For any epoch we choose as the contemporary time, there

will be an earlier generation such that all the contemporary segments descend from a single one of the segments in that earlier generation. In Fig. 2, for example, the bold lines show how all the segments present at gerneration 29 descend from a single common ancestral segment at generation 18; prior to generation 18, the segments in generation 29 are represented by a single line of descent. Similarly, the segments present at generation 59 are represented by a single line of descent prior to generation 32. Thus while the total evolutionary history looks like a continuous beard of divergences and extinctions, the evolutionary lineages leading to the segments in any one generation resemble an inverted tree.

Only mutations occurring in the branches descending from the most recent common ancestor are inherited by some of the contemporary segments and not by others, and can thus contribute to differences among the contemporary segments. Mutations occuring prior to the epoch of the most recent common ancestor-segment cannot contribute to the heterogeneity of the contemporary segments; most of these mutations are lost as the lineages in which they occur become extinct; the remaining mutations - those occurring in the single line of descent leading to the common ancestor - are inherited by all of the contemporary segments. These latter mutations, like the segments that inherit them, have become fixed in the segment-pool.

If (as we have postulated) natural selection neither favors nor disfavors mutations in the segment under consideration, the nucleotide sequences of the segments in the pool will diverge more and more from those of their ancestor-segments. Yet while these mutations accumulate, the segments present at any one generation will tend to remain relatively similar to one another, since heterogeneity can only accumulate during the limited time-lapse between that generation and the epoch of the most recent common ancestor. Thus we see that the change in characters within a species (the characters in this case being the nucleotide sequences of the segments) has direction even in the absence of natural selection. The absence of selection means, not that evolution has no direction, but rather that the choice of direction is random.

Natural selection superimposes a bias on the stochastic transmission of segments from one generation to the next, disfavoring the spread of deleterious mutations and favoring the spread of advantageous mutations. It can be seen as a driving force that increases the rate at which mutations become extinct and fixed. Hence natural selection tends to reduce the heterogeneity of the segments in a segment pool by reducing the transit-time between the occurrence of mutations and their ultimate extinction or fixation.

Several population geneticists, led by Kimura (KIMURA 1968), have pointed out that the very large heterogeneity that is actually observed at single loci in many species of higher organism finds a ready and simple explanation in the hypothesis that many mutations are nearly neutral as far as natural selection is concerned and therefore are fixed and extinguished at the leisurely

pace characteristic of genetic drift. I must warn you, however, that this view is highly controversial - as indeed it was when Sewall Wright (WRIGHT 1931) first propounded the theory of genetic drift. The "selectionists" (as the opponents of the neutral mutation theory are often called) have managed to devise schemes of selection which might account for the observed heterogeneity (MILKMAN 1976). Unaccountably they seem to triumph in thus rescuing conventional evolutionary theory from simplification by means of further complications. In fact, the selectionists have gone on to uncover some apparent discrepancies between the observations and non-darwinian theory. I say "apparent" discrepancies because I suppose that if the selectionists were to allow as much scope to the non-darwinian imagination in clearing up these discrepancies as they have to their own imaginations in devising their selective schemes, the discrepancies will largely disappear (KIMURA 1976). No doubt you will divine on which side of this controversy my sympathies lie.

GENETIC DRIFT OF INDIVIDUAL NUCLEOTIDES IN THE TOTAL NUCLEOTIDE POOL

So far I have been discussing the descent of segments as if they were indivisible units, but of course in reality they are composed of nucleotides. The evolutionary lineages I have been discussing ought properly to be regarded as *bundles* of lineages representing the descents of individual base-pairs. For the most part these lineages run strictly parallel to each other because of the extreme fidelity of DNA replication. There will be occasional irregularities, however. Homologous crossover between segments would be represented by reciprocal exchange of parts of the bundles corresponding to the two recombining segments; this can lead to non-parallelism between the lineages for different positions within the segment under consideration. This does not fundamentally alter the picture, however, since all the arguments can simply be interpreted as applying individually to subsegments that are sufficiently short - single base-pairs if necessary - that no intra-subsegment recombinants survive among the contemporary segments.

A more fundamental type of irregularity is due to deletion and duplication of nucleotides, which are known to occur spontaneously in DNA. These deletions and duplications cause an unevenness in the transmission of the nucleotides from one generation to the next in a lineage of segments that is exactly analogous to the unevenness in the transmission of segments from one generation to the next in a population of organisms. The same conclusions therefore apply: nucleotides become fixed in the "linear pool" of nucleotides that constitutes a DNA segment. At the same time, of course, the segments themselves are continually becoming fixed in the segment pool. Ultimately, then, the descendants of a single nucleotide-pair present at an arbitrary starting time will spread until they come to occupy all the positions in all the allelic segments in the segment pool. In other words, the total nucleotide pool - that is, the pool of all base-pairs in all the allelic segments in all the individuals in a species - will be connected to the remote past by the

precarious thread of a single lineage of nucleotide-pairs.

Truly it is a precarious thread. If it could be traced back far enough, it ought to come to an end. This end corresponds to the spontaneous insertion of one or more nucleotides, another process known to occur in DNA sequences. Such inserted nucleotides, unlike the extra nucleotides added to DNA as a result of some form of duplication, cannot be said to have descended from any prior nucleotides; from an evolutionary point of view insertions are *de novo* creations. These nucleotidyl *parvenus* have quite as good a chance of being ultimately fixed as their neighbors of more venerable lineage in the extreme democracy of non-darwinian theory. Eventually, of course, some of them *will* be fixed, and the total nucleotide pool will hang suspended from one or more tiny acts of spontaneous creation. We have come full circle to my starting image of a beard of life, whose nethermost hairs are unconnected to the chin of the First Creator.

A NON-DARWINIAN THEORY OF REPETITIVE DNA

Personally I find this image appealing, but I doubt that in reality any segment of DNA survives long enough without either accidental extinction or acquisition of some physiological role for the full non-darwinian development outlined above to occur. Nevertheless certain portions of the theory, such as the slow fixation of neutral mutations, may illuminate observable phenomena, as we have already seen. I would like to focus on another such part of the theory.

We saw above that the deletion and duplication of nucleotides lead to the fixation of nucleotides in what I called the "linear pool" of base-pairs that constitutes a DNA segment. One way such deletions and duplications are known to arise is by unequal crossover. As shown in Fig. 3, an unequal crossover produces one recombinant molecule with a deletion of a block of nucleotides and another recombinant molecule with a tandem duplication of the same block of nucleotides. Such unequal crossovers may be one explanation of how tandem duplications arise in nonrepetitive DNA. However, since the participating DNA molecules are not homologous at the point of crossover, this type of unequal crossover is presumably very much less likely than equal crossover between homologously aligned DNA molecules. But once an array of two or more tandem repeats arises in the DNA sequence (whether by nonhomologous unequal crossover or by some other mechanism), unequal crossover can occur between different repeats aligned in register, as shown in Fig. 4. Since DNA molecules aligned in this fashion are homologous at the point of crossover, this type of unequal crossover ought to be relatively efficient compared to other modes of deletion and duplication. This homologous but unequal crossover has the effect of deleting or tandemly duplicating a block of adjacent repeats in the tandem array, as shown in Fig. 4. Because of the efficiency of homologous but unequal crossover relative to other forms of duplication, once arrays of two or more tandem repeats arise in a segment of DNA, duplications ought to take the form predominantly

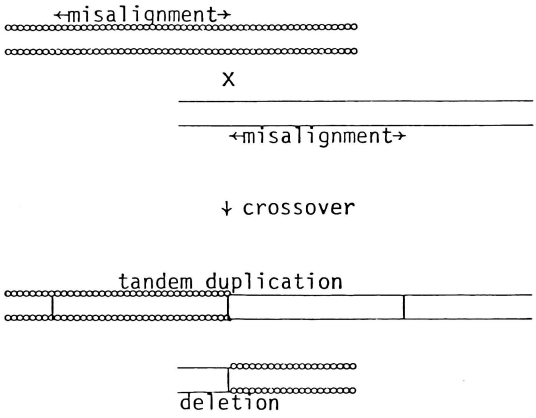

Figure 3. Unequal crossover results in deletion or tandem duplication of a block of adjacent nucleotides, depending on which (if either) of the two recombinant chromosomes survives random extinction in future generations.

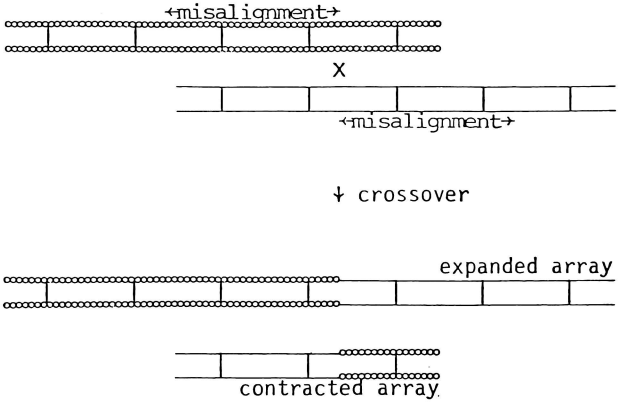

Figure 4. Homologous but unequal crossover between tandem arrays results in deletion or tandem duplication of a block of adjacent repeats. Because of extensive homology at the point of crossover, this type of crossover is presumed to be much more likely than unequal crossover between non-repetitive DNA sequences.

of expansion of some of these tandem arrays. These expansions are more or less compensated by the contraction of other tandem arrays and by the deletion of non-repetitious portions of the DNA sequence. Ultimately, therefore, a single array of tandem repeats must expand to occupy the entire segment. In other words, a segment of DNA which evolves entirely at random ought

inevitably to become tandemly repetitious. Computer simulations confirm this deduction. Initially random, non-repetitive DNA sequences inevitably turn into tandem repeats when subjected to multiple rounds of random mutation and random unequal crossover (SMITH 1976).

Tandem arrays such as these abound in the DNA of eukaryotes. The non-darwinian theory provides a simple explanation for their prevalence in terms of the random accumulation of random mutations and random unequal crossovers. Alternative theories of repetitious DNAs attribute them to special replication mechanisms or to natural selection resulting from some physiological role the repetitious sequences play (BRITTEN & KOHNE 1968; FRY & SALSER 1977). In thus invoking special mechanisms or physiological roles *ad hoc* to explain the occurrence of sequences which would have a strong tendency to evolve anyway, these theories must, I think, be accounted gratuitously complicated and therefore inferior.

THE SCIENTIFIC STATUS OF UNCORROBORABLE THEORIES

From another point of view, however, the foregoing non-darwinian theory of repetitive DNA is very unsatisfactory. An enormous variety of different repetitive patterns could arise in the manner described above, and few conceivable patterns could be taken as contradictory to the theory. This capacity to explain nearly everything is not, as might be thought at first, a strength, but a weakness, of the theory. The philosopher Karl Popper made this point quite well, I think, in a discussion of the all-explanatory theories of Marx, Freud, and Adler, which were all the rage in the Vienna of his youth. "The most characteristic element in this situation," Popper writes, "seemed to me the incessant stream of confirmations, of observations which 'verified' the theories in question....A Marxist could not open a newspaper without finding on every page confirming evidence for his interpretation of history; not only in the news, but also in its presentation...and especially of course in what the paper did *not* say. The Freudian analysts emphasized that their theories were constantly verified by their 'clinical observations'. As for Adler, I was much impressed by a personal experience. Once, in 1919, I reported to him a case which to me did not seem particularly Adlerian, but which he found no difficulty in analysing in terms of his theory of inferiority feelings. Slightly shocked, I asked him how he could be so sure. 'Because of my thousandfold experience,' he replied; whereupon I could not help saying 'And with this new case, I suppose, your experience has become thousand-and-one-fold'" (POPPER 1963).

Popper's philosophy of science seeks to make a distinction between genuine science and pseudoscience, which, though it is often embedded in a confusing welter of observations (one thinks of the astrologer's charts, or of the tracts of the Transcendental Meditators) bears a different relationship to its empirical base than does genuine science. "Every 'good' scientific theory," Popper claims, 'is a prohibition: it forbids certain things to happen. The more a theory forbids, the better it

is....Confirming evidence should not count *except when it is the result of a genuine test of the theory*; and this means that it can be presented as a serious but unsuccessful attempt to falsify the theory....Some genuinely testable theories, when found to be false, are still upheld by their admirers - for example, by introducing *ad hoc* some auxiliary assumption.... Such a procedure...rescues the theory from refutation only at the price of destroying, or at least lowering, its scientific status" (POPPER 1963).

Popper's view has come to be the conventional wisdom in science today, though Popper himself is given much less credit for it than he deserves. The idea that a theory should be testable has become so much a truism that it is not impossible for a lecturer to propose some wildly implausible theory and then avert boos and catcalls by intimidatingly pointing out at the end of his talk that his theory has the virtue of being testable.

Judged by Popper's criterion, I'm afraid the foregoing non-darwinian theory of repetitive DNAs is sadly lacking in scientific status. It is predictable that evidence contrary to it will accumulate steadily (FRY & SALSER 1977). It would be surprising, for example, if some repetitive DNAs which now seem purposeless do not turn out to have physiological roles and to be maintained to some degree by natural selection. Such an empirical observation would certainly refute the most extreme form of the non-darwinian theory. But it would not damage a more relaxed theory that (for example) would allow occasional repetitive arrays that arose by chance to acquire physiological roles adventitiously. Such *ad hoc* modifications, as Popper points out and as I'm sure we all agree, weaken a scientific theory by making it even less predictive and more indefinite than before. But in my opinion it is a grave error to be too hasty to abandon a theory because of *ad hoc* modifications. If the modifications are plausible, the theory may still be scientifically important - perhaps much more important than more testable but less plausible rival theories - even if it is less corroborable. The importance of such quasi-uncorroborable theories lies in providing possible solutions to problems. In the present instance, the widespread occurrence of repetitive DNAs posed the problem of their origin. At first it seemed that the solution would take one of two forms: either a fundamental physiological role would be found for them, so that their origin could be attributed to the driving force of natural selection, or some special replicative mechanism would be shown to operate in chromosomes to produce tandem arrays. The non-darwinian theory, even burdened with *ad hoc* modifications, provides a plausible third solution, according to which no physiological role or special replication mechanism need be invoked. The availability of this solution might protect us from a prolonged search for physiological roles or special replicative mechanisms that may not exist, or that if they are found, may still not be the true solution to the problem at hand.

I think it is clear that the scientific weight of a theory, both as an explanation valued for its own sake and as a pragmatic guide to future research, must take into account its *a priori*

plausibility as well as the degree (if any) to which it is supported by any empirical evidence. The pragmatic importance of *a priori* considerations might be shown by two contrasting examples. On the one hand, the theory of natural selection, which holds that heritable changes arise preadaptively without regard to the biological need they may happen to serve, ought (I'm sure you all will agree with me here) to deter us from wasting our careers searching for a direct causal connection between biological needs and heritable changes. By contrast, the theory of Special Creation, which holds that God created the species, ought not to deter us from searching for evolutionary links among the past and present species of living things. Testability is irrelevant to the relative scientific merits of these two theories. Natural selection is a notoriously untestable theory: what possible observation about living things would we as working biologists, thoroughly indoctrinated with the Darwinian view, be willing to accept as a refutation of the theory? Special creation is a much more testable theory: it is clearly refutable by the discovery of a single example of a "missing link" between species. Of course, it is maintained by "evolutionists" (if I may so call the adherents of the materialistic religion of Darwin and his disciples) that a few such missing links have been found, and they have devised ingenious explanations for why missing links are rare in the fossil record. Understandably, they become quite annoyed when their hard-won evidence for missing links is disputed, often with a fine disregard for details, by the Special Creationists. But the fact that Special Creationists have rendered their own hypothesis irrefutable by a fixed policy of refusing to consider contrary evidence seriously does not vitiate their argument that evolutionism is no less an irrefutable system of belief than Special Creation. If untestable theories are to be relegated to the realm of metaphysics, what are we to answer to the Special Creationists' contention that the teaching of evolutionism to the exclusion of Special Creation in the public schools amounts to the establishment of an atheistic religion?

Even in the field of the empirical corroboration of theories, to which it particularly refers, Popper's theory of theories seems defective (GROVER 1974). Despite the emphasis he and most scientists give to testability, in practice empirical corroboration of theories, especially in biology, seldom involves a true test in Popper's sense. For it is very seldom that a theory cannot be rescued from contrary evidence by plausible *ad hoc* assumptions that do not greatly diminish its scientific merit. Corroborating evidence has much more significance in the progress of science (at least biological science) than contrary evidence, and even very unpredictive, irrefutable theories can nonetheless be corroborated. These empirical corroborations have I think a common logical structure, which I will try to cast in a probabilistic form. Corroboration starts from empirical data bearing on the theory. On the one hand, one computes, at least intuitively and unconsciously, the likelihood that one would obtain those particular data given the theory in question; I will call this the probability of the evidence given the theory, and symbolize it <evidence|theory(0)>, where the argument 0 distinguishes the theory of interest from the alternative

theories. On the other hand, one computes what I may call the *a priori* likelihood of the same empirical data. This *a priori* likelihood of the evidence can be decomposed into a sum of the form

$$\sum_i \langle \text{theory}(i) \rangle \langle \text{evidence} \mid \text{theory}(i) \rangle ;$$

here <theory(i)> is the *a priori* likelihood of theory(i), which is based on our background knowledge (including the results of previous observations), and the summation is over all theories, including the theory of interest theory(0). The ratio of these two probabilities

$$\frac{\langle \text{evidence} \mid \text{theory}(0) \rangle}{\sum_i \langle \text{theory}(i) \rangle \langle \text{evidence} \mid \text{theory}(i) \rangle}$$

I will call the corroboration index. If the index is greater than 1, the evidence tends to corroborate the theory; if it is less than 1, it tends to discorroborate it.

Even a highly unpredictive theory, for which no particular empirical observation follows with a high probability, can nevertheless be strongly corroborated provided the empirical data have an even smaller *a priori* probability. Let me illustrate the use of the corroboration index. We all probably agree that observing Mendelian segregation of a character strongly corroborates Mendel's theory of inheritance. This cannot be because Mendel's theory *entails* Mendelian segregation of characters: there are lots of reasons why a character might not segregate in a Mendelian fashion, such as multigenic inheritance, incomplete penetrance, etc. Indeed, I imagine that the large majority of characters one might choose at random would not turn out to segregate according to Mendel's laws; thus, I assume that Mendelian segregation is quite improbable , even given the truth of Mendel's laws. We nevertheless accept the rare instances of Mendelian segregation as strong corroboration of Mendel's laws because the *a priori* probability of Mendelian segregation is even smaller than its probability given the truth of Mendel's laws, so that the corroboration index (which is the ratio of the two probabilities) is very much larger than 1. Conversely, we dismiss the frequent instances of non-mendelian segregation as counter-evidence, because non-mendelian segregation is not much less likely given the truth of Mendel's theory than it is *a priori*, so that the corroboration index is not much less than 1.

I can now explain more fully why I think it may be very difficult to corroborate the non-darwinian theory of repetitive DNA that I described earlier, even if that theory is basically correct. As a stochastic theory, it is fully consistent with an infinite variety of particular repetitive patterns. I cannot now think of a particular observation that would be much more probable, given the truth of the non-darwinian theory, than the same observation would be if we were unenlightened by the theory. Hence, for all evidence that I conceive of as likely to be obtained, the corroboration index

$$\langle \text{evidence} \mid \text{non-darwinian theory} \rangle / [\langle \text{non-darwinian theory} \rangle \langle \text{evidence} \mid \text{non-darwinian theory} \rangle + \sum_i \langle \text{alternative theory(i)} \rangle \langle \text{evidence} \mid \text{alternative theory(i)} \rangle]$$

would not be much larger than 1; such evidence would not weigh heavily in favor of the non-darwinian theory.

A THEORY OF SCIENTIFIC THEORIES

I would like to summarize my reflections on theories of scientific theories by proposing a unified framework in which the strength of a scientific theory might be assessed. It is by no means an original proposal; in particular, my approach is very similar to that of Grover Maxwell (MAXWELL 1974). I suggest that in principle we might be able to compute the probability that a particular scientific theory [theory(0)] is true, given the evidence we have on hand. In my proposal, this probability (symbolized <theory(0)|evidence>) equals the product of the *a priori* probability of the theory (given our background knowledge, including any previous evidence) times the corroboration index for the new evidence. Hence

$$\langle \text{theory(0)} \mid \text{evidence} \rangle = \frac{\langle \text{theory(0)} \rangle \langle \text{evidence} \mid \text{theory(0)} \rangle}{\sum_i \langle \text{theory(i)} \rangle \langle \text{evidence} \mid \text{theory(i)} \rangle} .$$

The probability <theory(0)|evidence> becomes a new *a priori* probability of theory (0); similarly the same evidence allows the calculation of <theory(i)|evidence> for all the other theories; these numbers become the new *a priori* probabilities of those alternative theories. These new *a priori* probabilities can then be used in calculating the effects of even newer evidence. This formulation accords with the intuitive notion of scientific progress as the continual reevaluation of the likelihood of alternative theories in the light of accumulating evidence.

Actually, assuming all the quantities involved are meaningful, my proposal is true logically. In that case, it is identical to a certain form of Bayes' theorem, which is proved in elementary books on probability. But it is admittedly a thorny question what meaning can be attached to the *a priori* probabilities of theories.

It seems to me that in general this Bayesian formula more accurately reflects the way working scientists actually weigh the merits of their theories than do conventional ideas about scientific method, even though it is often in terms of the latter that scientists write. And when the Bayesian formulation departs most radically from conventional ideas - namely in the equal intrinsic weight it gives to *a priori* plausibility and empirical corroboration in the overall assessment of a theory - it seems to me to represent a more rational approach.

I'm afraid I'm not, like the type of lecturer I disparaged earlier this evening, going to be able to close this talk by

pointing out that my proposal has the virtue of being testable. I suppose, then, that boos and catcalls are now in order.

ACKNOWLEDGMENTS

This work was supported by grant GM22686 to the author from the U.S. National Institutes of Health.

LITERATURE CITED

BRITTEN, R. J. and D. E. KOHNE 1968 Repeated sequences in DNA. Science 161: 529-540.

FRY, K. and W. SALSER 1977 Nucleotide sequence of HS-alpha satellite DNA from Kangaroo rat Dipodomys ordii and characterization of similar sequences in other rodents. Cell 12: 1069-1084.

KIMURA, M. 1968 Evolutionary rate at the molecular level. Nature 217: 624-626.

KIMURA, M. 1976 The neutral theory as a supplement to Darwinism. Trends in Biochem. Sci. 1: 248-249.

MAXWELL, G. 1974 Corroboration without demarcation. Pp. 292-315. The Philosophy of Karl Popper. (Schilpp, P. A., Ed.) Open Court, Lasalle, IL.

MILKMAN, R. 1976 Molecular evolution - selection is the major determinant. Trends in Biochem. Sci. 1: 152-154.

POPPER, K. 1963 Conjectures and Refutations: The Growth of Scientific Knowledge. Basic Books. New York.

SMITH, G. P. 1976 Evolution of repeated DNA sequences by unequal crossover. Science 191: 528-535.

WRIGHT, S. 1931 Evolution in Mendelian populations. Genetics 16: 97-159.

Drs. Donald L. Riddle and George P. Smith at the Symposium