CHAPTER 12
A Biochemical Echo of Typology

> . . . as we here and there see a thin straggling branch springing from a fork low down in a tree and which by some chance has been favoured and is still alive on its summit so we occasionally see an animal like Lungfish which in some small degree connects by its affinities two large branches of life.

We have seen that at a morphological level the pattern of nature seems to correspond reasonably well with the old nineteenth-century typological model. Nearly all known groups appear to be isolated and well defined and clear sequential patterns whereby one class is linked to another through linear series of transitional forms are virtually unknown. Moreover, classification procedures invariably result in orderly hierarchic schemes from which overlapping classes indicative of sequential relationships are emphatically absent.

However, no matter how much the diversity of nature may appear to conform to the theory of types at a morphological level, no matter how much all cats, all birds, all angiosperms, all mammals or all vertebrates may seem to be equally representative of their respective groups, there is no way of quantifying such conclusions. Judging relationships in terms of morphological characteristics is bound to involve an element of subjectivity. On purely morphological grounds there is no way of measuring the *exact* distance between two organisms in strictly mathematical terms. We cannot, for example, quantify the difference between a cat and a dog and compare it with, say, the difference between a cat and a mouse. We assume that a cat and a dog are closer than a cat and a mouse, but how secure are such judgments?

There is also simply no way of making a quantitative measurement of complexity at a morphological level. A mammal may "look" more complex than a fly, but whether this is true and, if it is, how much

more complex in strictly quantitative terms cannot be determined on grounds of morphology. Again, the vertebrate central nervous system gives every "appearance" of being a vastly complex system, but for all we know it may require less genetic information to specify for the vertebrate brain than for the pentadactyl limb!

From the founding of modern biology by Linnaeus in the eighteenth century right up to the 1960s, the only way biologists had of classifying organisms and assessing the differences between species was by comparing their structure at a gross morphological level. Comparative biology was no more nor less than comparative anatomy.

The molecular biological revolution has dramatically changed this situation by providing an entirely new way of comparing organisms at a biochemical level. In the late 1950s it was found that the sequence of a particular protein, such as, say, haemoglobin, was not fixed but varied considerably from species to species. The amino acid sequence of a protein from two different organisms can be readily compared by aligning the two sequences and counting the number of positions where the chains differ. In exactly the same way two sequences of letters can be compared. For example, sequences A and B in the diagram below differ at four positions.

(A) C D K N I A A T Y L V G H I T T E N B Y
(B) C B K N I D A T Y L V G H I C T E M B Y
1 2 3 4

There are twenty letters in each of the two sequences above so they can be said to exhibit a twenty percent sequential divergence.

Similarly, the differences between two proteins can be quantified exactly and the results of these measurements can provide an entirely novel approach to measuring the differences between species. The list below gives part of the sequence of the protein cytochrome C from a variety of species:

Horse	Gly-Leu-Phe-Gly-Arg-Lys-Thu-Gly-GluNH_2-Ala-Pro
Pig or cow	Gly-Leu-Phe-Gly-Arg-Lys-Thr-Gly-GluNH_2-Ala-Pro
Kangaroo	Gly-Ile-Phe-Gly-Arg-Lys-Thr-Gly-GluNH_2-Ala-Pro
Human	Gly-Leu-Phe-Gly-Arg-Lys-Thr-Gly-GluNH_2-Ala-Pro
Chicken	Gly-Leu-Phe-Gly-Arg-Lys-Thr-Gly-GluNH_2-Ala-Glu
Tuna	Gly-Leu-Phe-Gly-Arg-Lys-Thr-Gly-GluNH_2-Ala-Glu
Moth	Gly-Phe-Gly-Arg-His-Thr-Gly-GluNH_2-Ala-Pro-Gly-Phe-Tyr
Yeast	Gly-Ile-Phe-Gly-Arg-His-Ser-Gly-GluNH_2-Ala-GluNH_2

As work continued in this field, it became clear that each particular protein had a slightly different sequence in different species and that closely related species had closely related sequences. When the haemoglobin sequences in different mammals, such as man and dog, were compared the sequential divergence was about twenty percent, while, when the haemoglobin in two dissimilar species such as man carp were compared, the sequential divergence was found to be about fifty percent.

It was also found that different types of proteins exhibited different degrees of interspecies variation. Cytochrome C, for example, varied less between species than haemoglobin. While the haemoglobin sequences of man and dog differed by twenty percent, their cytochrome sequences varied by only five percent, and while the haemoglobin sequences of man and carp varied by fifty percent, their cytochrome sequences varied by only thirteen percent. Yet whichever protein was chosen, organisms that were close in terms of their haemoglobin sequences were also close in terms of their cytochromes, and the same was true of all other proteins examined.

These results showed that not only did organisms vary at a morphological level in terms of their gross anatomy, but that they also varied at a molecular level as well. It became increasingly apparent as more and more sequences accumulated that the differences between organisms at a molecular level corresponded to a large extent with their differences at a morphological level; and that all the classes traditionally identified by morphological critera could also be detected by comparing their protein sequences. Among the vertebrates, for example, all the major classes identified by morphological criteria can also be readily identified on the basis of molecular comparisons.

Armed with this new technique, biology at last possessed a strictly quantitative means of measuring the distance between two species and of determining the patterns of biological relationships. If it is true, as typology implied, that all the members of one type, however superficially divergent, always conform exactly to the basic eidos of their type, all possessing equally and in full measure all the defining character traits of their type and all standing therefore equidistant in all important aspects of their biological design from the members of other types, might this principle of equidistance be revealed by these new molecular studies? If the divisions in the nature were really as orderly as early nineteenth-century biologists insisted, might this

overall orderliness be confirmed by the new field of comparative biochemistry?

On the other hand, the new molecular approach to biological relationships could potentially have provided very strong, if not irrefutable, evidence supporting evolutionary claims. Armed with this new technique, all that was necessary to demonstrate an evolutionary relationship was to examine the proteins in the species concerned and show that the sequences could be arranged into an evolutionary series. In the diagram below it is obvious that sequence B is intermediate between sequences A and C.

A	*B*	*C*
A	A	A
B	B	C
T	T	T
W	W	W
V	S	S
Y	Y	Y
H	H	H
K	L	L
D	P	P
E	E	T

It is possible to arrange the letter strings in a series where B is intermediate between A and C and to postulate either that A evolved into B and B into C or that B evolved into A and C or C into B into A.

$$A \rightarrow B \rightarrow C \qquad \begin{matrix} A & & C \\ & \searrow \nearrow & \\ & B & \end{matrix} \qquad C \rightarrow B \rightarrow A$$

Whichever theory is correct, such sequential arrangements suggest evolutionary relationships.

The prospect of finding sequences in nature by this technique was, therefore, of great potential interest. Where the fossils had failed and morphological considerations were at best only ambiguous, perhaps this new field of comparative biochemistry might at last provide objective evidence of sequence and of the connecting links which had been so long sought by evolutionary biologists.

However, as more protein sequences began to accumulate during the 1960s, it became increasingly apparent that the molecules were not going to provide any evidence of sequential arrangements in

nature, but were rather going to reaffirm the traditional view that the system of nature conforms fundamentally to a highly ordered hierarchic scheme from which all direct evidence for evolution is emphatically absent. Moreover, the divisions turned out to be more mathematically perfect than even most die-hard typologists would have predicted.

To understand the subsequent pattern that has been revealed by these comparative studies we might start by reviewing the evidence provided by the protein cytochrome C, one of the proteins intimately connected with the production of cellular energy. Because of its fundamental role in biological oxidation it occurs in a wide range of organisms ranging from bacteria to mammals.

All cytochrome C molecules are about one hundred amino acids long, have the same 3D conformation and possess an identical active site or hydrophobic pocket specifically designed to complex tightly with the small iron-containing organic compound haem. Yet, despite the profound correspondence in the basic design of all cytochrome C molecules, their amino acid sequences vary in different organisms. The amino acid sequences of cytochrome C have now been determined in a wide variety of organisms including bacteria, fungi, higher plants and vertebrates.

When comparing a considerable number of sequences, it is convenient to present the data in the form of a percent sequence difference matrix. In the *Dayhoff Atlas of Protein Structure and Function* (1972 edition) there is a matrix with nearly 1089 entries showing the percent sequence difference between thirty-three different cytochromes taken from very diverse species. Part of this matrix is shown in Figure 12.1.

Examination of the percent sequence difference matrix reveals that it is possible to use the cytochrome sequences to classify species into groups and that these groups correspond precisely to the groups arrived at on traditional morphological grounds. It is also apparent that the sequential divergence becomes greater as the taxonomic distance between organisms increases, a finding that would again have been predicted from traditional taxonomic considerations. For example, between horse and dog (two mammals) the divergence is six percent, between horse and turtle (two vertebrates) the divergence is eleven percent, and between horse and fruit fly (two animals) the divergence is twenty-two percent.

However, the most striking feature of the matrix is that each identifiable subclass of sequences is isolated and distinct. Every

		Horse	Dog	Kangaroo	Penguin	Pekin Duck	Pigeon	Turtle	Tuna	Bonito	Carp	Lamprey	Screw-worm	Silkworm	Horn Worm	Castor	Sunflower	Wheat	C. krusei	D. kloeckeri	Yeast	R. rubrum C_2
MAMMALS, BIRDS, REPTILES TELEOSTS, CYCLOSTOMES	Horse	0	6	7	12	10	11	11	18	17	13	15	20	27	26	40	41	41	46	40	42	64
	Dog	6	0	7	10	8	9	9	17	16	11	13	19	23	23	38	39	39	45	38	41	65
	Kangaroo	7	7	0	10	10	11	11	17	17	13	16	22	26	26	38	39	42	46	41	42	66
	Penguin	12	10	10	0	3	4	8	17	17	14	18	22	25	25	40	41	41	45	40	40	64
	Pekin Duck	10	8	10	3	0	3	7	16	16	13	17	20	25	25	38	39	41	45	40	41	64
	Pigeon	11	9	11	4	3	0	8	17	17	14	18	21	25	24	38	39	41	45	40	41	64
	Snapping Turtle	11	9	11	8	7	8	0	17	16	13	18	22	26	27	38	39	41	47	42	44	64
	Tuna Fish	18	17	17	17	16	17	17	0	2	8	18	22	30	28	42	43	44	43	42	43	65
	Bonito	17	16	17	17	16	17	16	2	0	7	18	23	31	29	41	41	42	42	41	41	64
	Carp	13	11	13	14	13	14	13	8	7	0	12	20	25	24	41	41	42	45	39	42	64
	Lamprey	15	13	16	18	17	18	18	18	18	12	0	26	30	31	45	44	46	50	43	45	66
INSECTS	Screw-worm Fly	20	19	22	22	20	21	22	22	23	20	26	0	13	11	40	40	40	43	39	44	66
	Silkworm Moth	27	23	26	25	25	25	26	30	31	25	30	13	0	5	40	40	40	43	39	44	65
	Tobacco Horn Worm Moth	26	23	26	25	25	24	27	28	29	24	31	11	5	0	39	40	38	42	39	42	64
PLANTS	Castor	40	38	38	40	38	38	38	42	41	41	45	40	40	39	0	10	12	45	43	42	66
	Sunflower	41	39	39	41	39	39	39	43	41	41	44	40	40	40	10	0	13	47	44	43	67
	Wheat	41	39	42	41	41	41	41	44	42	42	46	40	40	38	12	13	0	45	41	42	66
YEASTS	Candida krusei	46	45	46	45	45	45	47	43	42	45	50	43	43	42	45	47	45	0	23	25	72
	Debaryomyces kloeckeri	40	38	41	40	40	40	42	42	41	39	43	39	39	39	43	44	41	23	0	27	67
	Baker's Yeast	42	41	42	40	41	41	44	43	41	42	45	42	44	42	42	43	42	25	27	0	69
BACTERIA	Rhodospirillum rubrum C_2	64	65	66	64	64	64	64	65	64	64	66	64	65	64	66	67	66	72	67	69	0

Figure 12.1: The Cytochromes Percent Sequence Difference Matrix. (from Dayhoff)[1]

MAMMALS	
Human	65
Monkey	64
Pig	64
Horse	64
Dog	65
Whale	65
Rabbit	64
Kangaroo	66

BIRDS	
Chicken	64
Penguin	64
Duck	64
Pigeon	64
REPTILES	
Turtle	64
Rattlesnake	66
AMPHIBIANS	
Bullfrog	65

TELEOSTS	
Tuna	65
Bonito	64
Carp	64
ELASMOBRANCHS	
Dogfish	65
CYCLOSTOMES	
Lamprey	66

INSECTS	
Fruit Fly	65
Screw-worm	64
Silkworm	65
Tobacco Horn Worm Moth	64

ANGIOSPERMS	
Mung-bean	66
Sesame	65
Castor	69
Sunflower	69
Wheat	66

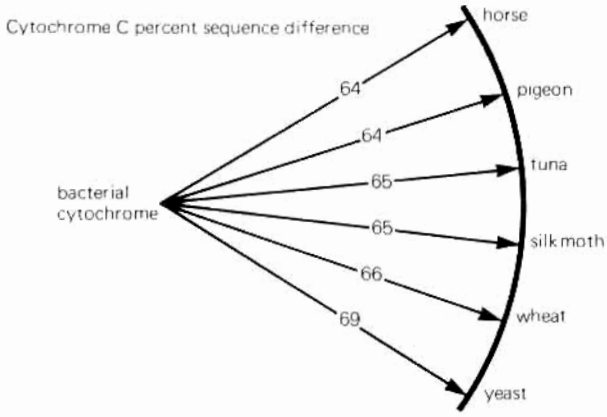
YEASTS	
Candida krusei	72
Debaryomyces kloeckeri	67
Baker's yeast	69
Neurospora crassa	69

Figure 12.2: The Molecular Equidistance of all Eucaryotic Organisms from Bacteria. *Percent sequence divergence between the cytochrome C_2 of Rhodospirillum rubrum and various eucaryotic cytochromes.* (from Dayhoff)[2]

sequence can be unambiguously assigned to a particular subclass. No sequence or group of sequences can be designated as intermediate with respect to other groups. All the sequences of each subclass are equally isolated from the members of another group. Transitional or intermediate classes are completely absent from the matrix.

A table which illustrates the dramatic absence of intermediates is seen in Figure 12.2 which lists thirty-three comparisons between the bacterial cytochrome C of *Rhodospirillum rubrum* and cytochromes of a wide variety of eucaryotic organisms. (Eucaryotic refers to organisms whose cells possess a nucleus, ie, all non-bacterial groups.) These comparisons indicate that all the eucaryotic cytochrome sequences are almost exactly the same distance from their bacterial homologue.

In the list in Figure 12.2, if three yeasts are excluded from the list, then the remaining eucaryotic cytochromes, from organisms as diverse as man, lamprey, fruit fly, wheat and yeast, all exhibit a sequence divergence of between sixty-four percent and sixty-seven percent from this particular bacterial cytochrome. Considering the enormous variation of eucaryotic species from unicellular organisms like yeasts



to multicellular organisms such as mammals, and considering that euraryotic cytochromes vary among themselves by up to about forty-five percent, this must be considered one of the most astonishing findings of modern science.

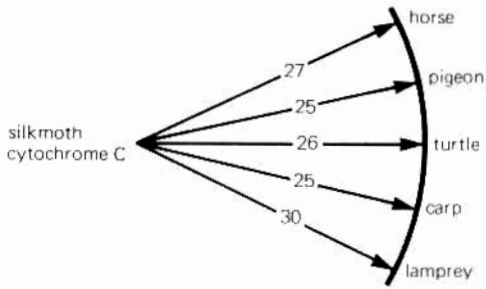
It means that no eucaryotic cytochrome is intermediate between the bacterial cytochrome and other eucaryotic cytochromes. As far as the bacterium is concerned, all the eucaryotes are equally distant. All the eucaryotic cytochromes are as a class isolated and unique. No intermediate type of cytochrome exists to bridge the discontinuity which divides the living kingdom into these two fundamental types. The bacterial kingdom has no neighbour in any of the fantastically diverse eucaryotic types. The "missing links" are well and truly missing.

But even among the eucaryotic cytochromes at a slightly lower taxonomic level the same isolation and uniqueness of subclasses, the same lack of intermediates are observed. Examination of the sequential divergence among eucaryotic cytochromes reveals three basic subgroups: the yeasts, the plants, and the animals (see matrix in Figure 12.1). Each type is quite isolated. Just as there are no intermediates to bridge the gap between procaryotes and eucaryotes so there are no intermediate types among these three basic eucaryotic groups. Although the distance among the three eucaryotic types is less than that between the procaryotes and the eucaryotes, the divisions among the three fundamental types are no less clear and unambiguous. Each type is just as unique and isolated from the others. The yeast cytochromes are uniformly isolated from the cytochromes of all other eucaryotes. The same ordered isolation is seen when the plants are

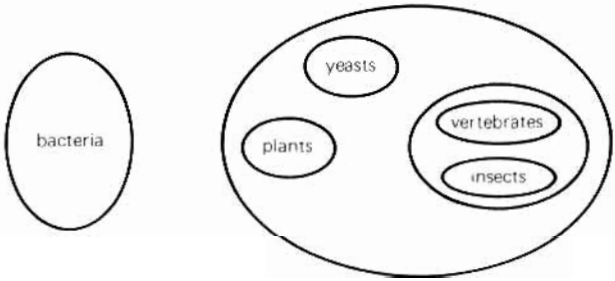
compared with other eucaryotes. Similarly, no animal cytochrome is intermediate between the animals and the other two eucaryotic groups.

At a still lower taxonomic level the same phenomenon is observed. From the matrix in Figure 12.1 it is clear that the insects and vertebrates are closely related, but when comparisons are made between insect species and a variety of vertebrate groups, no vertebrate group is primitive or in any sense a link between phylum Arthropoda and phylum Vertebrata. All the many diverse vertebrate types, including cyclostomes and mammals, are uniformly distant from the insects.

The diagram below gives the percentage sequence divergence between the silk moth and various vertebrate cytochromes:[3]

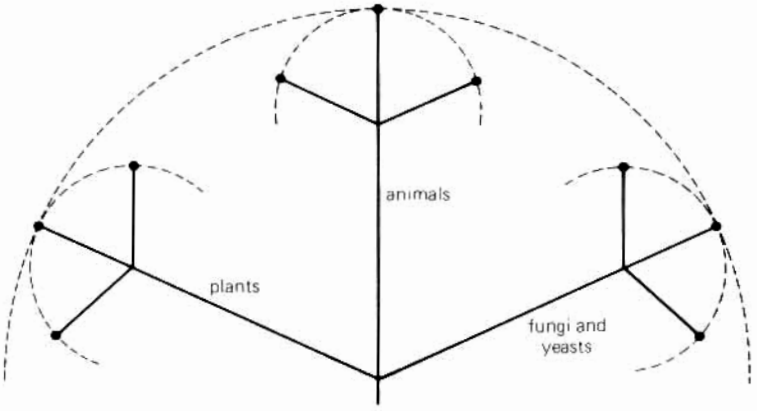


From the sequential divergence of their cytochromes it is possible to classify the living kingdom into various divisions. The primary division is clearly between bacteria and eucaryotes. The eucaryotes are subdivided into three distinct classes, yeasts, plants and animals; the animals can be subdivided into two further subclasses, insects and vertebrates.



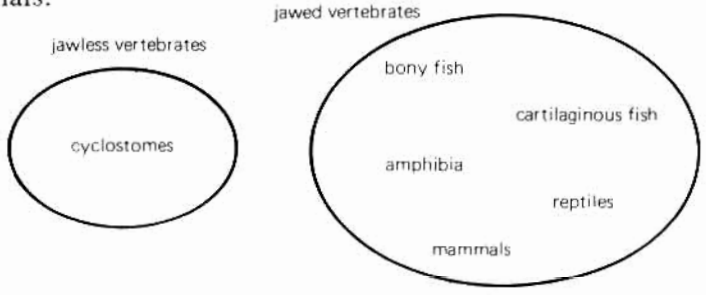
Each class is isolated and unique. No classes are intermediate or partially inclusive of other classes. The isolation of each class becomes greater as the taxonomic hierarchy is ascended, but even relatively closely related classes such as insects and vertebrates are still clearly distinguished.

The pattern of nature implied by these findings is depicted in the diagram below. In terms of their cytochromes, the three major eucaryotic kingdoms may be thought of as equidistant from a common hypothetical archetype, while within each group all the members are similarly equidistant from the hypothetical archetype of their group.

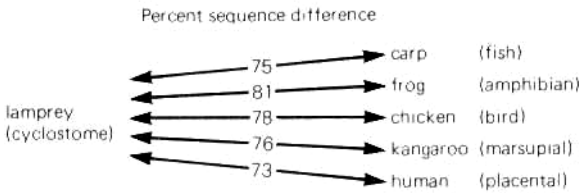


Note how closely the cytochrome pattern seems to correspond to the circumferential model of nature of the early nineteenth-century typologists.

Consider next the divisions within the vertebrate phylum. Based on the degree of similarity of their proteins, the vertebrates can be clearly divided into two fundamental divisions, the jawless cyclostomes and the higher jawed vertebrates – the fish, amphibia, reptiles and mammals.

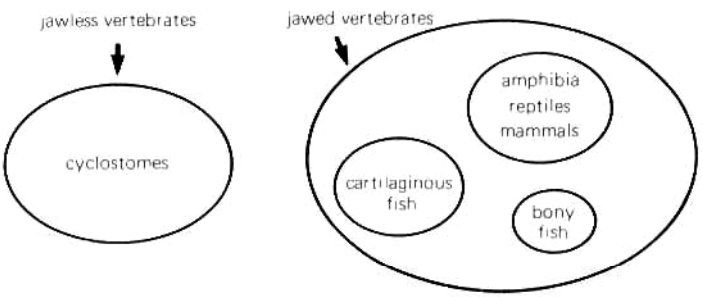


In itself, the existence of this fundamental division is not surprising, as it corresponds exactly with a traditional division based on morphological characteristics. But the strange thing about the division is the fact that although the proteins in the higher jawed vertebrate groups – fish, amphibia, reptiles and mammals – are widely divergent when they are compared with those of cyclostomes, invariably the degree of difference is always the same. The almost mathematical perfection of the isolation of the two fundamental classes at a molecular level is astonishing! The figure below gives the percent sequence difference between the haemoglobin of the lamprey and a variety of jawed vertebrates, taken from a sequence difference matrix of the vertebrate globins in the *Dayhoff Atlas of Protein Structure and Function:*[4]



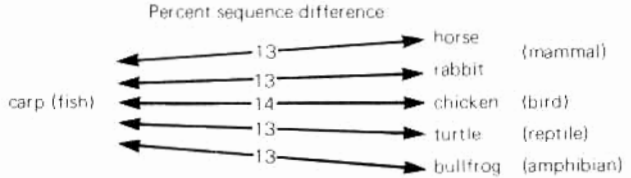
There is not a trace at a molecular level of the traditional evolutionary series: cyclostome → fish → amphibian → reptile → mammal. Incredibly, man is as close to lamprey as are fish! None of the higher jawed vertebrate groups is an any sense intermediate between the jawless vertebrates and other jawed vertebrate groups.

The higher vertebrate groups can also be divided into subgroups based on their degree of molecular similarity. Several basic subdivisions are revealed. One contains the many types of bony fish, another



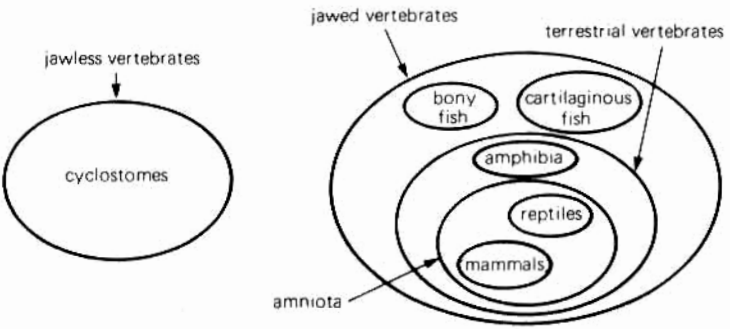
the cartilaginous fish and another the terrestrial vertebrate groups, amphibia, reptiles, and mammals.

These subdivisions are not very surprising as they correspond exactly with traditional subdivisions derived from morphological studies. But again there is the same strangely ordered aspect to the pattern of the molecular divisions. When the various terrestrial vertebrate groups, amphibia, reptile, or mammal, are compared with fishes all are equally isolated. The figure below gives the percent sequence difference between cyclostome C in carp and various terrestrial vertebrates.



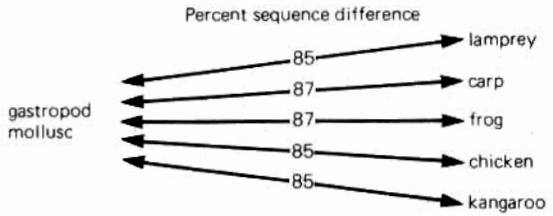
Again, an extraordinary mathematical exactness in the degree of isolation is apparent. So, although cytochrome C sequences varied among the different terrestrial vertebrates, all of them are equidistant from those of fish. At a molecular level there is no trace of the evolutionary transition from fish → amphibian → reptile → mammal. So amphibia, always traditionally considered intermediate between fish and the other terrestrial vertebrates, are in molecular terms as far from fish as any group of reptiles or mammals! To those well acquainted with the traditional picture of vertebrate evolution the result is truly astonishing.

The terrestrial vertebrates can themselves be divided into two basic classes, by virtue of their molecular similarities. One class contains the amphibia, the other the reptiles and mammals. Again the subdivision corresponds to that based on classical morphological grounds, but whichever species are taken for comparative purposes the distance between amphibian species on the one hand and mammalian and reptilian species on the other is always the same. No amphibian species is midway between other amphibia and the reptiles and the mammals. Similarly no reptilian or mammalian species is closer to amphibia than any of the others. The vertebrate classification scheme could be represented thus:



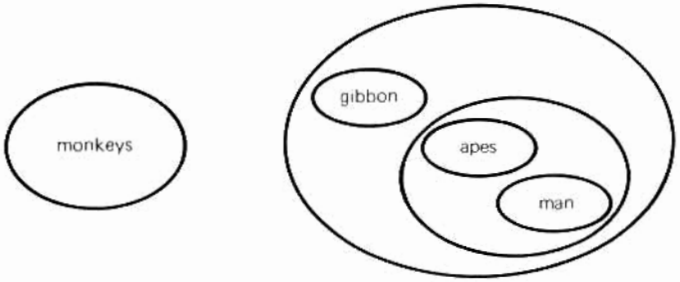
The classification system that is derived from these comparative molecular studies is a highly ordered non-overlapping system composed entirely of groups within groups, of classes which are inclusive or exclusive of other classes. There is a total absence of partially inclusive or intermediate classes, and therefore none of the groups traditionally cited by evolutionary biologists as intermediate gives even the slightest hint of a supposedly transitional character.

The molecules give no support to the traditional view of the vertebrates as a series of increasingly advanced classes leading from the cyclostomes to the mammals. In fact, when the vertebrates are compared with non-vertebrate organisms, all types are equidistant apart. The diagram below gives the percent sequence divergence between the haemoglobin in a snail and that of various vertebrate species.



On the evidence of the protein sequences we cannot classify the lamprey as primitive with respect to other vertebrates, nor in any sense as intermediate between higher vertebrates and the invertebrates. All we can safely infer about the cyclostomes is that they represent a highly specialized and isolated vertebrate group.

Furthermore, it is not only the major divisions which can be subdivided into non-overlapping classes; the same phenomenon holds to quite minor subdivisions of the animal kingdom, even where the actual biochemical differences between species is relatively trivial.[5] For example, when classifying the primates (the monkeys, apes and man) by comparing the differences in their protein sequences,[6] the result is again an entirely non-overlapping system of classes.



The same picture emerges when DNA or RNA sequences in different species are compared. A fascinating case which illustrates this is the recent work carried out by Carl Woese's group at the University of Illinois, which led to the discovery of a new primary kingdom consisting of an interesting group of anaerobic bacteria possessing the unique capacity to generate the gas methane.[7]

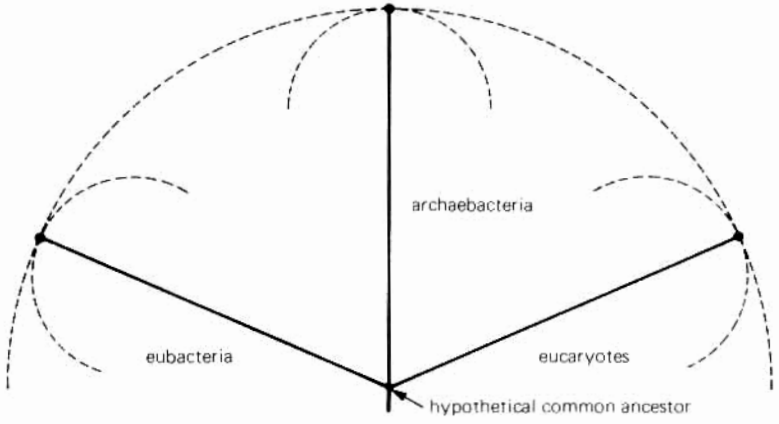
The discovery that the methanogens belonged to a new, quite distinct, living kingdom was made as Woese's group was comparing the RNA sequences of the large RNA molecules which form part of the protein synthetic apparatus (termed ribosomal RNA) in a wide variety of organisms. They showed that, on the basis of the degree of similarity of their RNA sequences, living organisms could be grouped into three primary kingdoms. The first contained all of the typical bacteria, named by Woese eubacteria. A second group contained all species of higher plants and animals, while a third contained the relatively unknown methanogens.

Professor Woese speculated that methanogenic bacteria may have been the first bacteria on earth because of their capacity to thrive in anoxic conditions and their unique capacity to manufacture methane from H_2 and carbon dioxide; and he has therefore termed them archaebacteria. There is however, no basis for believing that the methanogens really are archaebacteria on the grounds of sequential

comparisons. Because, just as in the case of classes derived from comparative protein sequences, none of the three classes, archaebacteria, eubacteria, or eucaryotes is intermediate with respect to the other classes, none of them can be designated ancestral or primitive with respect to the other.

The discovery that the methanogens belong to a quite new division of the living kingdom is yet another case which illustrates one of the main themes of this book, that whenever new types of organisms are occasionally discovered they never turn out to be ancestral to known groups but stand related only as sister groups in keeping with the thesis that nature's basic order is circumferential rather than sequential.

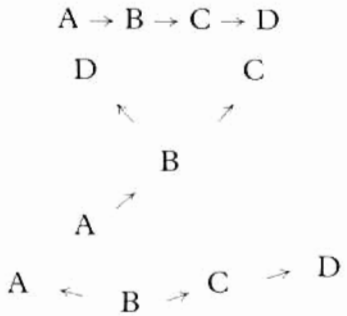
There is no evidence at all for evolutionary transformations in this sequencing data. The RNAs tell the same story as the proteins! On the basis of their RNA sequences, the three primary kingdoms stand equidistant apart, and equidistant from a theoretical common primeval ancestor.



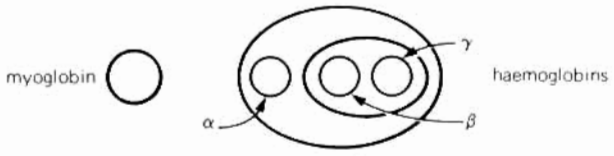
The above discussion has been concerned with comparisons of related molecules in different species. But there are many cases where a family of related proteins occur in the same species. For example, in man there are four related members of the haemoglobin family of proteins. There are three haemoglobins proper and of these α haemoglobin and β haemoglobin are found in the adult red blood cell while γ haemoglobin is found in the red cells of the fetus and newborn. Another member of the haemoglobin family is myoglobin

which is found in both adult and fetal muscle cells and acts as an oxygen reservoir. These four proteins are very similar in terms of function, overall 3-D configuration, as well as amino acid sequence. The only evolutionary explanation that makes sense and has ever been proposed to account for their close similarities is that all four proteins originally evolved from a common precursor. The details and merits of the various evolutionary schemes that have been proposed need not concern us. The really significant finding that comes to light from comparing the proteins' amino acid sequences is that it is impossible to arrange them in any sort of evolutionary series.

Where there are four related protein sequences, say A B C D, a number of different evolutionary arrangements might be theoretically envisaged.



However, when the sequences of the haemoglobins in man are compared we find that myoglobin is quite distinct from the α β and γ haemoglobins, but unfortunately α β and γ haemoglobins are all equally distant from myoglobin. Further, when we compare the sequences of the α β and γ haemoglobins we find that, although β and γ are much closer together than either are to α, again the distances between β-γ and between α-γ are equal. We may classify the haemoglobins but the classification that results is the same groups-within-groups that we have seen when we compare different species at a molecular level.



Thousands of different sequences, protein and nucleic acid, have now been compared in hundreds of different species but never has

any sequence been found to be in any sense the lineal descendant or ancestor of any other sequence. Anyone who doubts this need only consult the sequence difference matrices given in Dayhoff's standard reference book *Atlas of Protein Structure and Function*, available in any major library.

It is now well established that the pattern of diversity at a molecular level conforms to a highly ordered hierarchic system. Each class at a molecular level is unique, isolated and unlinked by intermediates. Thus molecules, like fossils have failed to provide the elusive intermediates so long sought by evolutionary biology. Again, the only relationships identified by this new technique are sisterly. At a molecular level, no organism is "ancestral" or "primitive" or "advanced" compared with its relatives. Nature seems to conform to the same non-evolutionary and intensely circumferential pattern that was long ago perceived by the great comparative anatomists of the nineteenth century.

One of the most remarkable features of these new biochemical discoveries is undoubtedly the way in which the pattern of molecular diversity seems to correspond to the predictions of typology. With very few exceptions the members of each defined taxa are always equally divergent whenever an outgroup comparison is made. Perhaps the only finding which does not seem to flow naturally from the typological model is that the degree of morphological divergence often does not seem to agree with the degree of molecular divergence. For example, the degree of molecular divergence among frogs, which are all morphologically very similar, is as great as that between mammals, which are morphologically very diverse.[8] Similarly, the proteins of conifers are as equally divergent as those of the flowering plants, a group which appears to be far more divergent than the conifers at a morphological level.[9] But despite those anomalies, all in all, the basic axioms of typology, that all the members of each type conform to type, that intratype variation is limited and type specific, so that when outgroup comparisons are made the subgroups of the type stand equidistant from more distantly related groups, hold universally throughout the entire realm of nature. This does not mean, of course, that typology is necessarily correct. But if we accept that closeness to empirical reality is the only criterion by which to judge alternative theories, we would, if strictly impartial, be forced to choose Aristotle and the eidos, in favour of Darwin and the theory of natural selection. There is little doubt that if this molecular evidence

had been available one century ago it would have been seized upon with devastating effect by the opponents of evolution theory like Agassiz and Owen, and the idea of organic evolution might never have been accepted.

This new era of comparative biology illustrates just how erroneous is the assumption that advances in biological knowledge are continually confirming the traditional evolutionary story. There is no avoiding the serious nature of the challenge to the whole evolutionary framework implicit in these findings. For if the ancient representatives of groups such as amphibia, lungfish, cyclostomes and reptiles manufactured proteins similar to those manufactured by their living relatives today, and if, therefore, the isolation of the main divisions of nature was just the same in the past as it is today, if for example ancient lungfish and ancient amphibia were as separate from each other as their present day descendants are, then the whole concept of evolution collapses.

There are of course simply no objective grounds for excluding this possibility and, ironically, it was widely believed by most evolutionists before the full impact of these new comparative studies was realized. Thus, writing in the *Scientific American* in 1963, Zuckerkandl speculates:[10]

> Contemporary organisms that look much like ancient ancestral organisms probably contain a majority of polypeptide chains that resemble quite closely those of the ancient organisms. In other words, certain animals said to be "living fossils", such as the cockroach, the horseshoe crab, the shark and, among mammals, the lemur, probably manufacture a great many polypeptide molecules that differ only slightly from those manufactured by their ancestors millions of years ago.

The only way to save evolution in the face of these discoveries is to make the *ad hoc* assumption that the degree of biochemical isolation of the major groups was far less in the past, that ancient lungfish, for example, were far closer biochemically to ancient amphibia than their present day descendants. There is, however, absolutely no objective evidence that this assumption is correct. The only justification for such an assumption would be if evolution is true, but this is precisely the question at issue!

Given the distinctness of most of the divisions of nature at a morphological level and the absence of *bona fide* ancestors, intermediates or transitional forms, the credibility of evolutionary claims

has had to depend traditionally very largely on "evidence" of a far from conclusive nature – on those instances where with the eye of faith it might be construed that, in Darwin's words, "a species or group like lungfish in some small degree connects by its affinities two large branches of life."[11] Thus the literature of biology is full of claims that this or that group, while not definitively intermediate or ancestral in any aspect of their biology, is at least so "in some degree" and its relationships with other groups may be interpreted in evolutionary terms.

So, according to Professor Romer, one of the leading vertebrate paleontologists:[12]

> The living cyclostomes and the fossil ostracoderms are members of a common stock of *primitive ancestral* vertebrates.
>
> [*emphasis added*]

and the opossums:[13]

> The opossums and their relatives found today in both Western continents are in almost every respect *ideal ancestors* for the whole marsupial group . . . In the late cretaceous of North America are found forms very similar to the living opossums.
>
> [*emphasis added*]

and the amphibians:[14]

> . . . are without question the *basal stock* from which the remaining group of land vertebrates have been derived.
>
> [*emphasis added*]

and among the placental mammals – the insectivores:[15]

> . . . There are still in existence a number of mammals, such as the shrews, moles and hedgehogs, which have retained . . . many *primitive characters*. These forms, grouped with related fossil types as the order Insectivora, are regarded as the most direct modern descendants of the primitive placentals.
>
> [*emphasis added*]

Similarly, despite the absence of clear-cut sequential arrangements, biologists have been able to allude to cases where nature does appear to fall approximately into a sequential pattern. One of the most celebrated cases of sequence is that of the vertebrate classes leading

from the cyclostomes, through fish, amphibia and reptiles to the mammals. While no evolutionist has ever claimed that any of the living representatives of any vertebrate class is directly ancestral with respect to another vertebrate group, it is definitely implied that in terms of their general biology and overall morphology there are clear grounds for viewing the series as a natural phylogenetic sequence.

Potentially, comparative biochemistry by the demonstration of underlying sequential patterns could have added substantially to the credibility of such claims. If the sequence of vertebrate proteins could have been arranged in a series like the letter strings below:

Z	A	B	B	B
G	B	B	D	D
D	D	C	C	C
T	T	T	V	L
G	K	K	H	H
G	G	G	G	K
W	W	T	T	L

CYCLOSTOME → FISH → AMPHIBIAN → REPTILE → MAMMAL

then this would have provided powerful confirmation of the traditional sequential interpretation of the vertebrate classes. But as we have seen, the molecules provide little support for this "sequential" interpretation of the vertebrate classes.

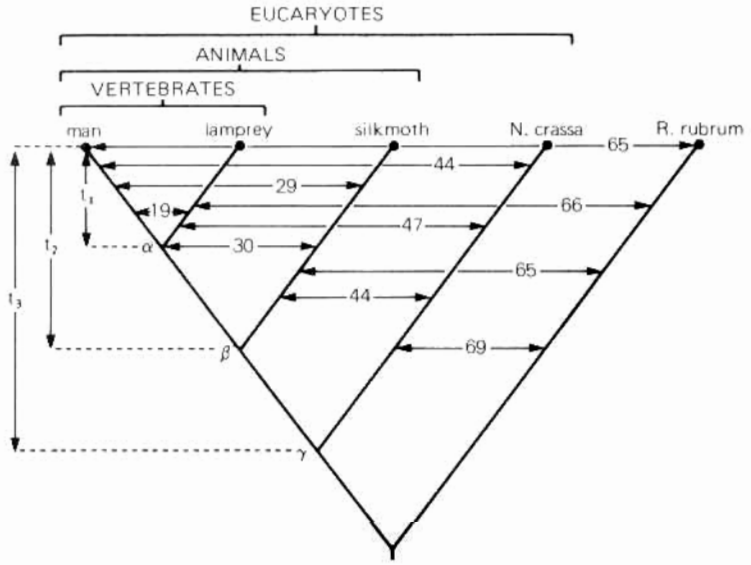
In terms of their biochemistry, none of the species deemed "intermediate", "ancestral" or "primitive" by generations of evolutionary biologists, and alluded to as evidence of sequence in nature, shows any sign of their supposed intermediate status. The *cyclostomes are not primitive biochemically*; they are no closer to any non-vertebrate species than any other vertebrate group. The *opposums are not ideal ancestors* for the whole marsupial group in terms of their biochemistry. At a molecular level they are as far away from any reptile as any other marsupial species. Similarly, *the insectivores are no less indubitably mammalian* at a molecular level than any other mammalian group.

There is ultimately nothing contradictory in the molecular and morphological evidence. As we have already seen, to what extent the vertebrate sequence is actually supported by comparative anatomy is open to debate. There are morphological grounds for arranging the vertebrate classes in a circumferential and typological arrangement which is perfectly in keeping with the molecular evidence.

The fact that lungfish, monotremes and all the other favourite links of evolutionary biology give no hint of their supposed transitional status at a molecular level is perfectly in keeping with the fact that there are many morphological features of their biology which have never been easy to reconcile with their supposedly transitional status and which have always suggested that they represent unique and isolated types.

But by far the most challenging aspect of this new biochemical picture as far as evolution is concerned is the incredible orderliness of all the divisions. We have seen that the sequences of a protein such as haemoglobin vary considerably between the different members of a particular group but that when the sequences of any one group are compared with those of a more distantly related class the sequential divergence is invariably the same. The only way to explain this in evolutionary terms is to propose that since all the different lines of a group diverged each particular protein, such as haemoglobin or cytochrome C, has continued to evolve in each of the lines at its own characteristic uniform rate.

This can be seen by examining the hypothetical evolutionary tree below, where the nodes α β and γ represent presumed points of evolutionary divergence at times t_1, t_2 and t_3, and where the numbers indicate the sequential divergence between the various contemporary cytochromes sequences.



Consider first the evolution of the vertebrate cytochromes. Because they exhibit a considerable degree of sequential divergence (nineteen per cent), yet are equally isolated from non-vertebrate cytochromes, it must be presumed that both sequences have undergone the *same net degree of change in the same time interval* t_1. This means that since their common evolutionary divergence at node α, both cytochrome sequences must have diverged at a net constant rate with respect to absolute astronomical time.

Consider next the evolution of the animal cytochromes. Again because of the considerable degree of intra-group variation (as evidenced by the sequential divergence of twenty nine per cent between silk moth and man, and thirty per cent between silkmoth and lamprey), and because all three sequences are uniformly isolated from the two non-animal cytochromes, then it must be presumed that the three sequences have undergone the *same net degree of change in the same time interval* t_2. Once again, since their common evolutionary divergence at node β, the animal cytochromes must have diverged at a *net* constant rate with respect to absolute astronomical time.

Finally, consider the evolution of the eucaryotic cytochromes. Because of the amount of intra-group variation (as evidenced by the average forty five per cent sequential divergence between *H. crassa* and the three animal cytochromes), and because of the fact that all four sequences are almost exactly the same distance from their bacterial homologue, it must be presumed that all the eucaryotic cytochromes have undergone the same net degree of change in the same time interval t_3. Again since their common evolutionary divergence at node γ, all four cytochrome sequences must have diverged at a net constant rate with respect to absolute time.

We see then that the highly ordered pattern of cytochrome diversity could only have been generated if the overall net rate of sequential divergence had been constant with respect to absolute time in *all* the diverse branches of every class since their common evolutionary origin. Moreover, only if such a strange rule had been repeated over and over again, throughout eucaryotic evolution, following each evolutionary divergence, could it have generated the highly ordered pattern and uniform isolation of each class of eucaryotic cytochrome sequences.

Only if the degree of evolution in a family of molecules such as the cytochromes had been constrained by some kind of time constant mechanism, so that in any one class the degree of change which

occurs is always proportional to the lapse of absolute time, can the ordered pattern of molecular diversity be explained. This remarkable concept is widely known as that of the 'molecular clock hypothesis'. But although such a clock is perfectly capable of accounting for the observed equal divergence of, say, all vertebrate cytochromes from those of insects, no one has been able to explain in precise terms exactly how such a time constant process could work. Rather than being a true explanation, the hypothesis of the molecular clock is really a tautology, no more than a restatement of the fact that at a molecular level the representatives of any one class are equally isolated from the representatives of another class.

The tautological nature of the molecular clock hypothesis is reminiscent of the explanations of the gaps in the fossil record. The proposal put forward to save evolution in the face of the missing links – that connecting links are missing from the fossil record because transitional species are very rare – is essentially tautological. If evolution is true then indeed the intermediates must be very rare. But unfortunately we can only know that evolution is true *after* we have found the transitional types! The explanation relies on belief in evolution in the first place. Similarly, if evolution is true then, yes indeed, the clock hypothesis must also be true. Again the hypothesis gets us nowhere. We save evolution because we believed it in the first place.

But there is an additional twist to the clock hypothesis. As we saw above, different proteins exhibit different degrees of interspecies variation. While haemoglobin sequences differ by fifty per cent between man and carp; cytochrome C differs by only thirteen per cent. To account for the fact that all the haemoglobin sequences of a particular group differ by fifty per cent from another group, while all the cytochrome C sequences differ by only thirteen per cent, it is necessary in evolutionary terms to presume that the molecular clock has ticked at a faster rate in the case of haemoglobin than in the case of cytochrome C; in other words, to propose two molecular clocks ticking at a different rate, one for the haemoglobin family and one for the cytochrome family. However, as there are hundreds of different families of proteins and each family exhibits its own unique degree of interspecies variation, some greater than haemoglobin, some far less than the cytochromes, then it is necessary to propose not just two clocks but one for each of the several hundred protein families, each ticking at its own unique and highly specific rate.

What sort of mutational mechanism might have generated uniform rates of evolution over vast periods of time in vastly dissimilar types of organisms? Basically, there are only two types of changes that can occur to the sequence of the genes specifying for functional proteins: neutral mutations which have no effect on function and are substituted by drift; and advantageous mutations which have a positive effect on function and are substituted by selection.

Unfortunately, neither evolution by genetic drift nor evolution by positive selection is likely to have generated anything remotely resembling a uniform rate of evolution in a family of homologous proteins.

The rate of genetic drift in any gene is directly related to and determined by the mutation rate. This is not controversial. The greater the mutation rate, the greater the speed of genetic drift. One fact that has perhaps lent a certain amount of support to the idea that drift might occur at approximately the same rate is the finding that in higher organisms the observed mutation rate per generation is approximately the same for many genes. The figure usually given for higher organisms is about 10^{-6}/gene/generation.[16] In closely related species such as man and chimpanzee, where the generation rates are similar, one might therefore expect approximately similar rates of drift to have occurred in homologous loci over several generations. But one would not expect similar rates of drift in more diverse types.

The proteins of small rodents, mice for example, are no more divergent than those in primates, elephants or whales, species which have very much longer generation times than rodents. A mouse may go through four to five generations in one year. The time taken by an elephant, a chimpanzee or a man to reach maturity is about fourteen, seven, ten years respectively. This means that at present the generation times of some mammalian species varies by a factor of nearly one hundred. Since the rodent order diverged from the primate, it is practically certain that the line leading to mouse has undergone nearly one hundred times as many reproductive cycles as that leading to man. If mutation rates are approximately constant per generation how then could drift have generated equal rates of genetic divergence in mice and men?

Among the insects there is an even greater diversity of generation times. A fruit fly may undergo a reproductive cycle in two weeks,[17] perhaps twenty generations per year. In the case of the Cicada one species has a generation cycle of seventeen years.[18] The generation rate of the Cicada is nearly one thousand times slower than that of

fruit fly. The time of origin of the modern insect orders, families, and genera is not known, but many insect species are practically identical to the fossils found in Scandinavian amber some fifty million years ago.[19] If the differential generation times observed in modern species had only been maintained for as much as fifty million years, the fruit fly would have undergone fifty thousand million more generative cycles than the Cicada. Yet the proteins of different insect orders are equally divergent from those of vertebrates!

The plausibility of uniform drift shrinks even more when more diverse types of organisms are compared. Some higher plants, trees for example, only reach sexual maturity after eighty years.[20] Micro-organisms such as yeast have generation times measured in minutes. The difference in generation times between such species is of the order of 10^5.

A uniform rate of drift in different lines is likely only if mutation rates in different organisms are uniform per unit time. This may be so in lines which have similar generation rates, such as man and chimpanzee or dog and fox. But many organisms often have vastly different generation times and all the evidence suggests that mutation rates per unit time are often very different in different species, varying by at least one to three orders of magnitude (see Figure 12.3).

Only if the rate of mutation in homologous proteins in different organisms, was for some mysterious reason adjusted so that it was constant with respect to absolute time would uniform rates of drift occur. As Ewens remarked at a recent symposium:[21]

> I note the well-known fact that the neutral theory predicts a constant rate of substitution per generation, whereas we appear to observe more a constant rate per year. In some of the species for which protein sequence comparisons have been made, there is a difference of one or even two orders of magnitude in generation time. It surely gets us nowhere simply to assume that the mutation rate adjusts itself in species of different generation time so that constant rates per year will arise.

Unfortunately, all the evidence suggests that in different groups of organisms the mutation rate per unit of absolute time is vastly different and this effectively excludes drift as a mechanism for the generation of uniform rates of evolution. On top of this there is the additional difficulty of envisaging how drift could have occurred at

ORGANISM	MUTATION RATE PER NUCLEOTIDE PER YEAR	
E. coli	0.7×10^{-6}	(a)
Drosophila	2.5×10^{-8}	(b)
Mouse	3.0×10^{-9}	(c)
Man	1.0×10^{-10}	(d)

Figure 12.3: Mutation Rates per Unit Time.
(a) assuming a mutation rate per base pair replication of 2×10^{-10} and 10 cell divisions per day.[22]
(b) assuming a mutation rate per cistron per generation of 10^{-6} that each cistron consists of 1000 nucleotides and a generation time of 2 weeks.[23]
(c) assuming a mutation rate per cistron per generation of 10^{-6} that each cistron consists of 1000 nucleotides and a generation time of 4 months.[24]
(d) assuming a mutation rate per cistron per generation of 10^{-6} that each cistron consists of 1000 nucleotides and a generation time of 10 years.[25]

different rates in different genes to account for the different rates of evolution in different families of homologous proteins.

One idea that has been put forward is that different proteins are under different functional constraints, which may have permitted some genes to have evolved faster than others. Alan Wilson, an authority in this field, wrote recently:[26]

> The proteins that evolve most slowly are supposed to have the highest proportion of sites at which the functional constraints are particularly severe. According to this view, nearly every mutation that could occur in the gene for histone 4* would be deleterious to the function of that histone.
>
> Conversely, the most rapidly evolving proteins are supposed to have the largest proportion of sites at which more than one residue would be

*The histones are a group of proteins which are intimately associated with the DNA in all eucaryotic, that is nucleated organisms which exhibit an astonishing invariance in their amino acid sequence. Histone 4 is the most invariant of all the histones.

compatible with function. Fibrinopeptides* are often cited as examples of the latter type.

Like many explanations of phenomena which are on the face of it difficult to reconcile with traditional evolutionary models, the "functional constraints" hypothesis is largely tautological. Although it is put forward as a solution to the problem of different rates of protein evolution in different families of molecules, the only evidence for the hypothesis is the observation it claims to explain.

Just because some vertebrate haemoglobins such as carp and man differ from one another at up to eighty amino acid sites – while their histones are identical, it cannot be inferred from this that the histones are under more stringent selective constraints. Similarly, we do not conclude that the selective pressures on vertebrate limbs are any less intensive than those on vertebrate spinal columns merely because the former exhibit much greater interspecies diversity than the latter.

Moreover, there is not a scrap of empirical evidence to suggest that there is any systematic difference in the tolerance of different functional proteins to mutational change. As Wilson concedes:[27]

> ... we are not aware of direct experimental evidence showing rigorously that histone function is especially sensitive to amino acid substitution or that fibrinopeptide function is especially insensitive to amino acid substitution. Experimental studies would require that quantitative in vitro assays for the specific functions of histone 4 and fibrinopeptides be available. These have not been developed for histones, fibrinopeptides, or, indeed, most of the proteins whose evolutionary rates are listed.

Further theoretical arguments can be advanced against the idea. Most of the functional criteria which must be satisfied in amino acid sequences, for example, those related to protein stability and the necessity for folding algorithms (see Chapter Ten), are of a general nature and are unlikely to differ in different proteins or in different species or at different times in the past.

The degree of stringency of the criteria for protein stability, for example, is almost certainly the same in all existing proteins and was probably the same in blue green algae 3,500 million years ago unless,

* The fibrinopeptides are two short amino acid sequences which are removed from the protein fibrinogen during the process of blood coagulation.

of course, one is prepared to presume that the basic physical constants have changed during geological time so that the nature of weak chemical interactions, hydrogen bonds etc, and their influence on the stability of α helices and β pleated sheet* formations are different today than in the past.

Similarly, the criteria for function of the active sites of enzymes could hardly vary much from protein to protein. *All* active sites depend on an exact atomic fit between the substrate and protein molecule. It is difficult to see why the tolerance levels should vary significantly.

Considering the vast complexity of all gene sequences it seems extremely unlikely that the functional constraints operating on the great majority of proteins, apart possibly from a few "junk" spacer sequences, vary according to any sort of systematic pattern. It is much, very much, more likely that the overall constraints on most protein sequences are somewhat similar in different proteins and in different species and would have been so over hundreds of millions of years of evolution.

Again, it is the sheer universality of the phenomenon – the necessity to believe that the functional constraints in *all* the members of a particular protein family, say A, in *all* diverse phylogenetic lines for *all* of hundreds of millions of years have remained precisely five times as stringent as those operating on the members of another protein family, say B – which fatally weakens the theory.

But if neutral drift gets us nowhere, selectionist explanations fare no better. It is very difficult to understand why all the members of a particular family of proteins, such as the haemoglobins or the cytochromes, should have suffered the same number of advantageous mutations since their common divergence. Selectionist explanations are particularly implausible in the case of the living fossils. While most species make only, what is on a geological time scale, a fleeting appearance in the fossil record, often no more than a few million years, some have persisted almost unchanged for hundreds of millions of years down to modern times – these we call the living fossils.

These great survivors have always held a great fascination for scientists because they are biological time capsules, preserving in their morphology, physiology and behaviour a pattern of life from

* α helices and β pleated sheet conformations are highly ordered configurations adopted by the folded amino acid sequences in many different kinds of proteins.

the remote past, but in the context of the molecular clock hypothesis they have taken on an added significance. The lungfish is a classic example. These remarkable fish are found in the swamps, rivers and lakes of central Africa. As well as gills the lungfish also possesses an efficient lung which it uses to survive during the severe droughts which periodically afflict the African plains. At the onset of the dry season, as the depth of the water decreases and while the mud is still soft enough for burrowing, the fish digs itself a bulbous cavity that opens to the surface through a small hole and there, just a few feet beneath the parched surface, curled up and encased in its cocoon of mud, it lies dormant. Over the months it slowly becomes severely dehydrated; its skin dries and wrinkles so that it has a lifeless mummified appearance; but as soon as the waters return it emerges from its tomb, takes on water, and returns to its active existence as a fish.

The modern lungfishes are members of an ancient group of fishes which are closely related to the rhipidistian fishes, the group considered almost directly ancestral to the amphibia. Lungfish almost identical to those of modern Africa are found as fossils in the rocks of the Devonian era 350,000,000 years ago alongside fossils of the earliest amphibians and the very fish groups from which the amphibia supposedly arose. Through millions of years since Devonian times, uninfluenced by all the massive changes in the Earth's crust and fauna, while the ancient super-continents of Gonwanaland and Laurentia fragmented, while the dinosaurs came and went, the lungfishes continued performing their unique ritual of survival.

In evolutionary terms the lungfish and other living fossils are in a very real sense like samples drawn an eternity ago from near the main course of the stream of vertebrate life. While the tree of evolution continued to grow in all directions above them, they remained the same, so that over the eons of time they were increasingly left behind, increasingly primitive and ancestral with respect to the newer groups.

Yet the proteins of lungfish are just as far from lamprey as any other fish, amphibian or mammalian group! If we are to explain this in terms of selection we must presume that the proteins of this living fossil have been subject to the same net rate of advantageous amino acid replacement over four hundred million years as the proteins of organisms which have been morphologically transformed out of all recognition over the same period of time. But this is verging on *reductio ad absurdum* because it necessitates a complete divorce between adaptive change at a molecular and at a morphological level.

Consider the case of the haemoglobin in man and lungfish. Since the two lines are presumed to have diverged in Devonian times, some four hundred million years ago, the line leading to man has undergone profound physiological and morphological changes, while the modern lungfish is still very close in terms of its morphology and physiology to the ancient fishes. The line leading to man has supposedly undergone three fundamental transformations, the amphibian, the amniotic, and the mammalian. During the course of these presumed transformations the cardiovascular system has undergone enormous and dramatic changes. The heart has changed from a simple tube-like organ to a four-chambered efficient pump. The gills and branchial arteries have been replaced by lungs and the pulmonary circulation. The system of oxygenation has been utterly transformed. At the same time, the red blood cells themselves have become completely different. From the large round red cells of diameter of approximately 20μ ($\mu = 10^{-6}$ meter) typical biochemically of relatively unspecialized eucaryotic cells possessing nucleus, mitochondria etc, they have changed into small plate-like structures of diameter 7μ without nucleus or mitochondria and containing very much more haemoglobin per unit volume.

While this dramatic series of morphological, physiological, histological and biochemical changes were supposedly occurring along the lineage leading to the mammals, the morphological, physiological, and histological organization of the cardiovascular system of the line leading to lungfish must have remained virtually unchanged.

It is very difficult to understand why a protein functioning in the basically unchanging physiological environment of the lungfish's red cell should have undergone precisely the same number of beneficial mutations as a related protein evolving in a line subject to such global adaptational changes. While selection at the morphological and molecular level may be relatively unlinked, it is surely inconceivable that they could be *absolutely* unrelated. All the biology of an organism, all its anatomical features, its physiological and metabolic functions are ultimately reducible to its constituent proteins. Because organisms are systematic wholes in which every component more or less interacts with every other component, because all the functional components of living things are all ultimately made up of proteins, then inevitably every physiological or structural change is bound to impinge on the functionality of proteins. Proteins cannot be isolated from the environment in which they function.

Unfortunately, the case of lungfish haemoglobin is not unique. The opossum is another classic living fossil, virtually unchanged morphologically from its ancient ancestors of the late Cretaceous period nearly one hundred million years ago. But when opossum haemoglobin is compared with the haemoglobins of other mammals it is in no way primitive with respect to other mammalian species. In fact, rather the reverse; if anything, opossum haemoglobin is actually slightly further away from presumed common ancestors of mammals such as fish and amphibia than other mammalian species. So this mammalian species, a living fossil, apparently unchanged morphologically for nearly a hundred million years, a species which predates the entire adaptive radiation of the placental orders, has a haemoglobin as far removed from presumed mammalian ancestors as any of the recently evolved mammalian types!

Of course, the implausibility of selectionist explanations do not stop with the haemoglobins of a few living fossils. As in the case of uniform drift it is the sheer universality of the phenomenon – the necessity to believe that since their common divergence every single family of homologous proteins have suffered the same number of adaptive substitutions over the same period of time in *all* phylogenetic lines – which fatally weakens selectionist explanations.

Perhaps one of the most difficult problems in this whole area is trying to provide an explanation of how a uniform rate of evolution could have occurred in amino acid sequences which apparently perform no function other than acting as spacer sequences linking together the functional regions of a protein. A classic example of this are two short amino acid sequences which are snipped out of the protein fibrinogen after it is activated during blood coagulation. These are known as fibrinopeptides A and B. As far as is known, neither of these two short peptides have any biological function, yet their percent sequence divergence in different mammalian groups conforms to the same ordered pattern as is found in all other proteins, ie, the fibrinopeptides in all the members of any group are equally isolated from all the fibrinopeptide sequences found outside their group. If we are to explain this in terms of evolution we must again assume that an equal degree of fibrinopeptide sequential change has occurred in all the diverse lines of a particular group since their common divergence.

If such sequences really are under no selective constraints then drift is the only agent that could have been responsible for the pattern

of interspecies differences. Neutral sequences are by definition outside the surveillance of natural selection but this leads to a serious dilemma. As we have seen above, there is no conceivable way in which a uniform rate of drift could have occurred in organisms as diverse as mouse and man and yet the fibrinopeptides in rodents are isolated to exactly the same degree as those in primates. Drift seems to be excluded.

But selectionist explanations seem to lead to absurd conclusions. Because the spacer sequences such as the fibrinopeptides exhibit the highest interspecies divergence of all proteins, if this is to be accounted for on purely selectionist grounds it is necessary to propose that they must have suffered adaptive changes very much more often than proteins such as the haemoglobins or the cytochromes. In other words, they must have been under the intense scrutiny of natural selection. Not only must such sequences have suffered more adaptive changes than other proteins but in addition, these substitutions must have occurred regularly.

The difficulties associated with attempting to explain how a family of homologous proteins could have evolved at constant rates has created chaos in evolutionary thought. The evolutionary community has divided into two camps – those still adhering to the selectionist position, and those rejecting it in favour of the neutralist. The devastating aspect of this controversy is that neither side can adequately account for the constancy of the rate of molecular evolution, yet each side fatally weakens the other. The selectionists wound the neutralists' position by pointing to the disparity in the rates of mutation per unit time, while the neutralists destroy the selectionist position by showing how ludicrous it is to believe that selection would have caused equal rates of divergence in "junk" proteins or along phylogenetic lines so dissimilar as those of man and carp. Both sides win valid points, but in the process the credibility of the molecular clock hypothesis is severely strained and with it the whole paradigm of evolution itself is endangered.

There is simply no way of explaining how a uniform rate of evolution could have occurred in any family of homologous proteins by either chance or selection; and, even if we could advance an explanation for one particular protein family, we would still be left with the mystifying problem of explaining why other protein families should have evolved at different rates. The more deeply the problem

is examined the less it appears amenable to solution in terms of chance and selection.

Despite the fact that no convincing explanation of how random evolutionary processes could have resulted in such an ordered pattern of diversity, the idea of uniform rates of evolution is presented in the literature as if it were an empirical discovery. The hold of the evolutionary paradigm is so powerful that an idea which is more like a principle of medieval astrology than a serious twentieth-century scientific theory has become a reality for evolutionary biologists.

Here is, perhaps, the most dramatic example of the principle that wherever we find significant empirical discontinuities in nature we invariably face great, if not insurmountable, conceptual problems in envisaging how the gaps could have been bridged in terms of gradual random processes. We saw this in the fossil record, we saw it in the case of the feather, in the case of the avian lung and in the case of the wing of the bat. We saw it again in the case of the origin of life and we see it here in this new area of comparative biochemistry.

What has been revealed as a result of the sequential comparisons of homologous proteins is an order as emphatic as that of the periodic table. Yet in the face of this extraordinary discovery the biological community seems content to offer explanations which are no more than apologetic tautologies.

NOTES

1. Dayhoff, M.D. (1972) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Silver Spring, Maryland, vol 5, Matrix 1, pD–8.
2. ibid. Matrix 1, pD–8.
3. ibid, Matrix 1, pD–8.
4. ibid, Matrix 7, pD–52.
5. ibid, Matrix 1, pD–8; and Matrix 12, pD–88.
6. ibid, Matrix 10, pD–56; and Matrix 12, pD–88.
7. Woese, C.R. and Fox, G.E. (1977) "Phylogenetic Structure of the Prokaryotic Domain: The Primary Kingdoms", *Proc. Natl. Acad. Sci. U.S.*, 74: 5088–90.
8. Ferguson, A. (1980) *Biochemical Systematics and Evolution*, Blackie, Glasgow, p156–57.
9. Prager, E.M., Fowler, D.P. and Wilson, A.C. (1976) "Rates of Evolution in Conifers (Pinaceae)", *Evolution* 30: 637–49.
10. Zukerkandl, E. (1965) "The Evolution of Haemoglobin", *Scientific American*, 213(5): 110–18, see p111.
11. Darwin, C. (1872) *The Origin of Species*, 6th ed (1962) Collier Books, New York, p137.

12. Romer, A.S. (1966) *Vertebrate Paleontology*, 3rd ed, University of Chicago Press, Chicago, p15.
13. ibid, p204.
14. ibid, p78.
15. ibid, p208.
16. Drake, J.W. (1970) *The Molecular Basis of Mutation*, Holden-Day Inc. San Francisco, Table 5–3. Sager, R. and Ryan, F.J. (1961) *Cell Heredity*, John Wiley and Sons, Inc, Table 2.3, p55. Cavalli-Sforza, L.L. and Bodmer, W.F. (1971) *The Genetics of Human Populations*, Freedman and Co, San Francisco, p108.
17. Demerec, M. and Kaufmann, B.D. (1961) *Drosophila Guide*, 7th ed, Carnegie Institute of Washington, p3.
18. Imms, A.D. (1957) *A General Textbook of Entomology*, Methuen and Co Ltd, London, p439.
19. Brues, C.T. (1951) "Insects in Amber", Scientific American, 185(5), pp56–61.
20. Thomson, J.A. (1934) *Biology for Everyman*, 2vols, J.M. Dent and Sons Ltd, vol2, p1241.
21. Ewens, W.J. (1973) "Comments on Dr Kimura's Paper", Genetics Supplement, *Genetics*, 73: 36–38, p36.
22. Drake, op cit.
23. ibid.
24. Cavalli-Sforza, op cit.
25. ibid.
26. Wilson, A.C. Carlson, S.S. and White, T.J. (1977), "Biochemical Evolution", *Ann. Rev. Biochem.*, 46: 573 · 639, see p611.
27. ibid.