J Mol Evol (1987) 26:34–46

Journal of Molecular Evolution

On the Molecular Evolutionary Clock

Emile Zuckerkandl

Linus Pauling Institute of Science and Medicine, 440 Page Mill Road, Palo Alto, California 94306, USA

Summary. The conceptual framework surrounding the origin of the molecular evolutionary clock and circumstances of this origin are described. In regard to the quest for the best available molecular clocks, a return to protein clocks is conditionally recommended. On the basis of recent data and certain considerations, it is pointed out that the realm of neutrality in evolution is probably less extensive than is now commonly thought, in the three distinct senses of the term neutrality—neutrality as nonfunctionality of mutations, neutrality as equifunctionality of mutations, and neutrality as a mode of fixation of mutations. The possibility is raised that complex sets of interacting components forming a system that is bounded with respect to its environment may quite generally display an intrinsic trend to a quasi-clockwise evolutionary behavior.

Key words: Molecular evolutionary clock — Protein clocks — Neutrality — Natural selection — Gene regulatory clocks

Introduction

Twenty-five years ago, when the molecular evolutionary clock hypothesis was put forward, one might have thought that within the next quarter of a century the basic issues surrounding the molecular clock hypothesis would be resolved. This is far from being the case. After a historical review of the concept, some comments will be made here on three issues relating to the clock: 1) Evolutionary clocks have been explored for different compartments of genomes. Where should one look for the best clock? 2) It is now widely considered that most substitutions, even in coding nucleotides, are functionally neutral. Is this legitimate? 3) Are there other biological evolutionary clocks and, if so, is there a common foundation for all biological evolutionary clocks?

Offprint requests to: E. Zuckerkandl

The Clock Postulate in Historical and Conceptual Perspective

The work of the initiators of the study of molecular evolution as a field had one basic shortcoming. At least at the time of their early writings, neither Ernest Baldwin (1936) nor Marcel Florkin (1944) were in a position to perceive the unique value for molecular evolution of the proteins and the nucleic acids. Only with this realization, pioneered by Christian Anfinsen (1959), did the study of molecular evolution find its center of gravity.

Linus Pauling, of course, was aware of how basic proteins were. When I joined Linus Pauling at the California Institute of Technology in 1959, he suggested that I work on the evolution of primate hemoglobin, in collaboration with Richard T. Jones, then one of his graduate students. Soon, the comparison of two-dimensional patterns of peptides resulting from the tryptic digestion of hemoglobins (a technique then called fingerprinting) permitted the inference that hemoglobins differed in primary structure at least roughly as a function of the phylogenetic distance between the organisms from which they had been obtained (Zuckerkandl et al. 1960). According to the evidence derived from tryptic digests of hemoglobin, humans, gorillas, and chimpanzees were more closely related to each other (there was no detectable difference between them) than either one was to the orangutan. This inference and some additional ones drawn from the same work suggested to me that there might be a proportionality

between the time elapsed since the period of the last common ancestor of two contemporary organisms on the one hand, and on the other hand the number of sequence differences between homologous polypeptide chains from these organisms as well as between their corresponding genes. I made this tentative assumption of a molecular evolutionary clock (first so designated in 1965, Zuckerkandl and Pauling 1965b) in late 1960 or in 1961.

In the first paper dealing with the molecular evolutionary clock, an underlying assumption was that the evolutionary rates of amino acid replacements in hemoglobin chains controlled on the one hand in different organisms by presumably corresponding gene duplicates and on the other hand within the same organism by distinct gene duplicates were *both* roughly—perhaps very roughly—proportional to time (Zuckerkandl and Pauling 1962). The two types of homology involved here, which we called duplication-independent and duplication-dependent homology (Zuckerkandl and Pauling 1965b) were later rechristened orthology and paralogy, respectively (Fitch 1970).

The paleontological record was used for the approximate calibration of absolute evolutionary time. It was estimated that about 7 million years (Myr) had elapsed per amino acid replacement in hemoglobin chains since the time of the last common chain ancestor. The same applied to the ancestor of organisms when corresponding polypeptide chains from different organisms were compared. On this basis it was estimated that the common ancestor of man and gorilla dated back perhaps 11 Myr. This figure would have been 7 million, had we not believed that there probably were two amino acid replacements that distinguished the α-globin chain of gorilla from that of humans (Zuckerkandl and Schroeder 1961), instead of one difference, as is now known to be the case (Goodman et al. 1983). As was emphasized at the time, such dating of ancestral genes and organisms could only be very rough, but the corrected figure for the ancestor of man and gorilla happens to be not far different from figures considered as probable at the present time (e.g., Chang and Slightom 1984; Sibley and Ahlquist 1984).

At the time the study of molecular evolution came into its own, the first important step was to extend to sequences of informational macromolecules the biological concept of homology, namely to pursue the comparative study of structures in a way such as to be able to conclude, if the conclusion applied, that the structures very likely descended from a common ancestral structure. The extension to informational macromolecules of the concept of homology was a precondition for, though quite distinct from, the molecular clock concept. In June 1960 it appeared to this author that not only corresponding hemoglobin polypeptide chains as found in different organisms and controlled by presumably identical loci, but also different hemoglobin chains found within the same organism and controlled by distinct loci were likely to be traceable to common ancestral genes which, in the latter case, had undergone duplication. This intuition was verified by comparing some preliminary sequence data on the human hemoglobin α- and β-chains. As a guest in Walter Schroeder's laboratory at the California Institute of Technology, I had access to the nascent sequence data of Walter Schroeder and his collaborators on the human α-globin chain as well as to 30 amino acids, including the N-terminal one, from the human β-globin chain put in sequence by Gerhard Braunitzer, a sequence fragment that Max Delbrück had just brought back to Caltech from a visit that he had paid to Braunitzer in Munich. Some of Schroeder's sequenced, but as yet unordered, tryptic peptides could indeed be fitted onto Braunitzer's sequence, provided a single-residue insertion or deletion was assumed to have occurred in the ancestry of one of the chains. This result strongly suggested homology between the hemoglobin α- and β-chains, and it was natural to think of the other human hemoglobin chains then known, the δ- and γ-chains, as homologous too, and of their last common molecular ancestor as dating back further in time the larger the number of sequence differences between the homologous polypeptide chains was found to be (see Zuckerkandl 1972). During this period, Vernon Ingram developed the same molecular evolutionary concept and on account of special circumstances was the first to publish it (Ingram 1961). In fact, however, neither Ingram nor this author were first to think of different globin genes as being homologous and having arisen as a consequence of gene duplication. This hypothesis was first advanced by Itano (1957), whose prophetic paper came to my attention after I had unwittingly followed in his footsteps. Though Itano was not in 1957 in a position to verify his idea, there is no doubt about his conceptual priority (see also Anfinsen 1959). However, in addition to homology between different hemoglobin polypeptide chains, I assumed that numbers of differences in sequence between homologous polypeptide chains could be expressed as the approximate evolutionary time since divergence.

This represented a conceptual jump. Observations on increasing differences between homologous proteins, including hemoglobins, from increasingly distantly related contemporary organisms had been made before. Reichert and Brown, as early as 1909, discovered that the protein moiety of hemoglobin was structurally different in different species, as

demonstrated by different axial ratios of hemoglobin crystals. Their important finding was that members of a genus and sometimes of a family generally shared crystallographic characteristics and generally did not share them with different genera or families. Substantive early contributions to a macromolecular taxonomy were subsequently made, notably by Boyden (1958), on the basis of immunological cross reactions, an approach dating back to Nuttall (1904), and that came of age later with the work of Goodman (1962, 1963) and eventually tied in with the molecular clock concept and its development through the accomplishments of Sarich and Wilson (1967a,b, 1973). What Reichert and Brown as well as Nuttall and his followers discovered is that proteins, even though their structure is microanatomical, behave like other anatomical characters: an analysis of graded differences between comparable proteins, i.e., proteins that show important chemical and functional similarities, more recently demonstrated to express homology, led to taxonomic groupings that are in general in accord with those drawn from higher levels of biological organization. Reichert and Brown themselves considered their comparative crystallographic analysis of hemoglobins as "a valuable adjunct to morphological data." They pointed out that it should be possible to use crystallographic characters of proteins to test phylogenetic relationships (at and below the family level).

Whether degree of observed differences between homologous structures measures time is quite another matter. Before the molecular clock concept was formulated and applied, not only had differences in macromolecular sequences not been correlated with evolutionary time, but also amounts of differences at any other levels, such as in morphological features, had not been surmised to be proportional to evolutionary time, and for a good reason: such a surmise seemed impossible to substantiate. In regard to morphological features, large changes in evolutionary rates even along one and the same lineage seemed to exclude anything remotely resembling a clocklike behavior in incrementally growing differences (Simpson 1953; Mayr 1982). Within certain limits, an evolutionary clocklike behavior may in fact be manifest at the level of certain morphological features (Turner 1986). (We shall return to this matter.) Twenty-five years ago, however, the diversity of rates in morphological evolution was the strongest argument, in the minds of biologists, against seriously considering the possibility that graded differences between informational macromoleclules might be proportional to evolutionary time. It did not occur to observers that the *structural differences at the anatomical level* might for the most part be traceable to *regulatory differences at the genetic level* (Zuckerkandl 1964a, 1968; Zuckerkandl and Pauling 1965b; King and Wilson 1975; Wilson 1975; Cherry et al. 1978), and that this left the *structural differences at the genetic level,* in the sense of differences between structural genes, *free* (though not obligated) to follow a clocklike behavior, without contradicting the observed absence of a clocklike behavior in morphological characters. To generate a rounded out picture, all that it was necessary further to assume was that, at the *genetic* level, the *regulatory* evolutionary changes, like, at their own level, the morphological changes, were not clocklike—an assumption that, now, ironically, has in turn begun to be questioned (Kettler et al. 1986). Indeed, at a time when the existence of anything that would deserve to be referred to as a molecular evolutionary clock continues to be challenged by certain analyses, evolutionary clocks seem, on the other hand, to pop up in quite unexpected niches. One cannot deny that the plot that runs through the clock story is imaginative.

At any rate, the conceptual jump that is required to pass on to the clock concept from the findings of Reichert and Brown and of Nuttall or from measurements of distances between primary polypeptide structures elicited strong adverse reactions. Kimura and Ohta (1974) wrote that the concept of the molecular clock was already implicit in the work of Ingram (1961) that I referred to. Yet this is not so for the reason given. Furthermore, Ingram was among those opposed to the idea of a molecular clock and rejected it, notably at a meeting held in Bruges in 1964. The first to support the clock hypothesis probably were in the United States Tom Jukes (1963), Allan Wilson and Vincent Sarich (Sarich and Wilson 1967a,b, 1973), and in Japan Motoo Kimura (1968). It should also be recalled that in 1964 Doolittle and Blomback examined the number of differences between fibrinopeptides as a function of the time of their last common molecular ancestor and found a correlation. Kimura (1968) recognized in the clock concept a major basis of the "neutral theory" and demonstrated the apparent span of its applicability (see also Derancourt et al. 1967). Laird et al. (1969) were perhaps the first to point to limitations of the clock. Sarich and Wilson undertook a series of studies that furnished the first spectacular applications of the clock principle to particular evolutionary problems and that greatly strengthened the methodology and theory of the clock. There have since been a number of important and sophisticated contributions to the clock concept and to its applications of at times diametrically opposed tendencies. It is not my purpose to review these contributions. Today the matter is as controversial as ever, and the issues are complex.

In 1962 Linus Pauling communicated to me his

idea of deducing ancestral polypeptide sequences from a comparison between contemporary homologous polypeptide sequences (Pauling and Zuckerkandl 1963). My first, though fleeting, reaction was to resist the concept: I was unduly mesmerized by the really rather powerless ghost of evolutionary convergence. I later concluded that evolutionary convergence *could* not decisively obscure the results of evolutionary divergence (unpublished). Soon, contemporary hemoglobin chains were individually compared with tentatively reconstructed ancestral sequences. The procedure contained an element of circularity, but the circularity was only partial, and the results encouraged the assumption that orthologous and paralogous genes evolved at comparable and nearly constant rates (Derancourt et al. 1967). The parsimony principle, which had independently been formulated by Cavalli-Sforza and Edwards (1964), was here for the first time applied to macromolecular sequences.[1] The amount of historical information locked into informational macromolecules (or "semantides," a handy term that was not adopted), the unique position in this regard of semantides among all molecules present in living systems, and their potential usefulness for the construction of phylogentic trees were apparently first emphasized by Zuckerkandl and Pauling (1964, 1965a), though Anfinsen (1959) indeed had already pointed to the importance for the study of molecular evolution of the proteins as reflections of the genes. The earliest examples of the building of molecular phylogenetic trees related to hemoglobin. They were to be found in Pauling and Zuckerkandl (1963), Jukes (1963), Zuckerkandl (1964b), and Zuckerkandl and Pauling (1965b). The presentation of an algorithm for the method and the demonstration of the power of the general approach as well as of that of the algorithm were the feat of Fitch and Margoliash (1967). They built a comprehensive phylogenetic tree on the basis of just one single type of protein, cytochrome c. This tree kept the promise of the approach in that it largely coincided with trees derived from morphology.

There were two obvious possibilities regarding rates of change in informational macromolecules: either their structure followed morphological evolution in regard to rate, or the structure of informational macromolecules was independent of morphological evolution. In the early 1960s there was no good evidence for either fundamental relationship. On theoretical grounds, both could be defended. During a time in 1964, I thought that it would turn out that in so-called living fossils a majority of structural genes have remained close to their ancestral forms (Zuckerkandl 1964b, 1965); they would more strongly diverge along lines of descent of organisms whose anatomy and physiology differed considerably from those of their ancestors. This was a partial, not a complete, reversal of position in regard to the molecular evolutionary clock, which was simultaneously affirmed (Zuckerkandl 1965). It was a concept of the linkage between rates of morphological and macromolecular evolution involving *a certain proportion* of polypeptide chains. Why, indeed, should rates of macromolecular evolution be completely disconnected from rates of morphological evolution? I have already said here why they *could* be disconnected and had referred to this matter for the first time in 1963. Nevertheless, in 1964, I held a view at variance with the earlier—and later—concept of such a disconnectedness. I returned to my first position later in the same year (Zuckerkandl and Pauling 1965b). Why this return to the assumption of an independence of rates at the organismal and macromolecular levels? The data that I had been able newly to bring together related to a few eutherian mammals and did not provide a test. Yet I had understood that presumably only a small fraction of all evolutionarily effective amino acid replacements in proteins were likely to be functionally innovative.[2] The majority of the replacements in hemoglobin chains seemed to fluctuate back and forth between certain subsets of amino acids at certain relatively highly variable molecular sites. They did this perhaps for functional reasons, even if the result resembled a random process. As Sneath (1980) stated: "If selection pressures fluctuate rapidly in direction, this will produce, by selection, changes that appear random." It was possible that many of these frequent substitutions had almost no functional effect and could therefore be called neutral, "if one dared pronounce this word" (Zuckerkandl 1968). In 1965 we did dare pronounce it (Zuckerkandl and Pauling 1965b), though neutrality was then about as popular in certain quarters as vitamin C is now in certain others, and proposed that "along lines of descent marked by high evolutionary sta-

[1] "It seems rational to distribute presumed ancestral residues over the molecular phylogenetic tree in a way such that the total number of evolutionarily effective mutations in the section of the tree that is being considered be reduced to a minimum in accordance with the transitions allowed by the genetic code" (Zuckerkandl, 1964b)

[2] In 1955 Fred Sanger and his associates (Brown et al. 1955) had shown that only three amino acid residues, all located in one region of the molecule, varied between cattle, pig, and sheep insulin. These residues then were not "particularly critical ones," as Anfinsen (1959) put it. Indeed, these structural variants all were able to function as insulin when injected into humans. Allan Wilson (personal communication) concluded from this observation as early as 1955 that most amino acid substitutions in proteins probably would turn out to be functionally unimportant. The question is: does "unimportant" mean "equifunctional and fixed by random drift" or "only slightly functionally divergent but probably fixed by selection?" In 1965 I believed the second

bility, the shuttle motion between functionally similar amino acids will also occur. The changes in amino acid sequence will, however, be limited almost exclusively to the functionally nearly neutral changes." This stance failed to foreshadow neutral mutation theory, in that we then contemplated the spread in populations of these nearly neutral mutations by selection rather than by drift.[3]

Whatever the mechanism of their spread in populations, apparently the frequent nearly neutral or moderately selective amino acid replacements deserved to be called the conservative ones, since they essentially conserved the protein's function regardless of the chemical and steric nature of the amino acid's side chain. Even though all change accompanying evolution is included in evolution, the important part of evolution—and one might be so bold as to say biological evolution, period—resides in functional evolution and in all that is connected with it. In this sense, important evolutionary changes had to be linked to important functional changes in proteins or to functionally significant changes in the regulation of their rate, time, or place of synthesis (or, as is apparent to me today, to changes in the organization of gene interaction networks). Whether the modification was in protein structure or regulation, important evolutionary changes would be brought about by a small minority of sequence changes in coding or noncoding DNA. Thus morphological changes, when functionally significant (and those that biologists think of in this connection usually are), would most likely be disconnected from the majority of fixed mutations and therefore from the molecular evolutionary clock.

Kimura (1969) in turn considered proteins from living fossils as excellent test objects for deciding between the two possibilities, namely that the rate of sequence change in proteins was linked to the rate of morphological evolution in organisms or that these two rates were independent, as they had already been shown to be by Sarich and Wilson (1967a). This independence was required by the neutrality theory. Subsequently, Stenzel (1974) showed that in the line of descent leading to the opossum, a well-known living fossil, the rate of sequence change in the hemoglobin α-chain was not lower (it was actually higher) than in the α-chains leading to a number of other vertebrates. Both the "clock" and the neutrality theory passed this test, though not with a brilliant grade (because of the excess in amino acid replacement rate), but a single protein could not demonstrate that the disconnectedness between rates of structural evolution of informational macromolecules and rates of morphological evolution was general as, since 1965, I thought likely. It remained for Allan Wilson and his associates to bring to bear on this point a large number of experimental data gathered with imagination and rigor (Wallace et al. 1971, 1973; Wilson et al. 1974; Maxson and Wilson 1975; and references cited above). The data went far toward vindicating the hypothesis of the generality of the molecular clock in genes, irrespective of rates of evolution at higher levels of biological integration.

It became clear early on that the rate at which the clock ticked differed not only between different informational macromolecules, notably between different sectors of DNA, but also between different parts of such molecules, in fact from molecular site to molecular site. The very first sequence comparisons between several hemoglobin chains from different species (1960; see Zuckerkandl 1972) showed that in some regions the hemoglobin molecule was evolutionarily much more variable than in others. Emmanuel Margoliash (1963) made the same observation on cytochrome c. It was not long before it was discovered, on the basis of the work of Max Perutz and John Kendrew, that the broadest generalization with respect to these differences was that the outside of globular proteins was evolutionarily more variable than their inside (namely than the amino acid residues to which water molecules had no access) (Perutz et al. 1965; Zuckerkandl and Pauling 1965b; Derancourt et al. 1967; Vogel and Zuckerkandl 1971b). At the level of DNA, there is likely to be a continuous variation in the rate of local sequence changes. At one extreme there may be no evolutionary sequence change at all over hundreds of millions of years at the first and second positions of codons. On the other hand the rate of fixation may be that presumed to correspond to the rate of fixation of neutral mutations. There may be sequence changes at rates even more rapid than the accepted neutral mutation rate when these changes

[3] Later, in 1967–1968 (see Zuckerkandl 1968, p. 271 ff.), independently of Kimura (1968) and King and Jukes (1969), I seriously considered the possibility that random genetic drift of neutral mutations played an important role in molecular evolution. In particular, I observed that apparently, during the evolution of hemoglobins and of cytochrome c, the fractional representation of amino acids tended to increase when it had been low and to decrease when it had been high, seemingly to approach "codon equilibrium," i.e., a representation of the different amino acids in approximate proportion to the numbers of codons that code for each of them. I presented these results with their neutralist conclusion at a joint meeting of the Belgian and French Biochemical Society in Brussels in 1968. This conclusion was however questioned by my colleague Helmut Vogel (Vogel and Zuckerkandl 1971a), who discovered that the results were also compatible with selection. Moreover he showed that concepts homologous to those used in thermodynamics could successfully be applied to a treatment of molecular evolution, and that on this basis a protein, like cytochrome c, displayed during evolution a functional constancy combined with a virtually complete structural adaptation to the function that left little space for functional neutrality

spread through genomes through special mechanisms (selfish DNA, Doolittle and Sapienza 1980; Orgel and Crick 1980; molecular drive, Dover 1982).

What type of sequences might provide the best molecular clock? As yet we cannot say, but there are indications now as to where to look for it.

Return to a Protein Clock?

Several observations suggest that the rate of compositional, and therefore sequence, change in single copy nuclear DNA (scnDNA) has not remained constant in evolution. Thermal stability measurements of hybridized "unique" sequences indicate that predominantly noncoding sequences evolve at different rates along different lines of descent (Britten 1986). Vawter and Brown (1986) also conclude from their study of rates of evolution in mitochondrial DNAs, which they compare to those of scnDNA, that the rate of evolutionary change in scnDNA has not remained constant. In this regard one should also consider the differentiation of chromosomal bands as revealed by Giemsa and other staining (Holmquist 1987), the appearance in the vertebrate line of isochores (DNA fragments of similar base composition) richer in GC than isochores known from the descendents of earlier vertebrate forms (Bernardi et al. 1985), and the relatively rapid change in GC content as a function of ambient temperature, as revealed by the comparison between closely related poikilotheric organisms adapted to different temperature ranges (Bernardi and Bernardi 1986). There is a possibility that a slowing in evolutionary rates can *in part* be accounted for by increases in functional density (Zuckerkandl 1986) of noncoding scnDNA in some branches of the vertebrate line.

Bernardi and Bernardi (1986) have shown that changes in isochore base composition affect all nucleotide positions, coding and noncoding; but that the first and second codon positions—especially the second—respond to regional compositional change at distinctly slower rates than the third codon position does. The difference between the first and second positions is by a factor of 0.75; that between the second and third postions by a factor of 2.7. The best molecular clocks may well be those that depend on nucleotide positions least affected by regional or genomic evolutionary changes in GC content, whatever the factors be that determine these changes. On this basis, the best clocks should be obtained with sufficiently large sets of second codon positions. Next best would be protein clocks. In a number of cases it has been found that clocks based on third codon positions are indeed poor (Perler and Efstratiadis 1980; Li et al. 1985; Wu and Li 1985; Britten 1986; Yokoyama and Gojobori 1987), but from a set of data that span a very large sector of the phylogenetic tree, including bacteria, Allan Wilson (personal communication) reaches the opposite conclusion.

It had been thought, and this was a logical stance, that the larger the number of sequences that are used in genetic distance measurements, the more reliable we may expect the molecular clock measurements to be. This has been the philosophy underlying the use of total populations of unique sequences in hybridization experiments (Sibley and Ahlquist 1984, 1986). The approach has contributed significantly to the knowledge of evolution and, thanks to the effectiveness of the method, has yielded a large volume of results. Yet, in regard to reliability of the clock and to taxonomic span over which a certain ticking rate applies, the best possible clock is probably based on a more selective approach to nucleotide choice. From this choice, the bulk of noncoding sequences should be excluded if further research warrants it. A study needs to be undertaken to establish the extent to which the molecular clock is likely to be affected by genomic variations in GC content (local and general). There are other processes, those subsumed under the name molecular drive, that are expected to interfere with a noncoding sequence clock (Dover, personal communication), though not indisputably with a "unique" noncoding sequence clock.

A number of considerations might, and in some cases (respect of orthologous relationships!) must, serve as guidelines for choosing the members of sets of proteins collectively best suited for testing the molecular clock hypothesis and for carrying out molecular evolutionary studies. Hopefully such sets will be established and indeed put to the test. One guideline to be considered involves the interference with the molecular clock by events of gene conversions (Zimmer et al. 1980). Even though over a large number of genes and a long evolutionary time the effects of such events on the apparent tick rate of the clock may average out, it is advisable to eliminate from the elected set of structural genes those that are likely to have been the object of gene conversion in relatively recent evolutionary time. All genes probably have undergone such processes at one time or another of their history, because all genes, at one time or another, probably have existed in the form of closely linked duplicates.[4]

[4] This follows from (1) the probable filiation of most proteins or, rather, most polypeptide domains back to a few original polypeptide domains (Zuckerkandl 1975); and (2) the constant origination and loss of gene duplicates, expressed in the fact that most surviving gene duplicates controlling the structure of polypeptide chains of very similar function are of relatively quite recent evolutionary origin (Petit and Zuckerkandl 1976). Except in the case of duplication of chromosomes or of fractions of

A further guideline for molecular clock studies is to use only those homologous proteins whose functional density and, more precisely, weighted functional density have remained constant over the phylogenetic sector under consideration (Zuckerkandl 1976a, 1986). Both evolutionary increases and decreases in functional density may occur (Zuckerkandl 1976b) and ought to be taken into account in tests of, or studies based on, the molecular clock. To do so would require a very detailed knowledge of the proteins. Yet major apparent irregularities in the rate of evolution of a given protein may well be attributable to a neglect of this factor.

On the Neutrality Issue

Neutrality as a modality of fixation of mutations, neutrality as functional equivalence, and neutrality as nonfunctionality are really three neutralities. They have interesting relationships, but should be carefully distinguished, lest one end up with the problems that always arise when a trinity is considered as one.

Functional constraints in genomes would eliminate neutrality as a mode of fixation if they were universally strong. This they are not. Hence functional constraints cannot eliminate neutrality as a mode of fixation. However, they can reduce the domain of applicability of this mode.

Despite great efforts at clarification (Ayala 1974; Kimura 1983), the respective contributions of selection and of neutrality as a mode of fixation are still in dispute, and this applies to noncoding as well as to coding sequences. In the case of the coding sequences, it is at least known that negative selection eliminates the majority of amino acid replacements in proteins. This, however, can account for the fate of only a few percent of base substitutions in genomes of higher organisms (see Walker 1983). Receptor DNA sequences to which regulatory proteins bind with a high degree of specificity add to this percentage, but, considering their shortness, surely only modestly. Whether or not a significant fraction of structural changes in DNA are subject to selection, negative or positive, then hinges upon the question whether or not the compositional constraints, observed by Bernardi and colleagues to be pervasive in genomes, are in large measure attributable to selective constraints. The hypothesis that such selective constraints are widespread in genomes is supported by the conservation in at least some types of satellite DNA of certain sequence features (Brutlag 1980), probably in part linked to proper DNA packaging (Strauss and Varshavsky 1984). The conservation within and between genomes of specific sequence features such as required for proper packaging, in the face of variations of other sequence features, is difficult to explain on any basis other than selection. Within the part of the noncoding sequences that are thought to be neutral in the sense of functionless, indications of functionality are creeping in under the eyes of the skeptics. Such developments involve proportions of DNA that go from rather modest in the case of "scaffold attached regions" (Gasser and Laemmli 1986), through respectable in the case of the evolutionary conservation over different families of rodents of a cell-type-specific nuclear localization of centromeric heterochromatin (Manuelidis 1984), to the more sweeping shifts in GC content in genomes as a function of ambient and internal temperature (Bernardi and Bernardi 1986). As to the genetic load argument against the idea of selection acting upon noncoding sequences, it may be weaker than its tenants think (Zuckerkandl 1986).

In the case of the coding sequences, where nonfunctionality of coding nucleotides is not assumed, but is replaced by the postulate of equifunctionality (Kimura 1983),[5] the relative contribution of mutations fixed by random drift and of positively selected mutations is also still in question. It becomes progressively clearer that even the most "ordinary" amino acid replacement may be functionally nonequivalent, at least potentially, and therefore may well be selected for or against, at least at times. Available evidence is not proof that this is so (Koehn et al. 1983), but is suggestive. For example, a recent comparison between neutral proteases of several species of *Bacillus* shows that certain amino acid replacements such as serine for threonine, alanine for serine, or arginine for lysine, which are among the most highly conservative replacements (under any definition of the term conservative), have a significant effect on the thermostability of the enzymes (Imanaka et al. 1986). Furthermore, for a number of enzyme loci, it has been shown that allele frequencies were highly significantly correlated with diverse environmental variables (Nevo et al. 1983; Philipp et al. 1985). A substitution that has no effect

chromosomes, gene duplicates originate in close linkage and are susceptible to gene conversion. Even when a gene has been duplicated by chromosome duplication it is probable that in addition it also has been duplicated at one time or other as a gene. Gene duplicates are conserved over long evolutionary periods only if before their elimination, mostly no doubt through their conversion into pseudogenes (originally called dormant genes, Zuckerkandl and Pauling 1962; Zuckerkandl 1972), they have managed to diverge sufficiently in function so as to be preserved (Petit and Zuckerkandl 1976)

[5] The distinction between nonfunctionality and equifunctionality of amino acid replacements, in relation to an apparent random process of fixation, was also made by Zuckerkandl et al. (1971)

on one function that a given protein molecule is able to carry out may be critical for another function of that protein, as has been shown for lactate dehydrogenase (Clarke et al. 1986; see also Wistow et al. 1987). In fish, certain isozymes controlled by allelic genes have been shown to be functionally nonequivalent in several respects, including swimming performance (DiMichele and Powers 1982; Powers et al. 1983; DiMichele et al. 1986). This does not demonstrate that selection of these alleles actually occurred. Yet the spectrum of the results just listed suggests that amino acid replacements in proteins may have significant functional effects more frequently than has been thought in recent years. It has now become fashionable to take for granted that most observed substitutions are functionally neutral, just as not too long ago similarly gratuitous selectionist statements were made without any evidence in the particular cases considered.

Though a functional effect produced by a mutation does not by itself decide about the mode of fixation or elimination of the mutation, except when the effect is very strong, it no doubt does alter the relative frequency at which a mutation will be fixed or eliminated by random drift. It is still possible that an important fraction of evolutionarily effective mutations at least in structural genes are much of the time selected for. This possibility received a serious analytical foundation as early as 1966 (Sneath 1966). Random drift and positive selection could indeed mingle in a systematic way (Zuckerkandl 1976b).[6]

The process of a selection/drift alternance may be illustrated by a succession of evolutionary events such as might occur in stem-loop structures of DNA or RNA. A first mutation that, at one base pair position, abolishes the normal base complementarity, may spread by random drift even though it is slightly deleterious. A second mutation elsewhere along the stem would be in a more ambiguous situation. It might still spread by a random effect, brought about for instance by a population bottleneck under genetic sufficiency (Zuckerkandl 1978a). There may be random drift of mutations that, in the presence of a fitter allele, would behave as frankly nonneutral, namely as deleterious, yet that are compatible with genetic sufficiency and spread in populations through random processes as long as there is no fitter allele to compete with these mutations. This probable looseness in the connection between functional impact and random drift leads to the necessity of keeping apart the notions of functional neutrality and neutrality as a process of fixation. In our example, genetic sufficiency becomes manifest in the absence of a competing population endowed with a "better" stem. Yet if the second or nth mutation spreads by drift, and reaches a momentary peak in the population, it is nevertheless likely that its effect will eventually be eliminated. This may happen by various pathways, such as a back mutation in the second mutational site or in the first, or as a compensatory mutation in the stem's other DNA strand at either site whereby base complementarity is reestablished, or as a mutation leading to base complementarity at a third site where such complementarity did not exist to start with. When a mutation reestablishes a base complementarity, the genetic sufficiency of the local situation of noncomplementarity will no longer insure the survival of the noncomplementary mutational state. Indeed, now that a competing stem structure is present in the population, the "fitness treadmill" will start rolling again, and the stem with more base complementarity will win out by *selection*.

Similar considerations should apply to secondary structures of proteins. One of the frequently exchanged amino acids in a loop or in a helix or in a β-structure can spread by drift, but a compensatory change in a spatially nearby site is then expected eventually to spread by selection (see Ohta 1973; Taylor 1986). *Drift may accumulate a "debt" that selection eventually repays.* Overall, the proportion of selected molecular change may be important.

Perutz (1983) agrees with Kimura (1979) in estimating that amino acid replacements in hemoglobin have in general no visible phenotypic effects and no obvious correlation with (external!) environmental conditions. The structural relationship between human, gorilla, and chimpanzee hemoglobins led me in 1968 to make a similar point regarding functional neutrality. However,

1) When a protein from two species differs by a number of amino acid replacements, yet its functional properties are "the same," as is the case according to Perutz (1983) of human and horse hemoglobins, it is likely that a proportion of the amino acid replacements have occurred in compensation for others, i.e., it is likely that, compared to the residues previously at the same positions, they were functionally nonequivalent. This hypothesis would be tested by the synthesis of a hemoglobin with amino acid residues as in human hemoglobin at one-

[6] This systematic aspect might weaken the objection of Kimura (1983) and Nei (personal communication), who correctly pointed out that in the selectionist approach to the molecular clock included in a paper published in 1976 (Zuckerkandl 1976b), I had failed to take into account the effects of changes in population size. In a selection/drift alternance, which, it seems to me, had already been implied in Ohta's (1973) model, the neutral clock component might furnish a backbone of rates sufficient to moderate to some extent the effects of changes in population size. At any rate, provided certain assumptions are modified, changes in population size are no longer universally considered to be an impediment to the possible existence of a clock of positive selection (Gillespie 1986)

half of the 42 sites, chosen at random, at which human and horse hemoglobins differ, and with amino acid residues as in horse hemoglobin at the other half of the sites. If there were no functional difference between this synthetic hemoglobin on the one hand and human and horse hemoglobin on the other, Kimura and Perutz's point would be considerably strengthened.[7]

2) Data referred to above underline the advisability of caution when a statement is made about "the same" function of two structurally diverged proteins. There may be differences in some parameters that have not been tested for and one cannot be confident that the untested for parameters will turn out to be functionally unimportant.

3) It may indeed be useful not to overlook the relationship between genetic sufficiency and the "fitness treadmill" (Zuckerkandl 1978a). An initial difference may have been conserved as genetically sufficient in the absence of a fitter allele in the population. Later a compensatory mutation may have been selected and now a back mutation at the original site would no longer be fitter, it would be less fit and may not be retained. In this fashion numerous amino acid differences could progressively develop between species in part by selection, and yet the proteins remain functionally similar.

4) Adaptive amino acid replacements may occur in response to changes in the *internal* environment, namely in direct or indirect response to sequence changes in other proteins (Zuckerkandl 1976b), notably in interacting proteins such as haptoglobin in the case of hemoglobin or in proteins competing for a factor or in noncoding sectors of nucleic acids that regulate the synthesis of these proteins.

Models that include a significant amount of selection should continue to be considered, even if it is not yet clear, or at any rate not generally accepted, how they could be compatible with the molecular clock in the presence of changes in effective population size.

The possibility of a selectionist basis for the molecular clock, in addition to its neutralist basis, does, it seems to me, deserve further exploration. Regarding a clock of *positive* selection, Hartl and Dykhuizen (1979) have adduced some experimental evidence for its occurrence, under conditions, it is true, of constant population size. Perler et al. (1980) argued for such a clock (see also Buonagurio et al. 1986). A statistical approach combined with the analysis of a model have been presented that support the concept of a clock of positive selection under certain circumstances where variations in effective population size average out (Gillespie 1986).

[7] Mas et al. (1986) interchanged the two protein domains of human and yeast phosphoglycerate kinase, whose amino acid compositions differ by 35%. They observed nearly normal activity in the hybrid enzymes. Domains indeed remain structurally coherent throughout evolution and amino acid residues at active sites are usually invariant. In the experiment outlined here, the differences in amino acid sequence would be sprinkled at random over the whole molecule

Gene Regulatory Evolutionary Clocks

Since the clocks that we are dealing with in evolution are not metronomical, but stochastic (Fitch 1976), we probably should not speak of clocks, but of quasi-clocks. As long as this is understood, it is convenient to refer to evolutionary clocks.

The question arises as to why the rates of evolutionarily effective sequence change in most of the noncoding unique sequence compartment of nuclear DNA are not faster than they are. The intrinsic potential for faster genetic change is present (e.g., Fitch and Atchley 1985a,b). Yet there may also be an intrinsic resistance to genetic change.

A general foundation for evolutionary clocks may be the following. The complexity of a system, in the sense of number of its interacting components, and the modes of interaction of these components combine in setting up a certain intrinsic stability of the system. This circumstance is particularly striking in the cases of homeostasis, where the system automatically corrects for a change. Complex systems must have originated through the workings of universal constants and through a succession of acts of selection. Once such a system is present there probably is, independently of the system's environment and therefore independently of any selection, a certain built-in resistance to the system's loss of identity. It is true that both kinds and rates of changes that complex systems undergo are often conditioned by factors in the systems' environment, but what is being emphasized here is the contribution to these kinds and rates made by the systems themselves. A range of values for built-in trends to change, which also express of course their inverse, the built-in resistance to change, should indeed, to a significant extent, be a property of the complex structures themselves. Environmentally determined rate changes would undoubtedly be manifest in biological systems, but would not be overriding within a certain range of environmental conditions. In altering the tick rate of clocks, they would, long term, play on the whole a secondary role. Not only in the disintegrating atomic nucleus, but also in biological systems, clocks would tend to be intrinsic clocks. The concept is not that all phenomena that display clocklike behavior are intrinsic clocks, but that all ultimately depend on intrinsic clocks. Selection, a type of environmental effect, may well not negate a clock over the long haul, provided that the frequency at which the environmental event has impact also

depends in part, as it must, on the structure of the system. This dependency is established via the frequency of opportunities for selection that the system provides. In these opportunities, generation time and population size probably both intervene (Kimura 1983), but not necessarily in a fashion that obliterates a quasi-clock. This should hold for systems at any level of biological organization.

This general concept could be applied, for instance, to the clocks of linguistic change, which really were the first "evolutionary" clocks to be conceived (Swadesh 1950). It could also apply to morphological clocks when such are found. Turner (1986) has shown that there is a dental clock, namely that 28 standard crown and root traits of human teeth vary over a period of at least the last 60,000 years as a function of time. Why does such a clock not apply to morphological change in general? Over certain periods of time, along certain lineages, anatomy *can* evolve at rather steady rates, as measured against rates of protein evolution (Selander and Johnson 1973; Avise 1974). Such correlations are radically contradicted when comparisons bear on different lines of descent, as has been observed to be the case within a wide variety of taxonomic groups, for instance frogs and placental mammals (Wilson 1975). In regard to Turner's dental clock the following points should be considered:

1) The heritable traits observed in "dentochronology" cannot help but be under the control of proteins.
2) Judging from normal rates of evolutionarily effective mutations, the time spans involved are too short for permitting the populations examined to differ in protein primary structure.
3) The morphological changes in question are therefore likely to be attributable to changes in gene regulation.

The findings of Turner thus point to the existence of a *gene regulatory evolutionary clock,* a possibility already considered by Whitt (1983), Parker et al. (1985), and by Kettler et al. (1986).

Clocklike progressive alterations could involve regulator protein/polynucleotide receptor pairs or, rather, more generalized regulatory interaction sets, namely regulatory interaction units, which have been called controller nodes (Zuckerkandl 1978a,b). The process would consist predominantly in incremental decreases in the mutual affinity of controller node components (Zuckerkandl 1976c, p. 407, 1978b, p. 132), reflected in changes in rate and timing of gene expression, with compensatory increases in affinity of course also occurring. In most cases, rates of morphological change probably are not proportional to rates of regulatory change in a single gene. Except when "master switches" are involved, morphological change must depend on combined changes in rates or timing of several genes, with effects of their interactions probably nonlinear. Morphological change then would not be expected to display quasi-clocklike behavior, even if the changes in rate of expression of each component gene did. A linear relationship between rates of morphological and gene regulatory change may hold in relatively simple cases such as the morphogenesis of teeth, where the differences in individual morphological characters may directly depend on the stronger or weaker expression of a single gene related to the morphological character through, say, the extracellular deposition of a greater or smaller amount of a gene product. Such morphological differences might be functionally quasi-neutral.

There are in fact two types of evolutionary changes in gene regulation (Zuckerkandl 1983): one concerns quantitative and temporal control of gene expression; the other bears on gene interaction patterns. The two types are interdependent via changes in temporal and spatial gene control. At this point, we are barely beginning to observe a clocklike behavior in gene regulatory changes of the first type. Might we expect a quasi-evolutionary clock to be at work also at the level of changes in gene interaction patterns? In linguistics, corresponding changes would presumably be those of syntactic relationships. A quasi-clock of changes in syntactic relationships would permit linguists to reach into a far more distant past of the evolution of languages than does the building of lexicographic trees. In contrast to linguistics, in molecular biology, many of the "words" (informational macromolecules), which are very ancient, are composed of enough letters and are often varying slowly enough to permit the construction of evolutionary trees for very high taxa. Likewise, gene interaction complexes would reveal relationships between very high taxa. *Subsets of gene interaction complexes can probably be as ancient as structural genes* (Zuckerkandl 1983). On the other hand, quantitative aspects of gene interaction and of gene expression apparently vary very much more rapidly and frequently during evolution than do the topological aspects of gene interaction. That these different types of evolutionary events will indeed be found to contribute to the ticking of an increasing number of quasi-clocks around us can be anticipated, but remains to be seen. Changes in gene interaction topology are unlikely to escape selective forces, whereas, in a larger proportion of cases, changes in rates of gene expression might. This circumstance may be thought to determine whether or not to expect that the evolutionary events conform to a "quasi-clock." However, the question is really open.

Is the difference in the variance of rates of mo-

lecular and morphological evolution as large as it has been thought to be? If the answer were no, that would bring a new twist to the old story of the clock, as told here. Looking at very large groups of taxa, such as vertebrates versus invertebrates (a singularly asymmetric grouping), morphological evolution appears slow and the differences in rate are only by a factor of two to three, invertebrates evolving faster than vertebrates (Gingerich 1983). These differences are of the same order, and even smaller, than those that have apparently been found in molecular evolution between certain taxa (Britten 1986, and others quoted above). Rates of morphological evolution observed over much shorter periods of time and smaller taxa ("post-Pleistocene mammalia") appear to be much faster. This reminds one of the "hidden substitutions" that can be inferred to have occurred over long enough spans of molecular evolution (see Ayala 1986). The same effect may be at work at the morphological level. Thus, the difference in the *mean* variance of evolutionary rates between these two levels may indeed not be as large as one had thought. In addition, the two levels may behave similarly in other ways. In particular, at both levels the magnitude of the differences in rate and of their changes can be seen to vary as a function of the "distance" between the observer and the observed, a distance that leads to either a fine-grained or a coarse-grained perception of events. Also, at both levels, the changes may often occur in spurts (Gould and Eldredge 1977; Gillespie 1984).

Clocks depend on the ability of structures and of their interactions to vary. Complementary to this ability is their invariance. How are variability and invariance distributed over the macromolecular and organismal levels? Though protein domains and exons are, for the most part, extremely ancient objects in terms of the homology groups to which they belong (see Zuckerkandl 1975), most of them probably almost as ancient as the most ancient rocks, there still is, in one sense, nothing equivalent in molecular evolution to the "living fossils" that have remained morphologically nearly unchanged, sometimes for hundreds of millions of years. In macromolecules, at sites where substitutions occur, perhaps in spurts, these spurts do not seem to be followed by periods of quiescence comparable to those observed in the morphology of certain animals. Primary structures of macromolecules continue to change when they are changeable. On the other hand, tertiary structures and, as I believe will eventually be found, also many narrowly circumscribed subsets within gene interaction systems are indeed "living fossils"; by themselves not living, of course, but thanks to life, always copied from an antique mold. Changes in that mold are few and far between, but they occur. To explore their distribution in time is for the future.

Acknowledgments. I gratefully acknowledge important critical contributions from the readers of earlier versions of this manuscript: Francisco Ayala, Wen-Hsiung Li, Linus Pauling, Gregory Whitt, and Allan Wilson, as well as the help of Mss. Constance Canevari, Ruth Reynolds, and Sandra Schwoebel. This work has been supported in part by The Japan Shipbuilding Industry Foundation.

References

Anfinsen CB (1959) The molecular basis of evolution. John Wiley & Sons, New York

Avise JC (1974) Systematic value of electrophoretic data. Syst Zool 23:465–481

Ayala FJ (1974) Biological evolution: natural selection or random walk? Am Sci 62:692–701

Ayala FJ (1986) On the virtues and pitfalls of the molecular evolutionary clock. J Hered 77:226–235

Baldwin E (1964) An introduction to comparative biochemistry. Cambridge University Press, Cambridge, England

Bernardi G, Bernardi G (1986) Compositional constraints and genome evolution. J Mol Evol 24:1–11

Bernardi G, Olofsson B, Filipski J, Zerial M, Salinas J, Cuny G, Meunier-Rotival M, Rodier F (1985) The mosaic genome of warm-blooded vertebrates. Science 228:953–958

Boyden AA (1958) Comparative serology: aims, methods and results. In: Cole W (ed) Serological and biochemical comparisons of proteins. Rutgers University Press, New Brunswick NJ

Britten RJ (1986) Rates of DNA sequence evolution differ between taxonomic groups. Science 231:1393–1398

Brown H, Sanger F, Kitai R (1955) The structure of pig and sheep insulins. Biochem J 60:556–565

Buonagurio DA, Nakada S, Parvin JD, Krystal M, Palese P, Fitch WM (1986) Evolution of human influenza A viruses over 50 years: a rapid uniform rate of change in NS gene. Science 232:980–982

Cavalli-Sforza LL, Edwards AWF (1964) Analysis of human evolution. In: Genetics today, Proc XI Int Congr Genetics. Pergamon, Oxford, England, pp 923–933

Chang L-YE, Slightom J (1984) Isolation and nucleotide sequence analysis of the β-type globin pseudogene from human, gorilla and chimpanzee. J Mol Biol 180:767–784

Cherry LM, Case SM, Wilson AC (1978) Frog perspective on the morphological difference between humans and chimpanzees. Science 200:209–211

Clarke AR, Wigley DB, Chia WN, Barstow D, Atkinson T, Holbrook JJ (1986) Site-directed mutagenesis reveals role of mobile arginine residue in lactate dehydrogenase catalysis. Nature 324:699–702

Derancourt J, Lebor AS, Zuckerkandl E (1967) Séquence des acides aminés, séquence des nucléotides et évolution. Bull Soc Chim Biol 49:577–607

DiMichele L, Powers DA (1982) Physiological basis for swimming endurance differences between LDH-B genotypes of *Fundulus heteroclitus*. Science 216:1014–1016

DiMichele L, Powers DA, DiMichele JA (1986) Developmental and physiological consequences of genetic variation at enzyme synthesizing loci in *Fundulus heteroclitus*. Am Zool 26:201–208

Doolittle RF, Blomback B (1964) Amino-acid sequence investigations of fibrinopeptides from various mammals: evolutionary implications. Nature 202:147–152

Doolittle WF, Sapienza C (1980) Selfish genes, the phenotype paradigm and gene evolution. Nature 284:601–603

Dover G (1982) Molecular drive: a cohesive mode of species evolution. Nature 299:111–117

Fitch WM (1970) Distinguishing homologous from analogous proteins. Syst Zool 19:99
Fitch WM (1976) Molecular evolutionary clocks. In: Ayala FJ (ed) Molecular evolution. Sinauer Associates, Sunderland MA, pp 160–178
Fitch WM, Atchley WR (1985a) Evolution in inbred strains of mice appears rapid. Science 228:1169–1175
Fitch WM, Atchley WR (1985b) Letter. Science 230:1408–1409
Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science 155:279–284
Florkin M (1944) L'evolution biochemique. Masson, Paris
Gasser SM, Laemmli UK (1986) Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. malanogaster*. Cell 46:521–530
Gillespie JH (1984) The molecular clock may be an episodic clock. Proc Natl Acad Sci USA 81:8009–8013
Gillespie JH (1986) Natural selection and the molecular clock. Mol Biol Evol 3:138–155
Gingerich PD (1983) Rates of evolution: effects of time and temporal scaling. Science 222:159–161
Goodman M (1962) Evolution of the immunologic species specificity of human serum proteins. Human Biol 34:104–150
Goodman M (1963) Serological analysis of the systematics of recent hominoids. Human Biol 35:377–436
Goodman M, Braunitzer G, Stangl A, Schrank B (1983) Evidence on human origins from haemoglobins of African apes. Nature 303:546–548
Gould SJ, Eldredge N (1977) Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology 3:115–151
Hartl D, Dykhuizen D (1979) A selectively driven molecular clock. Nature 281:230–231
Holmquist GP (1987) DNA sequences in G-bands and R-bands. In: Adolph KW (ed) Chromosome and chromatin structure. CRC Press, Boca Raton FL (in press)
Imanaka T, Shibazaki M, Takagi M (1986) A new way of enhancing the thermostability of proteases. Nature 324:695–697
Ingram VM (1961) Gene evolution and the haemoglobins. Nature 139:704–708
Itano HA (1957) The human hemoglobins: their properties and genetic control. Adv Protein Chem 12:215–268
Jukes TH (1963) Some recent advances in studies of the transcription of the genetic message. Adv Biol Med Phys 9:1–41
Kettler MK, Ghent AW, Whitt GS (1986) A comparison of phylogenies based on structural and tissue-expressional differences of enzymes in a family of teleost fishes (Salmoniformes: Umbridae). Mol Biol Evol 3:485–498
Kimura M (1968) Evolutionary rate at the molecular level. Nature 217:624–626
Kimura M (1969) The rate of molecular evolution considered from the standpoint of population genetics. Proc Natl Acad Sci USA 63:1181–1188
Kimura M (1979) The neutral theory of molecular evolution. Sci Am 241:94–104
Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge, England
Kimura M, Ohta T (1974) On some principles governing molecular evolution. Proc Natl Acad Sci USA 71:2848–2852
King JL, Jukes TH (1969) Non-Darwinian evolution. Science 164:788–798
King M-C, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. Science 188:107–116
Koehn RK, Zera AJ, Hall JG (1983) Enzyme polymorphism and natural selection. In: Nei M, Koehn RK (eds) Evolution of genes and proteins. Sinauer Associates, Sunderland MA, pp 115–136
Laird CD, McConaughy BL, McCarthy BJ (1969) Rate of fixation of nucleotide substitutions in evolution. Nature 224: 149–154
Li W-H, Wu C-I, Luo C-C (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol 2:150–174
Manuelidis L (1984) Different central nervous system cell types display distinct and nonrandom arrangements of satellite DNA sequences. Proc Natl Acad Sci UA 81:3123–3127
Margoliash E (1963) Primary structure and evolution of cytochrome c. Proc Natl Acad Sci USA 50:672–679
Mas MT, Chen CY, Hitzeman RA (1986) Active human-yeast chimeric phosphoglycerate kinases engineered by domain interchange. Science 233:788–790
Mayr E (1982) The growth of biological thought; diversity, evolution, and inheritance. Harvard University Press, Cambridge MA
Maxson LR, Wilson AC (1975) Albumin evolution and organismal evolution in tree frogs (*Hylidae*). Syst Zool 24:1–15
Nevo E, Lavie B, Ben-Shlomo R (1983) Selection of allelic isozyme polymorphisms in marine organisms: pattern, theory, and application. In: Rattazzi MC, Scandalios JG, Whitt GS (eds) Isozymes, current topics in biological and medical research, vol 10. Genetics and evolution. Alan R Liss, New York, pp 69–92
Nuttall GHF (1904) Blood immunity and blood relationship. Cambridge University Press, Cambridge, England
Ohta T (1973) Slightly deleterious mutant substitutions in evolution. Nature 246:96–98
Orgel LE, Crick FHC (1980) Selfish DNA: the ultimate parasite. Nature 284:604–607
Parker HR, Philipp DP, Whitt GS (1985) Gene regulatory divergence among species estimated by altered developmental patterns in interspecific hybrids. Mol Biol Evol 2:217–250
Pauling L, Zuckerkandl E (1963) Chemical paleogenetics: molecular "restoration studies" of extinct forms of life. Acta Chem Scand [B] 17:9–16
Perler F, Efstratiadis A (1980) The evolution of genes: the chicken preproinsulin gene. Cell 20:555–566
Perutz MF (1983) Species adaptation in a protein molecule. Mol Biol Evol 1:1–28
Perutz MF, Kendrew JC, Watson HC (1965) Structure and function of haemoglobin. II. Some relations between polypeptide chain configuration and amino acid sequence. J Mol Biol 13:669–678
Petit C, Zuckerkandl E (1976) Evolution, genetique des populations, evolution moleculaire. Hermann, Paris
Philipp DP, Childers WF, Whitt GS (1985) Correlations of allele frequences with physical and environmental variables for populations of largemouth bass, *Micropterus salmoides* (*Lacepede*). J Fish Biol 27:347–365
Powers DA, DiMichele L, Place AR (1983) The use of enzyme kinetics to predict differences in cellular metabolism, developmental rate, and swimming performance between LDH-B genotypes of the fish, *Fundulus heteroclitus*. In: Rattazzi MC, Scandalios JG, Whitt GS (eds) Isozymes, current topics in biological and medical research, vol 10. Genetics and evolution. Alan R Liss, New York, pp 147–170
Reichert ET, Brown AP (1909) The differentiations and specificity of corresponding proteins. In: Carnegie Inst Washington 116, 338 pp
Sarich VM, Wilson AC (1967a) Rates of albumin evolution in primates. Proc Natl Acad Sci USA 58:142–148
Sarich VM, Wilson AC (1967b) Immunological time scale for hominid evolution. Science 158:1200–1203
Sarich VM, Wilson AC (1973) Generation time and genomic evolution in primates. Science 179:1144–1147

Selander RK, Johnson WE (1973) Genetic variation among vertebrate species. Annu Rev Ecol Syst 4:75

Sibley CG, Ahlquist JE (1984) The phylogeny of the hominoid primates, as indicated by DNA–DNA hybridization. J Mol Evol 10:2–15

Sibley CG, Ahlquist JE (1986) Reconstructing bird phylogeny by comparing DNA's. Sci Am 254:82–92

Simpson GG (1953) The major features of evolution. Columbia University Press, New York

Sneath PHA (1966) Relations between chemical structure and biological activity in peptides. J Theor Biol 12:157–195

Sneath PHA (1980) The estimation of differences in protein evolution rates. Proc Geol Assoc 91:71–79

Stenzel P (1974) Opossum Hb chain sequence and neutral mutation theory. Nature 252:62–63

Strauss F, Varshavsky A (1984) A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell 37:889–901

Swadesh M (1950) Salishan internal relationships. Int J Am Linguistics 16:157–167

Taylor WR (1986) The classification of amino acid conservation. J Theor Biol 119:205–218

Turner CG (1986) Dentochronological separation estimates for Pacific Rim populations. Science 232:1140–1142

Vawter L, Brown WM (1986) Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. Science 234:194–196

Vogel H, Zuckerkandl E (1971a) Randomness and "thermodynamics" of molecular evolution. In: Schoffeniels E (ed) Biochemical evolution and the origin of life. North-Holland, Amsterdam, pp 352–365

Vogel H, Zuckerkandl E (1971b) The evolution of polarity relations in globins. In: Neyman J (ed) Darwinian, Neo-Darwinian, and non-Darwinian evolution, Proc. 6th Berkeley Symp. Mathematical Statistics and Probability, vol 5. University of California Press, Berkeley, pp 155–176

Walker R (1983) The molecular biology of enzyme synthesis; regulatory mechanisms of enzyme adaptation. John Wiley & Sons, New York

Wallace DG, Maxson LR, Wilson AC (1971) Albumin evolution in frogs: a test of the clock hypothesis. Proc Natl Acad Sci USA 68:3127–3129

Wallace, DG, King M-C, Wilson AC (1973) Albumin differences among ranid frogs: taxonomic and phylogenetic implications. Syst Zool 22:1–13

Whitt GS (1983) Isozymes as probes and participants in developmental and evolutionary genetics. In: Rattazzi MC, Scandalios JG, Whitt GS (eds) Isozymes, current topics in biological and medical research, vol 10. Genetics and evolution. Alan R Liss, New York, pp 1–40

Wilson AC (1975) Evolutionary importance of gene regulation. Stadler Symposium, University of Missouri 7:117–134

Wilson AC, Maxson LR, Sarich VM (1974) Two types of molecular evolution. Evidence from studies of interspecific hybridization. Proc Natl Acad Sci USA 71:2843–2847

Wistow GJ, Mulders JWM, de Jong WW (1987) The enzyme lactate dehydrogenase as a structural protein in avian and crocodilian lenses. Nature 326:622–624

Wu C-I, Li W-H (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci USA 82:1741–1745

Yokoyama S, Gojobori T (1987) Molecular evolution and phylogeny of the human AIDS viruses LAV, HTLV-III, and ARV. J Mol Evol 24:330–336

Zimmer EA, Martin SL, Beverly SM, Kan YW, Wilson AC (1980) Rapid duplication and loss of genes coding for the α-chains of hemoglobin. Proc Natl Acad Sci USA 77:2158–2162

Zuckerkandl E (1963) Perspectives in molecular anthropology. In: Washburn SL (ed) Classification and human evolution. Aldine, Chicago, pp 243–272

Zuckerkandl E (1964a) Controller-gene diseases: the operon model as applied to β-thalassemia, familial fetal hemoglobinemia and the normal switch from the production of fetal hemoglobin to that of adult hemoglobin. J Mol Biol 8:128–147

Zuckerkandl E (1964b) Further principles of chemical paleogenetics as applied to the evolution of hemoglobin. In: Peeters H (ed) Protides of the biological fluids. Elsevier, Amsterdam, pp 102–109

Zuckerkandl E (1965) The evolution of hemoglobin. Sci Am 212(5):110–118

Zuckerkandl E (1968) Hemoglobins, Haeckel's "biogenetic law," and molecular aspects of development. In: Rich A, Davidson N (eds) Structural chemistry and molecular biology: a volume dedicated to Linus Pauling by his students, colleagues, and friends. WH Freeman, San Francisco, pp 256–274

Zuckerkandl E (1972) Some aspects of protein evolution. Biochimie 54:1095–1102

Zuckerkandl E (1975) The appearance of new structures and functions in proteins during evolution. J Mol Evol 7:1–57

Zuckerkandl E (1976a) Evolutionary processes and evolutionary noise at the molecular level. I. Functional density in proteins. J Mol Evol 7:167–183

Zuckerkandl E (1976b) Evolutionary processes and evolutionary noise at the molecular level. II. A selectionist model for random fixations in proteins. J Mol Evol 7:269–311

Zuckerkandl E (1976c) Programs of gene action and progressive evolution. In: Goodman M, Tashian RE (eds) Molecular anthropology—genes and proteins in the evolutionary ascent of the primates. Plenum, New York, pp 387–447

Zuckerkandl E (1978a) Multilocus enzymes, gene regulation, and genetic sufficiency. J Mol Evol 12:57–89

Zuckerkandl E (1978b) Molecular evolution as a pathway to man. Z Morphol Anthropol 69:117–142

Zuckerkandl E (1983) Topological and quantitative relationships in evolving genomes. In: Helene C (ed) Structure, dynamics, interactions and evolution of biological macromolecules. Reidel, Dordrecht, pp 395–412

Zuckerkandl E (1986) Polite DNA: functional density and functional compatibility in genomes. J Mol Evol 24:12–27

Zuckerkandl E, Pauling L (1962) Molecular disease, evolution, and genic heterogeneity. In: Kasha M, Pullman B (eds) Horizons in biochemistry. Academic Press, New York, p 189

Zuckerkandl E, Pauling L (1964) Molecules as documents of evolutionary history. In: Problems of evolutionary and technical biochemistry. Science Press, Academy of Sciences of the USSR, pp 54–62 (in Russian); republished in English (1965a), J Theor Biol 8:357–366

Zuckerkandl E, Pauling L (1965b) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, New York, pp 97–165

Zuckerkandl E, Schroeder WA (1961) Amino acid composition of the polypeptide chains of gorilla haemoglobin. Nature 192: 984–985

Zuckerkandl E, Jones RT, Pauling L (1960) A comparison of animal hemoglobins by tryptic peptide pattern analysis. Proc Natl Acad Sci USA 46:1349–1360

Zuckerkandl E, Derancourt J, Vogel H (1971) Mutational trends and random processes in the evolution of informational macromolecules. J Mol Biol 59:473–490

Received April 25, 1987