J Mol Evol (1987) 26:87–98

Journal of Molecular Evolution

Transitions, Transversions, and the Molecular Evolutionary Clock

Thomas H. Jukes

Space Sciences Laboratory, University of California, Berkeley, California 94720, USA

Summary. Nucleotide substitutions in the form of transitions (purine–purine or pyrimidine–pyrimidine interchanges) and transversions (purine–pyrimidine interchanges) occur during evolution and may be compiled by aligning the sequences of homologous genes. Referring to the genetic code tables, silent transitions take place in third positions of codons in family boxes and two-codon sets. Silent transversions in third positions occur only in family boxes, except for A ⇌ C transversions between AGR and CGR arginine codons (R = A or G). Comparisons of several protein genes have been made, and various subclasses of transitional and transversional nucleotide substitutions have been compiled. Considerable variations occur among the relative proportions of transitions and transversions. Such variations could possibly be caused by mutator genes, favoring either transitions or, conversely, transversions, during DNA replication. At earlier stages of evolutionary divergence, transitions are usually more frequent, but there are exceptions. No indication was found that transversions usually originate from multiple substitutions in transitions.

Key words: Transitions — Transversions — Molecular evolutionary clock — Genetic code — Evolutionary divergence

Introduction

Historical Review of the "Clock"

The time taken by evolutionary changes is usually measured by combining the fossil record with geological data, including isotopic and radiometric dating. A molecular method of measuring the progress of evolution is by comparison of homologous sequences of DNA in terms of nucleotide differences or of protein molecules in terms of differences in amino acid residues. Nuttall (1904) showed that taxonomy was related to comparative serological tests, but the first molecular comparisons were made by crystallography of hemoglobins (Reichert and Brown 1909). This was long before it had been discovered (Sanger et al. 1952) that proteins consisted of chains of amino acids joined in peptide linkage. No one knew the chemical interpretation of the measurements by Reichert and Brown, except that a difference in crystallographic parameters meant a molecular difference. Nevertheless, the results were well accepted, and were published in leading textbooks of biochemistry many years ago, such as the text by Hawk (1907). Reichert and Brown examined their results on a taxonomic basis and found that hemoglobin crystals of hemoglobins from the various species of a genus belonged to the same crystallographic system and generally to the same crystallographic group. Cats, dogs, a bear, and a seal clearly could be placed in four different groups on the basis of the axial ratios in crystals of their reduced hemoglobins. A sample of blood labeled as that of a baboon was found by crystallographic examination to be that of a cat, and subsequent follow-up showed that the vial had been mislabeled (Reichert and Brown 1907).

Sequencing of proteins started about 1952, and in 1961, Ingram published a phylogenetic tree showing the successive separation of myoglobin and alpha, gamma, beta, and delta hemoglobins from common ancestors. Zuckerkandl (Zuckerkandl and Pauling 1965, p. 148) was the first to use the term "molecular evolutionary clock." This was in reference to the number of differences observed when

Offprint requests to: T.H. Jukes

the amino acid sequences of hemoglobins in various species were compared with each other. The horse–human difference was greater than the human–gorilla difference, corresponding to the times of evolutionary separation of two species from a common ancestor. Before this, no one had compared phylogenetic differences in sequences on the basis of time. The clock terminology came into wide usage, and may have led to the idea that amino acid substitutions take place at an exactly uniform rate for any one protein during evolution. There was never any doubt that the clock ran at different speeds in two different proteins, such as in cytochrome *c* and hemoglobin (Dickerson 1971).

Nonconstancy of the rate of the molecular evolutionary clock in different species was shown for cytochrome *c* (Jukes and Holmquist 1972a). The rate of replacement of amino acids was more than twice as fast in the rattlesnake as in another reptile, the snapping turtle. Subsequently, there have been many other reports of fluctuations in the molecular evolutionary clock, or even claims that it does not exist. It is now obvious that there are numerous irregularities in the rate at which the clock runs, and it is also evident that some but not all variations tend to "smooth out" when enough readings are taken over a prolonged time lapse. Given the many variable components that are involved in biological evolution, it could scarcely be expected that molecular changes would have a constant rate. At the same time, enough examples have been documented showing a roughly linear relationship between amino acid replacements and time in various proteins (see Dickerson 1971, Fig. 3) to justify the acceptance of the "molecular evolutionary clock" hypothesis, based on elapsed time and not on generation time. For if generation time, rather than elapsed time, controlled the clock in cytochrome *c*, it would be impossible to correlate amino acid replacements with time in species that vary in generation time from years (mammals), to 1 year (wheat), to weeks (insects), to hours (yeast). The molecular evolutionary clock in cytochrome *c* was cited as evidence for evolution in the amicus curiae brief of the National Academy of Sciences, USA for the US Supreme Court appeal against the Louisiana creationist law (Edwards versus Aguillard 1985).

The term "molecular evolutionary clock" is advantageous for drawing attention to steady substitutional changes in sequences of informational macromolecules during evolution. It has the disadvantage of implying an almost automatic, centrally controlled rate of change, similar to isotopic change in radioactive elements. This implication is incorrect. Perhaps the term "molecular evolutionary clock" should give way to "relation between molecular evolutionary changes and time."

In this communication, I examine the rates of occurrence of transitions (exchanges between purines and exchanges between pyrimidines) and transversions (purine–pyrimidine interchanges) in several genes.

Procedures

Evolutionary replacements of amino acids in the sequences of protein molecules vary in rate, depending on the changes in protein structure and function that these replacements produce. In genes, evolutionary nucleotide substitutions that produce amino acid replacements ("replacement substitutions") occur less frequently than silent substitutions. The difference was expressed quantitatively by Jukes and Bhushan (1986) for several genes.

Silent nucleotide substitutions in comparisons of the protein-coding regions of genes are either transitions or transversions. More information can be obtained if these two categories are subdivided according to the types of codons in which these substitutions take place, because in some types of codons, transversions are accepted more readily than in other types. This circumstance was briefly noted by Brown and Simpson (1982), and also in 1982, I presented a paper at the meeting of the Society for Study of Evolution at Stony Brook, New York (Jukes 1982), largely devoted to this point, in which both transitions and transversions were divided into silent, unconstrained ("optional") and constrained ("obligate") categories. Unconstrained transitions and transversions occur in silent, usually third positions of codons belonging to fourfold sets ("family boxes") in which all four codons specify a single amino acid. Unconstrained transitions also occur in third positions of two-codon sets, such as UUU and UUC, phenylalanine, but third-position transversions in these sets produce amino acid replacements and are therefore less frequent than transitions.

This difference between the properties of family-box codons and two-codon-set codons has marked effects on the occurrence of transitions and transversions in silent or potentially silent sites.

The following can give rise to silent nucleotide substitutions in the universal code:

1) Third positions in family-box codons (CUN, GUN, UCN, CCN, ACN, GCN, CGN, GGN) (N = A, C, G or U, R = A or G, Y = C or U).
2) Third positions in AUU, AUC, AUA.
3) Third-position transitions in two-codon sets (UUY, UUR, UAY, CAY, CAR, AAY, AAR, GAY, GAR, UGY, AGY, AGR).
4) First-position transitions in exchanges between CUR and UUR. The set YUR is therefore a two-codon set with respect to first-position transitions.

5) First-position C/A transversions in exchanges between CGR and AGR. First-position C/G transversions in CGR are replacement substitutions. First-position A/G transitions in AGR are replacement substitutions.

The same list applies to the vertebrate mitochondrial code, except that AUY, AUR, and UGR are added to the list of two-codon sets, and AGR is omitted from the list because AGA and AGG are chain terminators.

Some comparisons involve ambiguities. For example, in the case of TTC to TCA, the substitution in the third position may be either via TTC → TCC → TCA, a silent transversion in the third position, or via TTC → TTA → TCA, a replacement transversion. If the first of these two pathways is assumed, then introducing a replacement substitution in the third position is avoided. Since the first choice is the more probable one, it is used for the sake of simplicity, without attempting to estimate and include the lower probabilities for the second pathway.

Interchanges between a family-box codon and a two-codon-set codon are less frequent than interchanges between two family-box codons or between two two-codon-set codons.

Brown and co-workers (1982) proposed that transitions are strongly favored over transversions in mitochondrial evolution. They compared a sequenced mitochondrial DNA fragment, 896 base pairs in length, in various species. In closely related pairs of primates, 92% of the nucleotide substitutions were transitions, and, as divergence increased, the percentage of transitions fell, so that in comparisons between primates, and cow and mouse, 45% were transitions. These authors proposed that multiple substitutions at the same site tend to replace transitions by transversions. Their proposal was discussed and criticized by Jukes (1982) and Holmquist (1983). Holmquist concluded that "multiple substitutions *per se* do not lead to a decrease in transition differences with increasing evolutionary divergence." In this communication, more data are examined and further discussion of the proposals by W.M. Brown and colleagues is presented.

Multiple Hits and Revertants

Multiple nucleotide substitutions (multiple hits) at a site in DNA or in RNA may be estimated by the formula:

$$\mu = \tfrac{3}{4} \ln \frac{3}{3 - 4p} \qquad (1)$$

where μ is the mean number of nucleotide substitutions at a single site and p is the observed fraction of residues with single nucleotide differences (Jukes and Cantor 1969).

This formula assumes that transitions are numerically equal to transversions. Kimura (1980) pointed out that transitions are often more frequent than transversions in evolution "particularly when the third positions of codons are compared." He formulated the following equation for the ratio (r) of transitions to transversions

$$r = \frac{-\ln(1 - 2P - Q) + (\tfrac{1}{2})\ln(1 - 2Q)}{-\ln(1 - 2Q)} \qquad (2)$$

where P and Q are the observed fractions of transition type and transversion type differences between the two sequences compared.

Amino acids do not replace each other; their replacements are actually an expression of nucleotide substitutions, and although point mutations that change nucleotides may occur at random, their evolutionary fixation as substitutions is not randomized, but depends on constraints that vary with context.

Results

Genes of three proteins of different characteristics were used to study changes in numbers of silent substitutions accompanying increasing evolutionary divergence: (a) four cytochrome *c* genes, (b) five mammalian and two *Drosophila* mitochondrial cytochrome oxidase genes, component II (COII), and (c) trout and two sea urchin histone 2A and 3 genes. In all cases, in proceeding from the closely to the more distantly related gene sequences the increase in silent transversions was too great to be explainable by multiple substitutions at the same sites.

Cytochrome c

This slowly evolving protein, occurring in a range of organisms from yeasts to mammals, has long been a favorite subject for evolutionary studies.

Nucleotide sequences of four cytochrome *c* genes (Scarpulla et al. 1981; Limbach and Wu 1983, 1985a,b) were used for comparisons (Fig. 1). Table 1 summarizes some evolutionary changes in these four genes. In closely related species (rat/mouse), only silent transitions have occurred. Silent transversions in family-box sites then increase steadily as evolution proceeds from mammals to a bird and then to an insect, but silent transitions in family-box sites plateau at a level of about 10% and in two-codon-set sites at a level of about 20%. Most striking are the low values for transversions in two-codon-

```
Mouse          GGT GAT GTT GAA AAA GGC AAG AAG ATT TTT GTT CAG AAG TGT GCC CAG TGC CAC ACT GTG GAA AAG GGA GGC AAG CAT AAG ACT GGA CCA
Rat            GGT GAT GTT GAA AAA GGC AAG AAG ATT TTT GTT CAA AAG TGT GCC CAG TGC CAC ACT GTG GAA AAA GGA GGC AAG CAT AAG ACT GGA CCA
Chicken        GGA GAT ATT GAG AAG GGC AAG AAG ATT TTT GTC CAG AAA TGT TCC CAG TGC CAT ACG GTT GAA AAA GGA GGC AAG CAC AAG ACT GGA CCC
Drosophila 4   GGT GAT GTT GAG AAG GGA AAG AAG CTG TTC GTG CAG CGC TGC GCC CAG TGC CAC ACC GTT GAG GCT GGT GGC AAG CAC AAG GTT GGA CCC

Mouse          AAT CTC CAC GGT CTG TTC GGG CGG AAG ACA GGC CAG GCT GCT GGA TTC TCT TAC ACA GAT GCC AAC AAG AAC AAA GGC ATC ACC TGG GGA
Rat            AAC CTC CAT GGT CTG TTT GGG CGG AAG ACA GGC CAG GCT GCT GGA TTC TCT TAC ACA GAT GCC AAC AAG AAC AAA GGT ATC ACC TGG GGA
Chicken        AAC CTT CAT GGC CTG TTT GGA CGC AAA ACA GGA CAA GCT GAG GGC TTC TCT TAC ACA GAT GCC AAT AAG AAC AAA GGT ATC ACT TGG GGT
Drosophila 4   AAT CTG CAT GGT CTG ATC GGT CGC AAG ACC GGA CAG GCG GCC GGA TTC GCG TAC ACG GAC GCC AAC AAG GCC AAG GGC ATC ACC TGG AAC

Mouse          GAG GAT ACC CTG ATG GAG TAT TTG GAG AAT CCC AAA AAG TAC ATC CCT GGA ACA AAA ATG ATC TTC GCT GGA ATT AAG AAG AAG GGA GAA
Rat            GAG GAT ACC CTG ATG GAG TAT TTG GAA AAT CCC AAA AAG TAC ATC CCT GGA ACA AAA ATG ATC TTC GCT GGA ATT AAG AAG AAG GGA GAA
Chicken        GAG GAT ACT CTG ATG GAG TAT TTG GAA AAT CCA AAG AAG TAC ATC CCA GGA ACA AAG ATG ATT TTT GCG GGT ATC AAG AAG AAG TCT GAG
Drosophila 4   GAG GAC ACC CTG TTC GAG TAC CTG GAG AAC CCC AAG AAG TAC ATC CCC GGC ACC AAG ATG ATC TTC GCC GGT CTG AAG AAG CCC AAC GAG

Mouse          AGG GCA GAC CTA ATA GCT TAT CTT AAA AAG GCT AAT GAG TAA
Rat            AGG GCA GAC CTG ATA GCT TAT CTT AAA AAG GCT AAT GAA TAA
Chicken        AGA GTA GAC TTA ATA GCA TAT CTC AAA GAT GCC ACT AAG TAA
Drosophila 4   CGC GGC GAT CTG ATC GCC TAC CTG AAG TCG GCG ACC AAG TAA
```

Fig. 1. Nucleotide sequences of cytochrome *c* genes from mouse, rat, chicken, and *Drosophila melanogaster*

set sites, reflecting a resistance to accept amino acid replacements. Assuming uniform rates of nucleotide substitutions in genes, silent transversions in family-box sites are accepted at about 20 times the rate of replacement transversions in two-codon-set sites. This rejection of replacements is to be expected from the slow rate of amino acid replacements during the evolution of cytochrome *c* (Dickerson and Timkovich 1974).

The proposal can be tested by examining comparisons of cytochrome *c* from rat, mouse, and chicken in Table 1. There are about 49% of family-box sites in the codons of each of the three genes. The number of silent transversions per 100 family-box sites has increased during evolutionary divergence from 0 in rat : mouse comparisons to 33 (32 in 49 + 49 = 98 family boxes) in mammal : chicken. If these 33 resulted from multiple "hits," the number of such hits can be calculated from the percentage of identical sites, 47%, in family boxes. From Poisson distribution, for 47 identical sites, there should be 35 single hits and 18 multiple hits. Even if all the multiple hits produced transversions, which would not be remotely possible if transitions are favored, there should be only 18 rather than 33 transversions. Clearly, some other process rather than multiple hits must be responsible for the increase in transversions during evolutionary divergence.

Going to the next example, *Drosophila*, there is a further increase in silent transversions from approximately 16 (15 and 17) (mammal versus bird) to approximately 28 (27, 27, and 30) (vertebrate versus insect) (see Table 1, third column of figures). If these were produced by multiple substitutions, all of which resulted in transversions, about 107 hits for 100 sites would be needed, of which approximately 37 would be single hits, and approximately 28, multiple hits. In this case, all single substitutions would have to be transitions and all multiple substitutions transversions. Again, this seems highly improbable. Note that revertants from transitions can be produced only by transitions.

Table 1. Nucleotide substitutions in potentially silent sites of cytochrome *c*

	Silent sites					
	In family boxes			In two-codon sets		
Genes compared	Un-sbs.	Trt.	Trv.	Un-sbs.	Trt.	Trv.
Rat : mouse	47	2	0	54	7	0
Rat : chicken	23	11	15	42	16	1
Mouse : chicken	24	9	17	40	19	1
Rat : *Drosophila*	21	10	27	29	22	1
Mouse : *Drosophila*	21	8	27	33	19	2
Chicken : *Drosophila*	17	11	30	31	20	1

Unsbs. = unsubstituted; Trt. = transitions; Trv. = transversions

Changes in transitions and transversions during evolutionary divergence are shown in Table 2. In Table 2, the numerical values were calculated as follows: IC to TFB, MR:MC = 10, means that 10 identical codons in the mouse : rat comparison are

Table 2. Changes in transitions and transversions during evolutionary divergence of cytochrome *c* genes in mouse, rat, chicken, and *Drosophila melanogaster*

Changes	MR:MC	MR:RC	MC:CD4	RC:CD4
IC to TFB	10	10	2	4
IC to VFB	16	16	12	14
IC to T2CS	12	12	12	14
TFB to IC	0	0	0	0
VFB to IC	0	0	4	5
T2CS to IC	5	2	5	7
TFB to VFB	0	0	6	6
VFB to TFB	0	0	3	2
T2CS to VFB	0	0	3	2
T2CS to TFB	0	0	1	0
Net changes				
TFB	10	10	0	0
VFB	16	16	14	15
T2CS	7	10	3	5

IC = identical codons; TFB = transitions in family-box sites; VFB = transversions in family-box sites; T2CS = transitions in two-codon sets; M = mouse; R = rat; C = chicken; D4 = *Drosophila* cytochrome *c* 4

replaced by transitions in family boxes in the mouse : chicken comparison. Similarly, T2CS to IC = 5 means that 5 transitions in two-codon sets in the mouse : rat comparison are replaced by identical sites in the mouse : chicken comparison. The data in Table 2 show that silent nucleotide substitutions in these genes are in a state of flux, particularly in the case of the mammal : insect comparisons.

Perler and colleagues (1980) suggested that silent nucleotide substitutions in evolution increase rapidly to a saturation point beyond which they are constrained "due to the imposition of required secondary structure at the mRNA level" or "specific anticodon requirements." The sequences in Fig. 1 were examined for evidence of this idea. Thirty of the 109 potentially silent sites are unchanged, but if the four sequences are aligned with three other cytochrome *c* sequences, *Drosophila*-3 and yeast iso-1 and iso-2 (Smith et al. 1979; Montgomery et al. 1980; Limbach and Wu 1985b), all but four of these unchanged 30 sites have been substituted. The proposal by Perler and co-workers (1980) is, therefore, not substantiated.

Cytochrome Oxidase Subunit II (COII)

Cytochrome oxidase subunit II (COII) is coded by a mitochondrial gene, for which sequences have been determined in several species, including human, bovine, mouse, Norway rat [*Rattus norvegicus* (NR)], black rat [*R. rattus* (RR)] (Anderson et al. 1981, 1982; Bibb et al. 1981; Brown and Simpson 1982), and two species of *Drosophila* (*D. melanogaster* and *D. yakuba*) (de Bruijn 1983; Clary and Wolstenholme 1985). The gene evolves rapidly and is excellent subject material for studying transitions and transversions (Table 3). It contains about 256 potentially silent sites in 226 codons. The COII sequences in the cited references were used to make the comparisons shown in Tables 3 and 4.

Table 3. Substitutions of potentially silent nucleotide sites in COII genes

	Potentially silent sites					
	In family boxes			In two-codon sets		
Comparisons	Unsbs.	Trt.	Trv.	Unsbs.	Trt.	Trv.
NR : RR	85	19	7	112	25	—
Mouse : NR	49	19	41	109	42	—
Mouse : RR	51	13	45	101	45	1
Bovine : NR	44	21	49	93	42	4
Bovine : RR	50	16	53	87	42	4
Bovine : mouse	42	23	50	81	54	3
Human : NR	49	26	56	79	35	6
Human : RR	47	25	57	83	34	6
Human : mouse	49	19	60	77	39	7
Human : bovine	44	25	59	71	49	6
Dros M : Dros Y	89	6	16	133	17	—
NR : Dros M	50	17	67	55	46	15
NR : Dros Y	56	14	62	56	39	14
RR : Dros M	49	16	68	57	44	16
RR : Dros Y	57	13	62	56	44	15
Mouse : Dros M	54	18	60	60	40	18
Mouse : Dros Y	58	19	62	61	36	17
Bovine : Dros M	53	16	66	58	39	14
Bovine : Dros Y	55	11	67	64	39	13
Human : Dros M	39	27	76	37	57	10
Human : Dros Y	41	29	74	39	54	9

Unsbs. = unsubstituted; Trt. = transitions; Trv. = transversions; NR = *Rattus norvegicus*; RR = *Rattus rattus*; Dros M = *Drosophila melanogaster*; Dros Y = *Drosophila yakuba*

The silent substitutions are grouped in two categories in Table 3. In comparisons involving family-box sites, transitions and transversions should both be silent, and other things being equal, there are half as many opportunities for transitions as for transversions. The ratio between the two would therefore be 1:2. This expectation is borne out by the results. Excluding the comparison of the two rat species (see below), the ratio of transitions to transversions averages 1:2.5.

Two-Codon Sets

The constraints against transversions in two-codon sets of COII are evident in Table 3, which shows that only 8% of substitutions in these sets are transversions in comparisons between mammals. In the case of mammal : *Drosophila* comparisons, transversions in two-codon sets are much more frequent because amino acid replacements have increased markedly (Tables 3 and 4), showing that the con-

straints against replacement substitutions have become less with increasing divergence.

Table 4 shows the amino acid differences and the number of silent nucleotide substitutions in the six COII genes. There is a great disparity between the two sets of figures. Silent nucleotide substitutions accumulate very rapidly at first, then occur much more slowly. Amino acid replacements occur very slowly at first, as seen in the mouse : rat comparisons, but then become much more rapid. Mouse and rat diverged about 10 million years (Myr) ago (Catzeflis et al. 1987), human and the other three mammals about 75–100 Myr ago.

Table 4. Amino acid replacements and silent nucleotide substitutions during evolutionary divergence of COII genes

Comparison	Amino acid differences	Total replacement substitutions	Silent nucleotide substitutions
NR : RR	3	3	51
Mouse : NR	2	3	102
Mouse : RR	2	2	103
Bovine : NR	19	23	112
Bovine : RR	19	22	111
Bovine : mouse	20	23	127
Human : NR	61	75	117
Human : RR	59	76	116
Human : mouse	62	77	118
Human : bovine	58	71	133
Dros M : Dros Y	4	5	39
NR : Dros M	95	125	130
NR : Dros Y	97	130	115
RR : Dros M	95	124	127
RR : Dros Y	96	129	119
Mouse : Dros M	94	124	118
Mouse : Dros Y	97	131	117
Bovine : Dros M	92	126	121
Bovine : Dros Y	92	127	117
Human : Dros M	99	135	160
Human : Dros Y	99	132	157

NR = *Rattus norvegicus*; RR = *Rattus rattus*; Dros M = *Drosophila melanogaster*; Dros Y = *Drosophila yakuba*

Table 4 shows from amino acid differences in mitochondrial protein COII that human beings have diverged from mouse and rats in this gene about three times as far as mouse and rats are separated from cattle, as noted previously by Brown and Simpson (1982); see especially their Figure 3. This difference in evolutionary divergence has never been suspected from examination of nuclear-DNA-coded proteins such as hemoglobins and cytochromes, and is seen only in amino acid replacements, not in silent nucleotide substitutions.

This difference in rate of divergence between human and other mammals does not conform with the molecular evolutionary clock. The difference has disappeared from the *Drosophila* replacement comparisons (Table 4), but it is present in the list of silent nucleotide substitutions (~158 vs ~120), from which it is seen that the "plateau" of silent substitutions in lines 2–10 and 12–19 of Table 4 does not represent "saturation" as proposed by Perler and co-workers (1980). Indeed, of the 255 potentially silent sites in each of the seven COII genes, it was found that 230 have been substituted in one or more species when the sequences are aligned (not shown). Clearly, nearly all of the silent sites are available for substitution.

Brown and co-workers, in their discussion of mitochondrial DNA, suggested that one site might undergo five substitutions in the accumulation of transversions. This must be very rare. Only when the total number of hits is 140 per 100 sites can 1% of the sites be expected to receive five hits, and at

Table 5. Comparison of amino acid differences, replacement substitutions, and silent substitutions in cytochrome *c* and mitochondrial cytochrome oxidase II, per 100 sites

Comparison	Amino acid differences	Replacement substitutions in codons	Potentially silent sites					
			In family boxes			In two-codon sets		
			Unsbs.	Trt.	Trv.	Unsbs.	Trt.	Trv.
Cytochrome *c*								
Mouse : rat	—	—	41	2	—	51	6	0
Mouse : Dros 4	15	23	21	7	25	28	17	1
Rat : Dros 4	15	23	19	8	26	27	20	1
Mitochondrial cytochrome oxidase II								
Mouse : NR	1	1	20	5	18	39	18	—
Mouse : Dros M	43	55	24	8	27	27	18	8
Mouse : Dros Y	42	58	26	8	27	27	16	8
NR : Dros M	42	55	22	8	30	24	20	7
NR : Dros Y	43	58	25	6	27	25	17	6

Unsbs. = unsubstituted; Trt. = transitions; Trv. = transversions; NR = *Rattus norvegicus*; Dros 4 = cytochrome *c* 4, *Drosophila melanogaster*; Dros M = *Drosophila melanogaster*; Dros Y = *Drosophila yakuba*

```
RT AGC GGA AGA GGC AAA ACC GGA AAG GCC AGG GCG AAG GCA AAG ACA CGT TCA TCC CGT GCC GGG CTC CAG TTC CCC GTG GGC CGT GTG CAC
SP TCT GGC AGA GGA AAG AGT GGA AAG GCC CGC ACC AAG GCA AAG ACG CGC TCA TCC CGT GCA GGG CTC CAG TTT CCA GTG GGA CGT GTT CAT
PM TCT GGA AGA GGT AAA AGT GGA AAG GCC CGT ACC AAG GCA AAA TCT CGT TCA TCC CGC GCT GGT CTC CAG TTC CCA GTG GGA CGT GTT CAC

RT AGG CTG CTG CGT AAA GGC AAC TAC GCC GAG CGT GTG GGC GCT GGC GCA CCA GTG TAC CTG GCC GCA GTG CTC GAG TAC CTG ACT GCT GAG
SP CGG TTT CTC CGA AAG GGC AAC TAT GCA AAG AGG GTC GGC GGT GGA GCT CCT GTC TAC ATG GCT GCC GTC CTA GAG TAC CTC ACT GCC GAA
PM CGA TTT CTA CGC AAA GGC AAC TAT GCA AAG AGG GTC GGC GGT GGG GCA CCA GTC TAC ATG GCC GCT GTC TTG GAG TAC TTG ACT GCC GAA

RT ATC CTG GAG TTG GCC GGA AAC GCT GCC CGT GAC AAC AAG AAG ACT CGT ATC ATC CCC CGT CAC CTG CAG CTG GCA GTC CGT AAC GAC GAG
SP ATC TTG GAA CTC GCG GGC AAC GCT GCC CGC GAC AAC AAG AAA TCT AGG ATC ATC CCA CGC CAC CTT CAA CTC GCT GTG CGT AAT GAT GAA
PM ATT CTC GAG CTC GCT GGC AAC GCT GCT CGC GAC AAC AAG AAA TCT AGG ATC ATT CCC CGT CAT CTT CAA CTT GCC GTG CGC AAC GAC GAA

RT GAG CTG AAC AAA CTG CTT GGC GGC GTC ACC ATC GCT CAG GGT GGT GTT CTG CCC AAC ATC CAG GCA GTG CTG CTC CCC AAG AAG ACT GAG
SP GAA CTC AAC AAG CTT TTG GGT GGG GTG ACG ATC GCT CAA GGT GGT GTT CTG CCC AAC ATC CAA GCC GTG CTG CTC CCC AAG AAA ACC GCT
PM GAA CTC AAC AAG CTC CTC GGA GGG GTG ACG ATC GCC CAA GGT GGT GTC CTG CCC AAC ATC CAA GCC GTG CTG CTT CCC AAG AAG ACA GGC

RT AAG GCC GTC AAA GCC AAG TAA
SP AAA TCA AGC TAG
PM AAA TCA AGC TAA
```

Fig. 2. Nucleotide sequences of histone 2A genes from rainbow trout (RT) and sea urchins *Strongylocentrotus purpuratus* (SP) and *Psammechinus miliaris* (PM)

Table 6. Changes in transitions and transversions during evolutionary divergence of cytochrome oxidase component II genes in mitochondria of *Rattus norvegicus, Rattus rattus,* and mouse

Changes	NR:RR vs M:NR	NR:RR vs M:RR
IC to TFB	11	11
IC to VFB	31	29
IC to T2CS	27	30
TFB to IC	4	3
VFB to IC	1	1
T2CS to IC	12	12
TFB to VFB	10	10
VFB to TFB	0	0
T2CS to VFB	1	0
Net changes		
TFB	−2	−2
VFB	42	39
T2CS	14	18

IC = identical codons; TFB = transitions in family-box sites; VFB = transversions in family-box sites; T2CS = transitions in two-codon sets; NR = *Rattus norvegicus*; RR = *Rattus rattus*; M = mouse

this point only 25% of the sites remain unhit as compared with 51% in the case of the human : rodent comparisons in Table 2. Obviously, their explanation does not account for the progressive increase in silent transversions in COII.

Table 5 compares rates of molecular evolution of cytochrome *c* and COII. Silent nucleotide substitutions in mouse versus rat for cytochrome *c* are 9 in 112 sites (8%), and for COII are 103 in 256 sites (41%), about a fivefold difference in rate, not including multiple hits.

Table 6 shows totals of changes in transitions and transversions at individual sites during early evolutionary separation of COII genes. Transversions have increased greatly, but about 75% of the increase has come from transversions in previously unchanged sites, rather than from conversion of family-box site transitions to transversions. Many (57) transitions occurred in two-codon sets, but 24 transitions reverted to identical base pairs in such sets (Table 6).

Histones

The amino acid sequences of histones are highly conserved in evolution, but silent nucleotide substitutions are quite frequent. Nucleotide sequences for histones 2A and 3 in two sea urchins (*Strongylocentrotus purpuratus* and *Psammechinus miliaris*) (Schaffner et al. 1978; Sures et al. 1978) and rainbow trout (Connor et al. 1984; Winkfein et al. 1985) are aligned in Figs. 2 and 3, and Table 7 compares amino acid replacements with nucleotide substitutions.

These sequences are good subjects for comparing silent transitions and transversions in earlier (sea urchins) and later (rainbow trout versus sea urchin) evolutionary divergences. In the comparisons between sea urchins (family-box sites), there are more transitions than transversions, but transversions

RT GCC AGA ACC AAG CAA ACC GCT CGC AAA TCC ACC GGT GGC AAA GCA CCC AGG AAG CAG CTC GCC ACC AAG GCT GCG CGC AAG AGC GCC CCG
SP GCA CGC ACC AAG CAG ACC GCT CGC AAA TCT ACA GGA GGG AAG GCT CCC CGC AAG CAG CTG GCA ACC AAA GCT GCC AGA AAG AGT GCC CCC
PM GCA CGA ACC AAG CAG ACG GCT CGC AAA TCT ACA GGA GGT AAA GCC CCC CGC AAG CAG CTG GCA ACC AAG GCT GCC AGA AAG AGT GCC CCA

RT GCC ACC GGC GGC GTG AAG AAG CCT CAC CGT TAC AGG CCC GGC ACC GTG GCT CTT AGA GAG ATC CGT CGT TAC CAG AAG TCC ACT GAG CTG
SP GCC ACT GGA GGA GTC AAG AAG CCT CAT CGA TAC AGG CCT GGC ACA GTC GCC CTG AGA GAG ATT CGC CGC TAC CAG AAG AGC ACT GAG CTT
PM GCC ACC GGA GGA GTA AAG AAG CCT CAT CGC TAC AGA CCT GGT ACA GTC GCA TTG AGA GAG ATT CGT CGC TAC CAG AAG AGC ACC GAA CTT

RT CTG ATC CGC AAA CTG CCT TTC CAG CGC CTG GTG CGA GAA ATT GCC CAG GAC TTT AAG ACT GAC CTG CGC TTC CAG AGT TCC GCA GTG ATG
SP CTC ATC CGA AAA CTG CCA TTC CAG CGT CTA GTG CGT GAG ATT GCA CAG GAC TTC AAG ACA GAG CTA CGT TTC CAG AGT TCC GCT GTG ATG
PM CTT ATC CGA AAG CTA CCC TTC CAG CGC CTG GTA CGT GAG ATC GCT CAG GAT TTC AAG ACC GAG CTA CGT TTC CAG AGC TCC GCT GTT ATG

RT GCC CTG CAG GAG GCA AGC GAG GCT TAC CTG GTC GGC CTG TTC GAG GAC ACC AAC CTG TGC GCC ATC CAC GCC AAG AGG GTG ACC ATC ATG
SP GCC CTT CAA GAA GCC AGC GAG GCA TAC CTA GTT GGC CXX XXX XGA GAC ACC AAC CTG TGT GCC ATC CAC GCC AAG AGG GTT ACC ATC ATG
PM GCC CTC CAA GAG GCA AGC GAG GCC TAC CTG GTC GGT CTG TTC GAG GAC ACC AAT CTG TGC GCC ATC CAC GCC AAG AGG GTA ACC ATC ATG

RT CCC AAG GAC ATC CAG CTG GCC CGT CGT ATT CGC GGA GAG CGC GCA TAA
SP CCC AAA GAC ATC CAG CTC GCC CGT CGA ATC CGC GGA GAA CGC GCC TAG
PM CCC AAG GAT ATC CAA CTC GCT CGT CGA ATC CGC GGA GAA CGT GCC TAG

Fig. 3. Nucleotide sequences of histone 3 genes from rainbow trout (RT) and sea urchins *Strongylocentrotus purpuratus* (SP) and *Psammechinus miliaris* (PM)

predominate in the trout : sea urchin comparisons (Table 8). Transitions have remained stationary or even decreased slightly in family boxes. One might infer that transitions have been converted to transversions, but an examination of the sequences shows that this is not the case (Table 9). The increases in transversions have resulted predominantly from transversions occurring in previously unchanged family-box sites. In histones 2A and 3, 45 unchanged sites produced transversions, while transversions have actually been replaced by transitions in 18 family-box sites. Several codon changes each resulted in two silent transversions, including CGN to AGR and UCN to AGY substitutions.

Could the increase in transversions have resulted from multiple substitutions at the same site? This can be examined by means of the Poisson distribution, using the second row of comparisons in Table 8 as an example. In the average of trout : sea urchin family-box comparisons, there are 22.2 identical codons in 55.7 family-box comparisons, corresponding to 40 unhit sites, and 88 acutal hits per 100 sites (Poisson). Of the 88 hits, 37 will be single substitutions and 22 multiple substitutions at the same site. If *all* multiple substitutions result in transversions, which is improbable, this would account for 22 transversions in 100 sites, corresponding to 12 in 55.7 sites, as compared with an actual increase of 9.4–26.8 = 17.4 transversions. Clearly, the increase in transversions cannot be explained by multiple changes at the same site.

Table 7. Amino acid replacements and nucleotide substitutions in three histones

Comparison	Amino acid replacements		Nucleotide substitutions	
	Histone 2A	Histone 3	Histone 2A	Histone 3
Sea urchins SP : PM	2	0	49	45
Sea urchin SP : rainbow trout	9	1	75	72
Sea urchin PM : rainbow trout	10	1	61	60

SP = *Strongylocentrotus purpuratus*; PM = *Psammechinus miliaris*; 123 codons are compared in histone 2A and 132 in histone 3

The almost complete absence of transversions (producing amino acid replacements) in two-codon-set sites evidently results from constraints against such replacements in the evolution of this group of histones. Indeed, histones are one of the proteins showing fewest evolutionary amino acid replacements.

Discussion

Transversion Rates

Four pairs of genes are compared in Table 10 with respect to types of evolutionary change. The pairs

were selected so that each pair had about 50 silent nucleotide substitutions per 100 codons. The rates of transversions to transitions in family-box sites, shown in Table 10, show a wide range. The numbers of transversions and transitions in these sites are about equal in the bacterial genes examined, but in the mitochondrial pair, the ratio is 2.5:1. Two-codon sets all have a low incidence of transversions.

The proportion of transitions and transversions is affected by the DNA polymerase system. Originally, Watson and Crick (1953) proposed that "spontaneous mutation may be due to a base occasionally occurring in one of its less likely tautomeric forms." Later, transversions A·T to C·G were shown by Cox and Yanofsky (1967) to be favored by the Treffers mutator (*mut T*) gene in *Escherichia coli*, while in bacteriophage T4, a gene may favor G·C to A·T transitions (Speyer 1965). Such genetic differences might easily be responsible for variations in proportion between silent transitions and transversions listed in Table 10.

Ochman and Wilson (1987) estimate divergence time of *Salmonella typhimurium* and *E. coli* as between 120 and 160 Myr. This is not inconsistent with the comparisons in Table 10, which show this divergence as being on the same order of magnitude as that for cytochrome *c* (rodent : chicken). However, no reference point for divergence of bacteria is available from the fossil record.

Transitions and transversions in two-codon sets have ratios that vary greatly from protein to protein. The controlling factor in this case is the extent to which amino acid replacements are accommodated in the evolution of the protein in question. As would be expected, the resistance to such replacements is greatest in histones, with only one replacement in a two-codon set—glutamic acid to the similar aspartic acid—occurring in the trout : sea urchin comparisons. The ratio of transitions to transversions in potentially silent positions of two-codon sets is an index of the extent to which the amino acid sequence of a protein is conserved in evolution. The fact that the ratio is always so much wider than the ratio in family boxes is because silent nucleotide substitutions are much closer to neutrality than are amino acid replacements (Jukes and Bhushan 1986).

Brown and co-workers (1982) studied transitions and transversions in a segment of mitochondrial DNA, 896 base pairs in length, from human, chimpanzee, gorilla, orangutan, and gibbon. They found the rate of nucleotide substitutions to be 5–10 times higher than in nuclear DNA. The silent substitution

Table 8. Categories of nucleotide substitutions per 100 potentially silent sites during evolutionary divergence of histone 2A and 3 genes in sea urchins and trout

	Family boxes			Two-codon sets			
						Trv.	
Comparison	Unsbs.	Trt.	Trv.	Unsbs.	Trt.	R	S
Histone 2A							
SP : PM	28.2	19.8	9.4	35.6	9.4	—	—
T : SP	22.2	6.7	26.8	28.2	14.1	—	2.7
T : PM	18.8	12.8	22.8	32.2	10.1	—	2.7
Histone 3							
SP : PM	30.7	9.6	8.4	44.0	8.4	—	—
T : SP	21.0	8.4	19.3	41.6	7.2	0.6	3.0
T : PM	21.0	12.0	15.7	39.8	9.6	0.6	2.4

Unsbs. = unsubstituted; Trt. = transitions; Trv. = transversions; SP = *Strongylocentrotus purpuratus*; PM = *Psammechinus miliaris*; T = rainbow trout; total of potentially silent sites in each comparison: 2A = 149; 3 = 166; R = replacement transversions; S = silent transversions in two-codon sets

Table 9. Changes in transitions and transversions during evolutionary divergence in histone 2A and 3 genes in sea urchins and trout

	Histone 2A		Histone 3	
Changes	SP : PM vs T : SP	SP : PM vs T : PM	SP : PM vs T : SP	SP : PM vs T : PM
IC to TFB	1	4	5	5
IC to VFB	22	25	21	24
IC to T2CS	11	11	9	9
TFB to IC	3	3	3	5
VFB to IC	1	3	2	1
T2CS to IC	3	7	6	5
TFB to VFB	7	8	5	3
VFB to TFB	3	6	9	9
Net changes				
TFB	−2	1	6	6
VFB	25	24	15	17
T2CS	8	4	3	4

IC = identical codons; TFB = transitions in family-box sites; VFB = transversions in family-box sites; T2CS = transitions in two-codon sets; T = rainbow trout; SP = *Strongylocentrotus purpuratus*; PM = *Psammechinus miliaris*

Table 10. Comparisons of genes for four proteins for amino acid replacements, transitions, and transversions, per 100 codons

Gene	Comparison	Amino acid differences	Silent nucleotide substitutions	Family boxes			Two-codon sites	
				Trt.	Trv.	Trt./Trt.	Trt.	Trv.
Cyt *c*	Rodent : chicken	7.7	40.4	10.6	14.4	1.36	15.4	1.0
COII	Rodent : bovine	9.0	51.4	8.8	22.3	2.53	20.3	1.6
TS	*E. coli* : *S. typh*	8.1	47.7	16.8	17.3	1.03	13.5	1.4
H2A+3	Trout : sea urchin	4.3	43.8	10.0	21.2	2.12	10.2	0.3

Trt. = transitions; Trv. = transversions; cyt *c* = cytochrome *c*; COII = cytochrome oxidase component II; TS = tryptophan synthase; *S. typh* = *Salmonella typhimurium*; H2A+3 = histones 2A and 3

rate in the protein-coding portion was "4 to 6 times higher than the replacement rate, indicating strong functional constraints at replacement sites." Observations on mitochondrial genes (Jukes and Bhushan 1986) confirm this conclusion.

Brown and co-workers (1982) state that transitions "greatly outnumber transversions," especially in the most closely related species. The percentage of transitions falls with increasing divergence. The authors postulate that "the evolutionary process is heavily biased towards transitions," but that "multiple substitutions occur at the same rate when the time is long, obscuring the record of transitions." They illustrate this point by hypothesizing "10 point mutations at the same site along two diverging lineages."

However, they do not note that for only one site in a sequence of 1000 nucleotides to be "hit" 10 times, the sequence would have to absorb 3000 hits. Only 50 of the sites would not be substituted at least once. Comparisons of mitochondrial genes do not show such a high rate of substitution in silent sites.

This question was explored further by comparing the COII sequences of the two rats and mouse (M). A net increase of transversions is found in comparing NR/RR substitutions with NR/M substitutions (Table 6). Only nine of these result from conversions of transitions to transversions. The remaining 27 are transversions occurring in previously unsubstituted sites. There is a net increase of transitions produced by an increase of 13 in two-codon sets, all in previously unsubstituted sites, and a decrease of one transition in family boxes. Clearly the increase in transversions is not, for the most part, caused by conversion of transitions to transversions.

Recently, DeSalle and co-workers (1987) have stated that transitions accumulate about 10 times faster in mitochondrial DNA than in nuclear DNA, "but there is no significant rise in the ratio at which transversions accumulate." They suggest that vertebrate mitochondria may be deficient in the DNA repair system that eliminates transitions.

In marked contrast, Clary and Wolstenholme (1987) found that transversions were more common than transitions in comparing various regions of mitochondrial DNA sequences in *Drosophila virilis* and *D. yakuba*. In 117 nucleotide substitutions, occurring in 1577 total nucleotides compared, there were 40 transitions and 77 transversions; very close to the 1:2 ratio expected for random substitutions.

Li and associates (1985) classified nucleotide sites in homologous genes as nondegenerate, twofold degenerate, or fourfold degenerate, and nucleotide changes as either transitional or transversional. They found that the rate of nucleotide substitution was lowest at nondegenerate sites (i.e., replacement sites), intermediate at twofold sites (i.e., potentially silent sites in two-codon sets), and highest at fourfold degenerate sites (i.e., third-position sites in family boxes).

Table 2 in the paper by Li and co-workers (1985) has a compilation of nucleotide substitutions in the preceding three categories in 42 animal protein genes from 23 to 580 amino acids. Transitions at twofold degenerate sites per site per 10 years range from 0.43 to 3.79; transitions at fourfold degenerate sites fluctuate considerably, most are in the 1–4 range, and transversions at these sites are a bit lower, ranging from 0.37 to 3.19.

The authors point out that at twofold sites "the lower rate of transversional substitutions occurs because all transversional changes at twofold degenerate sites are nonsynonymous."

Perler and associates (1980) have compared the DNA sequences of rat, human, and chicken preproinsulin genes, and of human, rabbit, mouse, and chicken globin genes. They have discussed the comparisons in terms of nucleotide substitutions at replacement sites and silent sites. They did not divide the silent substitutions into family-box sites and two-codon-set sites. Perler and colleagues contrasted amino acid replacements "in terms of the neutralist position" in a comparison with the provisos of the selection hypothesis. These authors stated that the neutralist viewpoint "argues that some 99% of the changes observed in amino acid sequences are neutral changes which have no effect on the function of the protein, fixed in the population by random drift."

They said, correctly, that the selectionist hypothesis is that "the observed changes in protein structure have spread through the population by selection."

From these assumptions, Perler and co-workers (1980) proceed to a discussion of initial and subsequent rates of silent substitutions, saturation of silent changes, and the "clock provided by replacement sites." They "conclude that the evolutionary clock does not involve neutral changes occurring at a small number of replacement sites."

A simpler interpretation is readily available, once it is understood that there are various degrees of approximation of neutrality (Ohta 1973). It can be seen in Table 10 that silent nucleotide substitutions accumulate from 5 to 12 times more rapidly than amino acid replacements, and that amino acid replacements accumulate with time, even in the case of highly conserved proteins, such as histones.

Data in the tables are in their original state, not recalculated and revised for multiple hits according to either Eq. (2) or to the random evolutionary hit procedure (Holmquist et al. 1972; Jukes and Holmquist 1972b). Such recalculations would produce quantitative rather than qualitative changes, without altering the conclusions drawn.

The molecular evolutionary clock tends to persist despite great variations in its rate of keeping time, and despite criticisms by those who declare it to be nonexistent. The method of taxonomic classification of living organisms, started by Linnaeus in 1736, led to the use of phylogeny as a framework for evolution. As a result of the study of visible phenotypic characteristics, anatomical and physiological pathways of evolution have been traced, leading from common ancestors to a diversity of phyla, genera, and species. With the advent of protein sequencing, and, later, DNA sequencing, phylogeny was reexamined on the basis of the molecular evolutionary clock. The coincidence of the two unrelated approaches, each confirming the other, is one of the most remarkable findings in the history of evolution.

Acknowledgments. I thank Carol Fegté for editorial work, Vikas Bhushan for helpful discussions, and Drs. Richard Holmquist and Motoo Kimura for valuable suggestions. This work was supported by NASA grant 05-003-460.

References

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Anderson S, de Bruijn MHL, Coulson AR, Eperon IC, Sanger F, Young IG (1982) The complete sequence of bovine mitochondrial DNA: conserved features of the mammalian mitochondrial genome. J Mol Biol 155:683–717

Bibb MJ, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and gene organization of mouse mitochondrial DNA. Cell 26:167–180

Brown GG, Simpson MV (1982) Novel features of animal mtDNA evolution as shown by sequences of two rat cytochrome oxidase subunit II genes. Proc Natl Acad Sci USA 79:3246–3250

Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences in primates: tempo and mode of evolution. J Mol Evol 18:225–239

Catzeflis FM, Sheldon FH, Ahlquist JH, Sibley C (1987) DNA–DNA hybridization enhancement of the rapid rate of muroid–rodent DNA evolution. Mol Biol Evol, vol 4

Clary DO, Wolstenholme DR (1985) The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J Mol Evol 22:252–271

Clary DO, Wolstenholme DR (1987) *Drosophila* mitochondrial DNA: conserved sequences in the A + T-rich region and supporting evidence for a secondary structure model of the small ribosomal RNA. J Mol Evol 25:116–125

Connor W, States JC, Mezquita J, Dixon GH (1984) Organization and nucleotide sequence of rainbow trout histone H2A and H3 genes. J Mol Evol 20:236–250

Cox EC, Yanofsky C (1967) Altered base ratios in the DNA of an *Escherichia coli* mutator strain. Proc Natl Acad Sci USA 58:1895–1902

de Bruijn MHL (1983) *Drosophila melanogaster* mitochondrial DNA, a novel organization and genetic code. Nature 304: 234–241

DeSalle R, Freedman T, Prager EM, Wilson A (1987) Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J Mol Evol 26:157–164

Dickerson RE (1971) The structure of cytochrome *c* and the rates of molecular evolution. J Mol Evol 1:26–45

Dickerson RE, Timkovich R (1974) Cytochromes *c*. In: Boyer PD, Lardy H, Myrbäck K (eds) The enzymes, vol 18. Academic Press, New York, pp 1–280

Edwards v. Aguillard (1985) Brief for amicus curiae. The National Academy of Sciences Urging Affirmance No. 85-1512, US (5th Cir. 1985)

Hawk PB[1] (1907) Practical physiological chemistry. Blakiston, New York

Holmquist R (1983) Transitions and transversions in evolutionary descent: an approach to understanding. J Mol Evol 19:134

Holmquist R, Cantor C, Jukes T (1972) Improved procedures for comparing homologous sequences in molecules of proteins and nucleic acids. J Mol Biol 64:145–161

Ingram VM (1961) Gene evolution and the hemoglobins. Nature 189:704–709

[1] In the preface to the first edition, Hawk stated: "For the microphotographs of oxyhaemoglobin and haemin . . . the author is indebted to Professor E.T. Reichert of the University of Pennsylvania, who, in collaboration with Professor A.P. Brown is making a very extended investigation into the crystalline forms of biochemic substances." In the 13th edition (1954, p. 471) he stated: "Oxyhemoglobin may be crystallized and a specific form of crystal is obtained from the blood of each individual animal species. . . . Reichert and Brown studied oxyhemoglobin crystals prepared from the blood of more than 100 species of animals from the point of view of their crystallographic characteristics. . . . Since hemin crystals are identical no matter from what species the blood is obtained, the species differences in oxyhemoglobins must presumably be related to differences in the globin portion of the molecule."

Jukes TH (1982) Silent nucleotide substitutions in evolution. Paper presented at the Meeting of the Society for Study of Evolution, Stony Brook NY, June 23, 1982

Jukes TH, Bhushan V (1986) Silent nucleotide substitutions and G+C content of some mitochondrial and bacterial genes. J Mol Evol 24:39–44

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism, vol 3. Academic Press, New York, pp 21–132

Jukes TH, Holmquist R (1972a) Evolutionary clock: nonconstancy of rate in different species. Science 177:530–532

Jukes TH, Holmquist R (1972b) Estimation of evolutionary changes in certain homologous polypeptide chains. J Mol Biol 64:163–179

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Li WH, Wu CI, Luo CC (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

Limbach KJ, Wu R (1983) Isolation and characterization of two alleles of the chicken cytochrome *c* gene. Nucleic Acids Res 11:8931–8950

Limbach KJ, Wu R (1985a) Characterization of a mouse somatic cytochrome *c* gene and three cytochrome *c* pseudogenes. Nucleic Acids Res 13:617–630

Limbach KJ, Wu R (1985b) Characterization of two *Drosophila melanogaster* cytochrome *c* genes and their transcripts. Nucleic Acids Res 13:631–642

Montgomery DL, Leung DW, Smith M, Shalit P, Faye G, Hall BD (1980) Isolation and sequence of the gene for iso-2-cytochrome *c* in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 77:541–545

Nuttall GHF (1904) Blood immunity and blood relationship. Cambridge University Press, London

Ochman H, Wilson AC (1987) A universal substitution rate: evidence from bacteria. J Mol Evol 24:000–000

Ohta T (1973) Slightly deleterious mutant substitutions in evolution. Nature 246:96–98

Perler F, Efstratiadis A, Lomedico P, Gilbert W, Kolodner R, Dodgson J (1980) The evolution of genes: the chicken preproinsulin gene. Cell 20:555–566

Reichert ET, Brown AP (1907) Crystallography of hemoglobins from various species. Proc Soc Exp Biol Med 5:66–75

Reichert ET, Brown AP (1909) The differentiation and specificity of corresponding proteins and other vital substances in relation to biological classification and organic evolution: the crystallography of hemoglobins. Carnegie Institution, Washington DC

Sanger F, Thompson EOP, Tuppy H (1952) The structure of insulin. Symposium sur les hormones proteiques and derives des proteins. 2nd Int Congr Biochem, Paris, France, p 26

Scarpulla RC, Agne KM, Wu RJ (1981) Isolation and structure of a rat cytochrome *c* gene. J Biol Chem 256:6480–6486

Schaffner W, Kunz G, Daetwyler H, Telford J, Smith HO, Birnstiel ML (1978) Genes and spacers of cloned sea urchin histone DNA analyzed by sequencing. Cell 14:655–671

Smith M, Leung DW, Gillam S, Astell CR (1979) Sequence of the gene for iso-1-cytochrome c in Saccharomyces cerevisiae. Cell 16:753–761

Speyer J (1965) Mutagenic DNA polymerase. Biochem Biophys Res Commun 21:6–10

Sures I, Lowry J, Kedes L (1978) The DNA sequence of sea urchin (*S. purpuratus*) H2A, H2B and H3 histone coding and spacer regions. Cell 15:1033–1044

Watson J, Crick FHC (1953) Genetical implications of the structure of deoxyribonucleic acid. Nature 171:964–967

Winkfein RJ, Connor W, Mezquita J, Dixon GH (1985) Histone H4 and H2B genes in rainbow trout (*Salmo gairdneri*). J Mol Evol 22:1–19

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, New York, pp 97–166

Received February 2, 1987/Revised March 20, 1987