

# EVOLUTION OF MAMMALIAN DNA

D. E. KOHNE  
CARNEGIE INSTITUTION OF WASHINGTON  
J. A. CHISCON  
PURDUE UNIVERSITY  
and  
B. H. HOYER  
CARNEGIE INSTITUTION OF WASHINGTON

## 1. Introduction and rationale behind experiments

This communication describes the measurement of DNA nucleotide sequence changes which have occurred since the divergence of various primates. These measurements, coupled with paleontological evidence for times of divergence of primates, allow an estimation of the rate of nucleotide sequence change during different periods of primate evolution. These data strongly suggest the possibility that the rate of nucleotide sequence change may have been a function of the generation times of the species involved. The rate appears to be faster in species with short generation times. This suggests that the rate of molecular evolution has not been constant through evolutionary time.

A great deal of information about evolutionary events and processes has been inferred from careful studies of fossil records. Early DNA studies [1], [4], [11] generally corroborated classical evolutionary findings, and at the same time provided some new understanding of molecular processes in evolution. For the most part, the early work with DNA simply measured the quantities of DNA which would reassociate when mixtures of DNA from various species were allowed to react. Repeated sequences had not been recognized and their significance was not appreciated [2]. At present it is possible to separate the repeated sequences from the nonrepeated sequences and to measure relationships between species using either fraction.

DNA reassociation is the primary tool used for exploring the evolution of DNA. Since a wide variety of scientists are interested in evolution, an attempt has been made to provide enough background material to be helpful to those with little knowledge of DNA. Readers well versed in DNA lore could well ignore this material and skip to another section.

1.1. *Basic characteristics of DNA.* DNA is a long linear polymer constructed of four distinct chemical subunits called nucleotides. These nucleotides differ in their bases, containing either Adenine (A), Guanine (G), Cytosine (C), or Thymine (T). The biological information of the DNA molecule is stored in the ordered sequence of its bases. In its natural state DNA exists as a double strand

molecule with each strand winding around the other in a helical fashion. The two strands are held together by specific interactions between the individual nucleotides in each strand. A always interacts or pairs with T while G and C always pair together. A-T and G-C are called complementary base pairs. In DNA whose strands have never been separated (native DNA) each base in one strand is believed to be paired with its complementary base in the other. Perfect base pair matching is then presumed to be present between the strands. The complementary strands of a double strand DNA molecule are readily separated or dissociated into two single DNA strands. Under the proper conditions, these complementary DNA strands can reassociate and reform a stable double strand DNA molecule. This "sequence recognition" property of complementary DNA strands is extremely useful for studying the evolution of DNA, and is utilized to compare the genetic material of different organisms [4], [11].

Reassociation of DNA can be measured in several ways, all of which depend upon some physical chemical differences between dissociated (single strand DNA) and reassociated (double strand DNA). A particularly useful technique for measuring DNA reassociation utilizes the properties of hydroxyapatite crystals. Under the proper conditions hydroxyapatite will absorb double strand or reassociated DNA while single strand DNA does not absorb to the hydroxyapatite and can be washed away [2], [6].

1.2. *Thermal stability of DNA.* The thermal stability of double strand DNA can be used as an indicator of the degree of perfection of base pair matching between the strands of a double strand DNA molecule. Double strand DNA heated in solution will dissociate into single strands at a specific temperature, depending on the salt concentration and DNA composition. The thermal stability of reassociated DNA can be compared to that of native DNA which has perfect, or nearly perfect, matching between the bases of its component strands. Reassociated DNA can have perfection of base pair matching similar to that of native DNA. It is possible to alter nucleic acid strands so that a small fraction of the nucleotides can no longer interact with a complementary nucleotide. These partially complementary strands can then reassociate to form stable double strand molecules whose thermal stability is less than that of reassociated, unaltered molecules. Lack of complementarity in 1.5 per cent of the nucleotides will lower the thermal stability of the double strand molecule by about 1°C [10].

1.3. *Rationale of experiments.* To determine the rate of nucleotide sequence change since the divergence of various species, one must know (a) the time since the divergence of the species in question, and (b) the number of nucleotide changes which have occurred during that time. Paleontology provides us with estimates of (a), and the number of nucleotide changes can be experimentally determined. Complementary single DNA strands from different animal species can interact to form a "hybrid" double strand molecule. One strand of this "hybrid" molecule is radioactive and is from one species while the other is non-radioactive and from a different species. If the two strands are only partially complementary in the reassociated region, the "hybrid" reassociated DNA

will have a lower thermal stability than perfectly base pair matched DNA. The difference in thermal stability between the "homologous" and "hybrid" DNA is a measure of the extent of nucleotide changes which have occurred since divergence of the two species in question. A reasonable estimate of the actual percentage of nucleotide changes can be obtained by using the experimentally determined observation that 1.5 per cent nucleotide pair mismatches lowers the thermal stability 1°C. Figure 1 presents the rationale used for determining the extent of nucleotide substitutions which have occurred since the divergence of the two species.

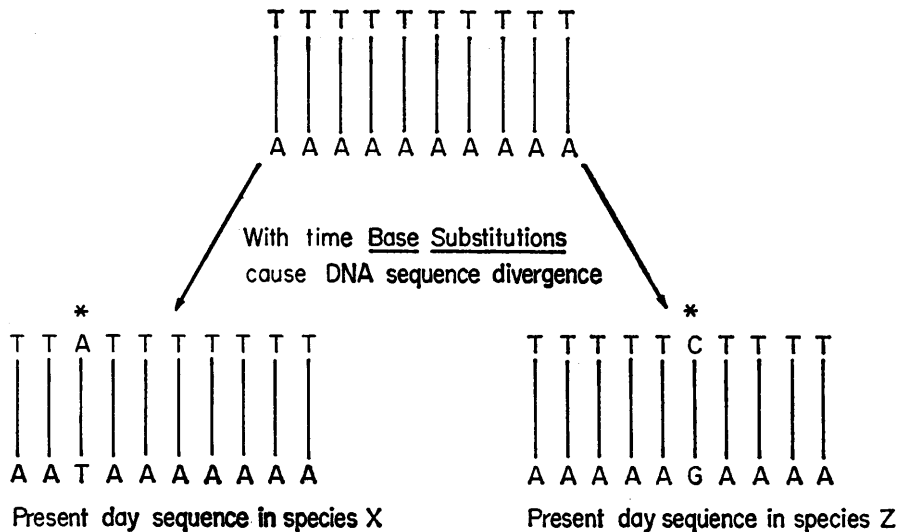
The value for per cent base changes which have occurred since divergence is an average divergence value for the many different nucleotide sequences present in the DNA used for the experiment. Furthermore, multiple changes have undoubtedly occurred at many sites in extensively diverged DNA's (for example, man *versus* galago). However, only those nucleotide changes existing today can be detected, so the extent of change seen between distantly related species are minimum values. Values obtained for closely related species should be very nearly correct since the probability of multiple changes occurring at any one site will be low.

A primary requirement for comparing the DNA's of different species is that one species of DNA forms predominately "hybrid" double strand DNA molecules during the reassociation period. This requirement can be met by properly adjusting the radioactive and nonradioactive DNA concentrations to control the rates of reassociation. The radioactive DNA used must be kept at a low enough concentration so that the collision of two radioactive complementary strands is very improbable. The nonradioactive DNA concentration must be high enough so that virtually all of the nonradioactive DNA sequences will find complementary sequences and reassociate during the time of incubation. Thus, a radioactive DNA strand will form a stable double strand "hybrid" molecule *if and only if there is a complementary sequence for it in the nonradioactive DNA*.

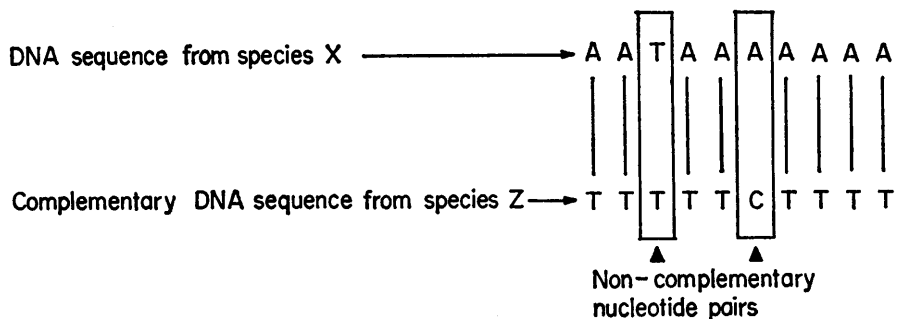
Not all of the DNA of higher organisms can be used for these experiments. All higher organisms examined thus far have contained large fractions or families of repeated DNA sequences (Table I) [2]. A family of repeated nucleotide sequences is composed of many member sequences, each of which can reassociate with any other member of that family. There are at least several families in each cell and the member sequences of each family are generally not identical to one another in sequence, but are similar enough to reassociate together. It is known that related species contain at least some of the same families of repeated DNA sequences, since the repeated DNA of one species will reassociate with that of another species [1], [4]. If two related species contain similar repeated DNA, it seems probable that they both inherited this repeated DNA from their most recent common ancestor. Since it is not known whether these particular family DNA sequences were identical or only similar at the time of divergence, it is not possible to determine how long it took to accumulate the

I. Original DNA sequence in the most recent common ancestor of species X and Z.

The asterisk indicates nucleotide substitution.



II. The DNA SEQUENCES from species X and Z can be compared in a test tube by forming a *HYBRID DOUBLE STRAND* DNA molecule. Single strand DNA sequences do not have to be perfectly complementary in order to reassociate to form a stable double strand hybrid molecule.



III. Double strand DNA is most *Thermal Stable* when all of the nucleotides in one strand are properly paired with their complementary nucleotide.

When all of the nucleotides in the double strand region are *not* paired their complementary nucleotide, the thermal stability is lowered.

*1.5 per cent noncomplementary nucleotide pairs = 1°C lowering of thermal stability.*

IV. The per cent of noncomplementary nucleotide pairs seen for a particular hybrid represents the extent of nucleotide sequence change which has occurred since the time of divergence of the species involved.

FIGURE 1

Determination of the extent of nucleotide substitutions which have occurred since the divergence of two species.

TABLE I  
CHARACTERISTICS OF MAMMALIAN DNA

---



---

A. Repeated DNA

- (1) Repeated DNA is composed of groups or families of DNA sequences. *Each* member of a family can interact and reassociate with *any other* member of that family.
- (2) There are at least several families in each cell. The number of members per family is about  $10^6$  for mammals.
- (3) Family members are generally similar but not identical to each other. There are exceptions to this.
- (4) Repeated DNA reassociates *very* rapidly because of the high concentration of each sequence in each cell. It is easily purified.

B. Nonrepeated DNA

- (1) Each nonrepeated sequence is present one time per haploid cell.
  - (2) Nonrepeated DNA sequences reassociate *very* slowly because of their low concentration per cell. They are readily purified.
  - (3) *Most* of the potential genetic information is contained in this fraction.
- 

number of differences observed in the "hybrid" reassociated repeated DNA. Thus, the comparison of DNA in its repeated form will not yield the true rate of nucleotide sequence change when correlated with the time since divergence.

## 2. Extents of nucleotide change since the divergence of various mammals

The basic experiments involve reassociating radioactive nonrepeated DNA from various animals with total nonradioactive DNA prepared from the same or from other organisms (Table II). Radioactive nonrepeated DNA's was obtained by reassociating the radioactive DNA (human, green monkey, mouse) for a sufficient time to reassociate the repeated DNA sequences. The DNA was then fractionated on hydroxyapatite and the nonrepeated fraction isolated [2], [6].

TABLE II  
PROCEDURE FOR COMPARING MAN DNA SEQUENCES WITH CHIMP DNA SEQUENCES

- 
- 
- I. Isolate man radioactive nonrepeated DNA sequences. There are about  $5 \times 10^9$  nucleotides per cell in this fraction. All mammals have about the same amount of DNA per cell. DNA used in these experiments has been sheared to *400 nucleotide long pieces*.
  - II. A. Mix a small amount of human *radioactive nonrepeated* DNA with a large amount of *nonradioactive* DNA from the Chimp. Denature and incubate.  
B. For the control, a small amount of *human radioactive nonrepeated* DNA is mixed with a large amount of *human nonradioactive* DNA. Denature and incubate. In this case the radioactive DNA will form perfectly nucleotide pair matched duplexes.
  - III. Isolate both the hybrid [Man-Chimp (A)] and control [Man-Man (B)] double strand molecules and determine their thermal stabilities (T.S.):  
 $\Delta \text{T.S.} = [\text{T.S. of Man-Man duplex}] - [\text{T.S. of Man-Chimp hybrid}]$ .
  - IV. Per cent noncomplementary nucleotide pairs (or per cent nucleotide substitutions since the divergence of Man and Chimp)

$$= [\Delta \text{T.S. } (^{\circ}\text{C})] \left( \frac{1.5\% \text{ noncomplementary nucleotide pairs}}{1^{\circ}\text{C lowering of T.S.}} \right).$$


---

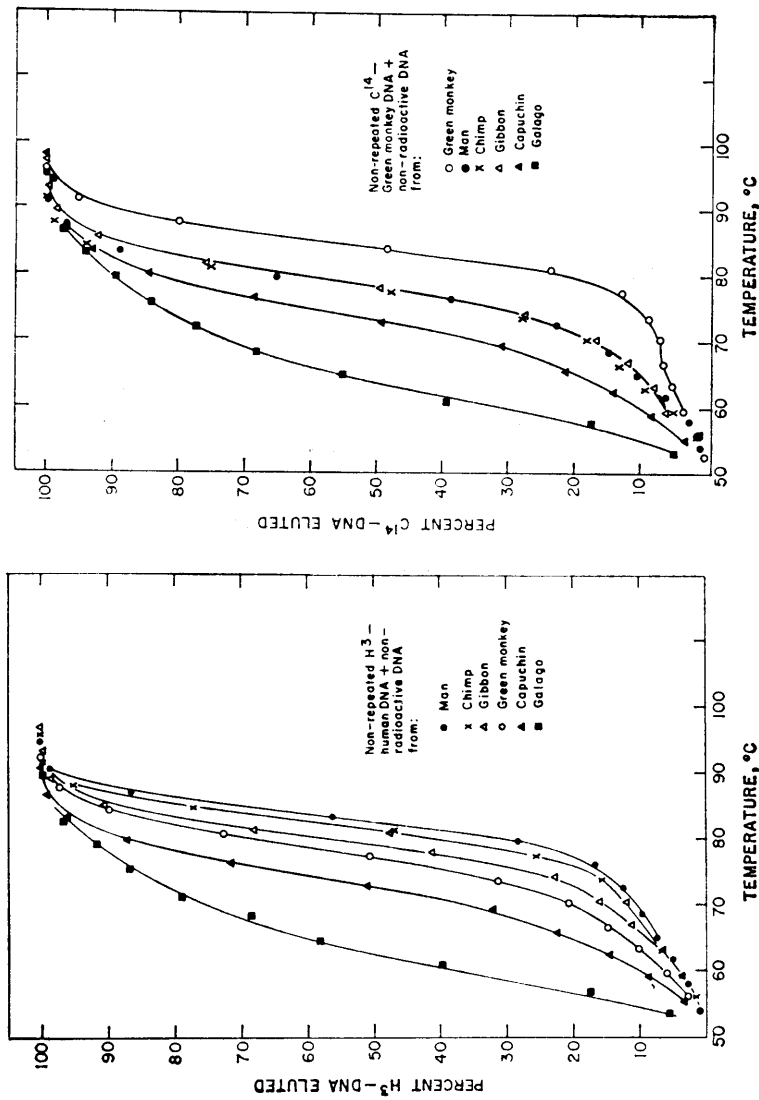


FIGURE 2

Thermal stability profiles.

On the left is shown per cent H<sup>3</sup>-DNA eluted against temperature, and on the right per cent C<sup>14</sup>-DNA eluted plotted against temperature. The basic experimental procedure followed for these studies is given in the text.

TABLE III

EXTENT OF NUCLEOTIDE SEQUENCE DIFFERENCE BETWEEN VARIOUS SPECIES

DNA compared	$\Delta$ T.S. ( $^{\circ}$ C) (Control T.S. - hybrid T.S.)	Per cent nucleotide substitutions observed since the time of divergence of the two species compared
Man-Chimp	1.7	2.5
Man-Gibbon	3.4	5.1
Man-Green monkey	6.1	9
Man- <i>Rhesus</i>	5.5	8.3
Man-Capuchin	10.5	15.8
Man-Galago	$\sim$ 28	42
Mouse-Rat	20	30
Cow-Sheep	7.5	11.2
Green monkey-Man	6	9
Green monkey-Chimp	6.2	9.3
Green monkey-Gibbon	6	9
Green monkey-Capuchin	11.5	17.3
Green monkey-Galago	$\sim$ 28	$\sim$ 42

Figure 2 presents thermal stability profiles obtained in these experiments, while Table III summarizes the data. The stability profiles of the hybrids show several qualitative features of significance. Almost all of the human DNA reacted with most of the species. It appears that no large class of DNA has diverged so rapidly that hybrids cannot be formed at this temperature of incubation. At the other extreme, almost all of the DNA has diverged appreciably, so there is very little of a highly conserved class of DNA sequences (for example, the man-galago comparison).

The basic experimental procedure followed for these studies was: mix a large amount of nonradioactive DNA (2 to 4 milligrams in less than one milliliter) with a small amount of nonrepeated radioactive DNA (1 to 2 micrograms). Add EDTA (pH = 8.0) to a final concentration of 2.5 millimolar and dissociate the DNA by placing the vial in a boiling water bath for one minute. EDTA helps prevent degradation of the DNA during incubation. Care should be taken to heat the DNA for as short a time as possible when the DNA is in a solution of a low salt concentration. After heating, make the DNA to 0.8 M PB (1.2 M Na<sup>+</sup>) and incubate the mixture at 63 $^{\circ}$ C. The DNA was incubated to a DNA  $C_{0t}$  (the product of the molar concentration of DNA monomers and time in seconds) equivalent to 10,000 in 0.14 M PB 51 $^{\circ}$ C (at 2 mg/ml of nonradioactive DNA in 0.8 M PB, where PB is a sodium phosphate buffer solution. This involves incubating the DNA for about 90 hours at 63 $^{\circ}$ C). The final volume was one milliliter. At the end of the incubation period, freeze the samples for later

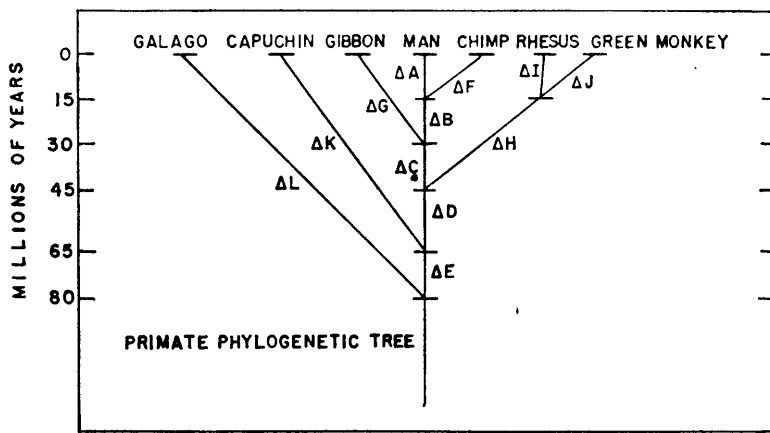
analysis. For analysis take an aliquot of the sample (about 400 micrograms of DNA) and adjust the salt concentration to 0.14 M PB and pass this sample over a water jacketed hydroxyapatite (two grams of Bio Rad HTP) column equilibrated to 51°C and 0.14 M PB. Wash the unreassociated DNA off the column. The thermal stability of these samples was measured by raising the temperature of the column automatically (0.9°C rise in 2.5 minutes) and collecting, with a fraction collector, the effluent (0.14 M PB) which pumped through the column at about five milliliters per minute. Each point on Figure 2 represents 40 milliliters of effluent (0.14 M PB) and a 1.8°C temperature rise. After collection, precipitate each fraction, collect on a filter, and assay for radioactivity with a scintillation spectrometer. In order to gain maximum reproducibility, the thermal elution procedure was completely automated. For the sake of efficiency, some of the incubation mixtures contained nonrepeated DNA from two different species (H<sup>3</sup>-man DNA and C<sup>14</sup>-green monkey DNA). These DNA's were at such low concentration that they were unable to react significantly with themselves (that is, H<sup>3</sup>-DNA with H<sup>3</sup>-DNA).

In comparing the divergence of different pairs of species, the same classes of DNA's must be compared. For the purposes of the calculations and considerations to follow, we have compared the average divergence of *all* the nonrepeated DNA. The measure of the average divergence is the temperature at which 50 per cent of the *total* radioactive DNA is in a hybrid form. This temperature is referred to as the thermal stability of the hybrid molecules. Two parameters are used to calculate the thermal stability of a hybrid pair: the extent of reaction of the radioactive DNA with the nonradioactive DNA, and the temperature elution profile of the hybrids formed (Figure 2).

2.1. *Interpretation of base substitution data.* Both the extent of reassociation between the various primate DNA's and the thermal stability measurements agree with the paleontological view of the relationships between the primates tested [6]. As expected, for example, chimp DNA is most closely related to man DNA, and *Rhesus* DNA to green monkey DNA. The nucleotide substitutions studied here are those which have been "fixed" during the evolution of these species. Further, the substitution values for the more distantly related species are likely to be underestimates due to multiple changes at the same site, as was discussed earlier.

Figure 3 depicts the present day view of the phylogenetic relationships among the primates studied here. Chimp and man, according to paleontological estimates, diverged from their most recent common ancestors about 10 to 20 million years ago. The extent of divergence of the man and chimp DNA's is, then, the sum of changes occurring since divergence in the chimp line plus those occurring in the human line. The difference seen between chimp and human DNA's today (2.5 per cent) is equal to  $\Delta A$  (the per cent nucleotide changes occurring in the human line) plus  $\Delta F$  (the per cent changes occurring in the chimp line). Using the data listed on Figure 3, it is possible to determine how many changes occurred in each line. The rationale for this is as follows. Form hybrids between radio-





<i>DNA's Compared</i>	<i>Per Cent Changes Since Divergence</i>
Man*-Chimp	$\Delta A + \Delta F = 2.5$
Man*-Gibbon	$\Delta A + \Delta B + \Delta G = 5.1$
Man*-Green monkey	$\Delta A + \Delta B + \Delta C + \Delta H + \Delta J = 9$
Man*-Rhesus	$\Delta A + \Delta B + \Delta C + \Delta H + \Delta I = 8.3$
Man*-Capuchin	$\Delta A + \Delta B + \Delta C + \Delta D + \Delta K = 15.8$
Man*-Galago	$\Delta A + \Delta B + \Delta C + \Delta D + \Delta E + \Delta L \approx 42$
GM*-Rhesus	$\Delta J + \Delta I = 2.9$
GM*-Man	$\Delta J + \Delta H + \Delta C + \Delta B + \Delta A = 9$
GM*-Chimp	$\Delta J + \Delta H + \Delta C + \Delta B + \Delta F = 9.3$
GM*-Gibbon	$\Delta J + \Delta H + \Delta C + \Delta G = 9$
GM*-Capuchin	$\Delta J + \Delta H + \Delta D + \Delta K = 17.3$
GM*-Galago	$\Delta J + \Delta H + \Delta D + \Delta E + \Delta L \approx 42$

FIGURE 3

Present day view of the phylogenetic relationships among several primates.

active green monkey nonrepeated DNA and man and chimp DNA's and then determine their thermal stabilities. The extent of difference between chimp and man DNA's has been measured to equal  $\Delta A + \Delta F = 2.5$  per cent. If all of the changes had occurred in the chimp line,  $\Delta F = 2.5$  per cent, while  $\Delta A = 0$ . In this case the man-green monkey hybrid difference would be  $\Delta A + \Delta B + \Delta C + \Delta H + \Delta J = X + 2.5$ . If the same number of changes had occurred in both the chimp  $\Delta F$  and human  $\Delta A$ , then  $\Delta A + \Delta B + \Delta C + \Delta H + \Delta J = \Delta F + \Delta B + \Delta C + \Delta H + \Delta J$ . Analysis of this type allows the assignment of rough values of extent of change for the various periods during the primate evolution. Table IV presents these values. The data used in these analyses is not yet good enough to precisely determine the values for each period. However, it is good enough so that large differences between values are meaningful. The

TABLE IV  
 RATES OF CHANGE DURING DIFFERENT EVOLUTIONARY ERAS  
 $\Delta D$ ,  $\Delta E$ ,  $\Delta K$ , and  $\Delta L$  are estimated.

Per cent nucleotide change during period	Length of period $\times 10^6$ years	Per cent change per $10^6$ years	
$\Delta A$	1.1	15	0.074
$\Delta B$	1.45	15	0.097
$\Delta C$	1.20	15	0.08
$\Delta D$	$\sim 4.45$	20	0.22
$\Delta E$	12.8	15	0.85
$\Delta F$	1.4	15	0.093
$\Delta G$	2.55	30	0.085
$\Delta H$	3.45	30	0.12
$\Delta I$	1.1	15	0.073
$\Delta J$	1.8	15	0.12
$\Delta K$	$\sim 8.2$	65	0.13
$\Delta L$	$\sim 21$	80	0.26

values of  $\Delta A$ ,  $\Delta B$ ,  $\Delta C$ ,  $\Delta F$ ,  $\Delta G$ ,  $\Delta H$ ,  $\Delta I$ ,  $\Delta J$  can be directly determined, while the values for  $\Delta D$ ,  $\Delta E$ ,  $\Delta K$ ,  $\Delta L$  can be estimated by assuming that after divergence the number of changes in each line was roughly the same. This, clearly, is the case in those periods where definite data is available. For example, after the divergence of human and green monkey the extent of change in the human line and the green monkey line is roughly the same. A similar situation exists for the human-chimp and human-gibbon lines. The estimated value for  $\Delta E$  is undoubtedly low due to multiple changes which have occurred at many of the nucleotide sites.

2.2. *Rate of nucleotide sequence change.* The average rate of nucleotide sequence change since the time of divergence of two species can be calculated from the data of Table III and estimates of the time since divergence of two species. Divergence times have been inferred from the fossil record. There is, however, by no means unanimous agreement among paleontologists on the assignment of divergence times from this record. The divergence times used in Table V are generally considered to be realistic values in the light of present knowledge [12], [13] (p. 208), [14], [15].

Table V presents the calculated average rates of change since the divergence of the various primate species. Table IV presents rates of change calculated for the different time periods during the development of the primate line. The absolute values for the rates calculated in Table IV are probably not very accurate. Much better data is needed to generate precisely accurate rate values by the method of analysis used. The data appear sufficiently good, however, to support the changing rate trend seen in Table IV. It appears that the rate of change was much higher during the early part of the primate line than during recent primate evolution.

Table V also presents rate of change data for rodent (rat *versus* mouse)

TABLE V

## RATES OF NUCLEOTIDE FIXATION DURING EVOLUTION

Estimates of time since divergence are from paleontological studies [12], [13], [14], [15].  
The last column gives average rates of nucleotide substitution.

DNA's compared	Per cent nucleotide substitutions observed since the time of divergence	Million years since divergence $\times 2$	Rate of nucleotide fixation since divergence (per cent/year) $\times 10^6$
Man-Chimp	2.5	30	0.08
Man-Gibbon	5.1	60	0.08
Man-Green monkey	9.0	90	0.1
Man- <i>Rhesus</i>	8.3	90	0.09
Man-Capuchin	15.8	130	0.12
Man-Galago	42	160	0.26
Mouse-Rat	30	20	1.5
Cow-Sheep	11.2	50	0.22

comparisons. Calculated on an absolute time basis, the rate of change of rodent nucleotide sequences is about 10 times faster than that for the primates or bovids. These data plus the rate data of Table IV suggest that the rate of nucleotide sequence change has not been the same for all species over evolutionary time. It must again be emphasized that the large difference in the rates seen for primates and rodents depend upon the divergence times used for these calculations. Wilson and Sarich [16] have recently questioned the paleontologically determined times of divergence for man and chimp. They contend from amino acid sequence and immunological evidence that the actual times of divergence of these species has been much more recent than the bulk of the fossil evidence indicates. The contention basic to their interpretation is that the *rate of molecular evolution has been the same for all species and has also been constant over absolute time*. Obviously, if this can be shown to be true, the degree of molecular change can be used as an evolutionary clock and times of divergence can be calculated with good accuracy from biochemical data, if just one fairly accurate time of divergence is known. Wilson and Sarich used their data and a time of divergence for old world monkeys (for example, *Rhesus* and green monkey) and man of 30 million years to calculate that the man-chimp divergence time was about five million years ago. Table VI shows the nucleotide sequence change data treated in a similar manner. The divergence time estimate of mouse-rat derived in this way is far larger than that estimated from the fossil record.

It must again be emphasized that the observed large difference in rates of nucleotide sequence divergence for primates and rodents depends on the divergence times used for these calculations.

TABLE VI

## CALCULATIONS OF DIVERGENCE TIMES PREDICTED FROM ASSUMPTIONS

The assumptions are that rate of DNA molecular evolution is constant with absolute time and the same in all species, and further, that man and green monkey diverged as early as 30 million years ago. (The fossil dates of divergence can't easily be reconciled with those calculated from a constant rate assumption.)

DNA's compared	Per cent nucleotide substitutions since divergence	Million years since divergence	Divergence time estimates based on fossil evidence (million years)
Man-Green monkey	9	30	30-45
Man-Chimp	2.5	$\frac{2.5}{9} \times 30 = 8$	10-25
Mouse-Rat	30	$\frac{30}{9} \times 30 = 100$	5-20

2.3. *Generation time correlation of rate of nucleotide change.* Table VII shows the values obtained when the rate of nucleotide sequence change is calculated in terms of generation time. The generation time used for the calculation was in each case the shortest generation time of the two animals compared. The rates calculated on this basis are surprisingly similar. Table VII also includes data for the rat-mouse nonrepeated DNA comparisons and also data on cow-sheep comparisons done by Laird, McConaughy, and McCarthy [10]. The rates of divergence (corrected for generation time) seen in these lines are very similar to the primate values, even though wide differences in generation times exist. The data of Table VII give a strong suggestion that the mutational events seen in evolution may be generation time dependent [9]. It seems highly improbable

TABLE VII

## GENERATION TIME BASED RATE OF NUCLEOTIDE SUBSTITUTION

DNA's compared	B		Per cent nucleotide substitutions per generation ( $A \times B$ )
	A	Estimated generation time using shortest generation time of pair	
	Per cent nucleotides fixed per year since divergence		
Man-Chimp	$0.08 \times 10^{-6}$	~10 years	$0.8 \times 10^{-6}$
Man-Gibbon	$0.08 \times 10^{-6}$	~10	$0.8 \times 10^{-6}$
Man-Green monkey	$0.1 \times 10^{-6}$	2-4	$0.3 \times 10^{-6}$
Man-Rhesus	$0.09 \times 10^{-6}$	2-4	$0.27 \times 10^{-6}$
Man-Capuchin	$0.1 \times 10^{-6}$	2-4	$0.3 \times 10^{-6}$
Man-Galago	$0.26 \times 10^{-6}$	1-2	$0.39 \times 10^{-6}$
Mouse-Rat	$1.5 \times 10^{-6}$	0.33 year	$0.5 \times 10^{-6}$
Cow-Sheep	0.22	1-2 years	$0.33 \times 10^{-6}$

that fluctuating mutational events in each of the three groups could fortuitously give essentially the same result. In calculating the values for Table VII, it is assumed that the present day generation times provide a reasonable estimate of the generation time history of the various species' lines.

The changes in the absolute time based rates of nucleotide change seen in Table IV could also be explained by taking into account the probable generation time history of the species' lines. If the generation times have increased with the development of the present day primates, the slowing of the absolute time based rate of change would be due to a decrease in the number of generations per million years as the primate lines evolved.

It appears from the fossil records that generation times may have been lengthening during the development of the primate line. The "precursor" primate was apparently a small rodent-like creature which is presumed to have had a short generation time. There is today a general, but not absolute, correlation between animal size and generation time. In general the larger the animal the longer its generation time.

The data on nucleotide substitution rates suggests that the rate of molecular evolution may have been different for different species. If true, this would be a troublesome observation. An assumption basic to evolutionary studies has been that morphological and molecular similarity of two species is an index of the temporal relationship between those species. In this view, species which have diverged most recently would be more similar molecularly and morphologically. This situation is certainly true if the rate of molecular evolution has been the same for all species at any time during evolution. If, however, the rate of molecular and morphological change can be different for different species' lines during the same evolutionary period, the situation becomes difficult. In this case it is possible (particularly in the case of closely related species) that the similarity between species will give an erroneous view of the relative times of divergences of the species in question. Table VIII illustrates this point. It is

TABLE VIII

## EFFECT OF RATE MODEL ON INTERPRETATION OF SIMILARITY DATA

- 
- 
- I. Data. Of the three species *X*, *Y*, and *Z*, *X* and *Y* are most similar to one another. Species *Z* is equally dissimilar to species *X* and *Y*.
  - II. Assumption. The rate of morphological and molecular change has been the same for all species during any evolutionary period.
    - A. From this assumption, the pattern of relatedness among the three species is that species *X* and *Y* had diverged more recently than *X* and *Z* or *Y* and *Z*.
  - III. Assumption. The rate of morphological and molecular change has been different for different species during any evolutionary period.
    - A. From this assumption, the pattern of relatedness could be as is described in IIA. If the rate change in different species is variable enough, species *X* and *Z* could have diverged more recently than species *X* and *Y* or *Z* and *Y*.
-

important, then, to know something about the rate at which change has occurred during the evolution of the various species' lines, since different rate models give different interpretations for the similarities seen between species.

### 3. The nature of the nucleotide substitutions

A wide range of effects of nucleotide substitutions on an organism is possible [5]. Nucleotide substitutions in DNA sequences which have *no function* will not affect the organism. It is not known how much of the total cell DNA is actually needed. There may be a class of DNA which can tolerate no nucleotide changes. Such a class of DNA will suffer nucleotide substitutions but these changes will be selected against and not passed along. It is known that very little of the DNA falls in this class. Another class of DNA may determine the templates for proteins which require an unchanged sequence of amino acids. Roughly, one fourth of the total nucleotide substitutions in this class will have no effect on the amino acid sequence, since in many triplets the third base can be changed to any other base without changing the amino acid inserted into the protein (for example, UCU, UCC, UCA, UCG all code for the amino acid serine). The remainder of the DNA (not in one of the above three classes) can be considered in a class intermediate between these defined extremes.

As is apparent from the above paragraph, little is known about the effect of the nucleotide substitutions which have occurred during evolution. The above discussion implies that many base changes which have occurred could have little if any effect on the phenotype of the organism, and are not likely to have much selection pressure on them. There is a chance, then, that many, if not most, of the nucleotide substitutions measured in the experiments reported here may be neutral or close to neutral in their effect on the organism. If this is so, an enormous number of nucleotide substitutions can accumulate in the DNA with little effect on the organism.

Perhaps the best available indication that many changes can accumulate in a genome lies in the behavior of the Treffers mutator strain of *E. coli* [3]. This bacterial strain has a mutation rate hundreds of times higher than the normal *E. coli*. It has been shown that an average of seven nucleotide substitutions per bacterium per cell division occurs in the mutation strain of bacteria [3]. These bacteria grow virtually as well and as rapidly on minimal media as do wild type bacteria, even after many thousands of cell divisions. This directly indicates that many nucleotide substitutions can occur in a living organism without appreciably affecting its short term viability. Over the period during which the mutator strain has been studied, it would not be unreasonable to say that most of the base changes which have occurred are "neutral." Perhaps it would be better to say that the great majority of nucleotide substitutions have had little selective value on them. What the effect of these changes will be over a longer time is not known.

Many base changes can accumulate in the Treffers bacteria without affecting

the viability of the bacteria. This indicates that the fraction of the total base changes which produce a lethal effect is relatively small. A rough estimate of the ratio of lethal base changes to the total base changes can be made. Each bacterium contains on the average seven new base substitutions after every cell division. If the ratio of lethal to total base changes is one seventh, the bacterial population would not increase in number. It is known that the mutator strain grows virtually as well as the wild type. The available data indicate that in minimal media cultures roughly 10 per cent of the mutator bacteria are auxotrophs (they have mutated so that they will no longer grow on minimal media). This indicates that roughly one in 25 nucleotide changes is lethal. It seems reasonable to say further that the majority of the other 24 changes has little effect on the viability of the bacteria, since they grow so well. The 1/25 ratio is a rough one and the ratio may well be much smaller.

It does seem, then, that in a relatively simple organism such as a bacterium a large number of base changes can occur and accumulate, apparently with little effect on the bacterial viability. This raises the very real possibility that a similar situation exists in higher organisms. If so, most of the nucleotide sequence changes studied here may well have had little effect on the evolution of the organism. Moreover the rates of nucleotide fixation observed may be a reasonable reflection of the actual base substitution rate during evolutionary time.

3.1. *Divergence of DNA sequences expressed in a specific tissue.* Recent application of DNA:RNA hybridization technique have enabled the isolation from specific tissues of DNA sequences which have been expressed as RNA [8]. These isolated expressed DNA sequences (E-DNA) have been utilized to ask whether E-DNA sequences diverge at the same average rate as do the bulk of the DNA sequences [8]. This approach was tried because we do not know the function of the nonrepeated DNA sequences examined in the earlier paragraphs. The one function which some of these nonrepeated DNA sequences have is that they are expressed as RNA. Unfortunately, we do not know what the function of the RNA is, but at least the E-DNA sequences have been shown to be active in present day creatures.

The experiments to determine the rate of divergence of the E-DNA relative to the bulk DNA are similar to those already described earlier for the bulk DNA sequences. Cow radioactive nonrepeated E-DNA (from brain or liver) was reacted with sheep and pig nonradioactive DNA's. The thermal stabilities of the hybrid molecules were then determined and compared to the thermal stability of hybrids formed by reacting the total radioactive nonrepeated cow DNA with sheep and pig nonradioactive DNA.

About six per cent of the nonrepeated DNA was isolated as being expressed in the liver while about three per cent was detected as being expressed in the brain (these values are minimum indications of the actual extent of expression). Both liver and brain E-DNA diverged at the same average rate as the total nonrepeated DNA. Expression, then, does not appear to confer any great selective advantage to a DNA sequence.

#### 4. Conclusion

The current view of the role of DNA indicates that evolutionary changes at the organismal level must have been preceded by some quantitative or qualitative change in the DNA. Detectable in the DNA are repeated DNA sequences, nonrepeated DNA sequences, base changes, and translocations. These are the present day "residues" of historical events which occurred during the evolutionary history of the DNA. Correlating molecular changes with organismal changes is always a difficult task. Comparison of the pattern and character of these "DNA fossils" with the patterns derived from the classical fossil record should, however, provide new insight into the nature of evolutionary forces. The evidence available thus far allows a very tentative correlation to be made. During the period where the classical fossil record indicates that extensive speciation occurred in the primates it appears that (a) DNA was added to the presumptive human genome at a much faster rate than during other periods for which data is available [7], and (b) the rate of nucleotide sequence divergence was also much higher relative to other periods.

It is not known whether these increased rates are the result or the cause of the extensive speciation seen during this period. The correlation of nucleotide sequence change with generation time suggests that the generation time of the early primate ancestors of man which lived during this period was relatively short. The rate of mixing of genetic material to provide new genetic combinations (for example, translocations, inversions, and so forth) should also be generation time dependent. It does not seem unreasonable to speculate that the rapid addition, divergence, and mixing of the DNA may have played a role in the burst of speciation seen during this period. Much more extensive work needs to be done, however, before any strong statements can be made.

#### REFERENCES

- [1] E. T. BOLTON, R. J. BRITTEN, T. J. BYERS, D. B. COWIE, B. H. HOYER, Y. KATO, B. J. MCCARTHY, M. MIRANDA, and R. B. ROBERTS, "Biophysics section report," *Carnegie Institution of Washington Year Book*, Vol. 63 (1964), pp. 366-398.
- [2] R. J. BRITTEN and D. E. KOHNE, "Repeated sequences in DNA," *Science*, Vol. 161 (1968), p. 529.
- [3] E. C. COX and C. YANOFSKY, "Altered base ratios in the DNA of an *Escherichia coli* mutator strain," *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 58 (1967), pp. 1895-1902.
- [4] B. H. HOYER and R. B. ROBERTS, "Studies of DNA homology by the DNA-agar technique," *Molecular Genetics*, New York, Academic Press, 1967, pp. 425-479.
- [5] J. L. KING and T. H. JUKES, "Non-Darwinian evolution: random fixation of selectively neutral mutations," *Science*, Vol. 164 (1969), pp. 788-798.
- [6] D. E. KOHNE, "Isolation and characterization of bacterial ribosomal RNA genes," *Biophys. J.*, Vol. 8 (1968), pp. 1104-1110.
- [7] ———, "Evolution of higher organism DNA," *Quart. Rev. Biophys.*, Vol. 3 (1970), pp. 317-375.
- [8] D. E. KOHNE and M. J. BYERS, "Studies on the evolutionary divergence of expressed DNA sequences," in preparation.



- [9] D. E. KOHNE, J. A. CHISCON, and B. H. HOYER, "Evolution of primate DNA," in preparation.
- [10] C. D. LAIRD, B. L. McCONAUGHY, and B. J. McCARTHY, "Rate of fixation of nucleotide substitutions in evolution," *Nature*, Vol. 224 (1969), pp. 149-154.
- [11] J. MARMUR, R. ROWND, and C. L. SCNILDKRAUT, "Denaturation and renaturation of DNA," *Progr. Nucl. Acid Res. Mol. Biol.*, Vol. 1 (1963), pp. 232-300.
- [12] M. McKENNA, Personal communication.
- [13] A. S. ROMER, *Vertebrate Paleontology*, Chicago, University of Chicago Press, 1966.
- [14] E. SIMONS, "The early relatives of man," *Sci. Amer.*, Vol. 211 (1964), pp. 50-62.
- [15] G. G. SIMPSON, "The nature and origin of supraspecific taxa," *Cold Spring Harbor Symp. Quant. Biol.*, Vol. 24 (1959), pp. 255-272.
- [16] A. C. WILSON and V. M. SARICH, "A molecular time scale for human evolution," *Proc. Nat. Acad. Sci., U.S.A.*, Vol. 63 (1969), pp. 1088-1093.