11 April 1975, Volume 188, Number 4184 SCIENCE

Evolution at Two Levels in Humans and Chimpanzees

Their macromolecules are so alike that regulatory mutations may account for their biological differences.

Mary-Claire King and A. C. Wilson

We suggest that evolutionary changes in anatomy and way of life are more often based on changes in the mechanisms controlling the expression of genes than on sequence changes in proteins. We therefore propose that regulatory mutations account for the major biological differences between humans and chimpanzees.

Soon after the expansion of molecular biology in the 1950's, it became evident that by comparing the proteins and nucleic acids of one species with those of another, one could hope to obtain a quantitative and objective estimate of the "genetic distance" between species. Until then, there was no common yardstick for measuring the degree of genetic difference among species. The characters used to distinguish among bacterial species, for example, were entirely different from those used for distinguishing among mammals. The hope was to use molecular biology to measure the differences in the DNA base sequences of various species. This would be the common yardstick for studies of organismal diversity.

During the past decade, many workers have participated in the development and application of biochemical methods for estimating genetic distance. These methods include the comparison of proteins by electrophoretic, immunological, and sequencing techniques, as well as the comparison of nucleic acids by annealing techniques. The only two species which have been compared by all of these methods are chimpanzees (*Pan troglodytes*) and humans (*Homo sapiens*). This pair of species is also unique because of the thoroughness with which they have been compared at the organismal level—that is, at the level of anatomy, physiology, behavior, and ecology. A good opportunity is therefore presented for finding out whether the molecular and organismal estimates of distance agree.

The intriguing result, documented in this article, is that all the biochemical methods agree in showing that the genetic distance between humans and the chimpanzee is probably too small to account for their substantial organismal differences.

Indications of such a paradox already existed long ago. By 1963, it appeared that some of the blood proteins of humans were virtually identical in amino acid sequence with those of apes such as the chimpanzee or gorilla (*1*). In the intervening years, comparisons between humans and chimpanzees were made with many additional proteins and with DNA. These results, reported herein, are consistent with the early results. Moreover, they tell us that the genes of the human and the chimpanzee are as similar as those of sibling species of other organisms (*2*). So, the paradox remains. In order to explain how species which have such similar genes can differ so substantially in anatomy and way of life, we review evidence concerning the molecular basis of evolution at the organismal level.

Dr. King, formerly a graduate student in the Departments of Genetics and Biochemistry, University of California, Berkeley, is now a research geneticist at the Hooper Foundation and Department of International Health, University of California, San Francisco 94143. Dr. Wilson is a professor of biochemistry at the University of California, Berkeley 94720.

Similarity of Human and Chimpanzee Genes

To compare human and chimpanzee genes, one compares either homologous proteins or nucleic acids. At the protein level, one way of measuring the degree of genetic similarity of two taxa is to determine the average number of amino acid differences between homologous polypeptides from each population. The most direct method for determining this difference is to compare the amino acid sequences of the homologous proteins. A second method is microcomplement fixation, which provides immunological distances linearly correlated with amino acid sequence difference. A third method is electrophoresis, which is useful in analyzing taxa sufficiently closely related that they share many alleles. For the human-chimpanzee comparison all three methods are appropriate, and thus many human and chimpanzee proteins have now been compared by each method. We can therefore estimate the degree of genetic similarity between humans and chimpanzees by each of these techniques.

Sequence and immunological comparisons of proteins. During the last decade, amino acid sequence studies have been published on several human and chimpanzee proteins. As Table 1 indicates, the two species seem to have identical fibrinopeptides (*3*), cytochromes c (*4*), and hemoglobin chains [alpha (*4*), beta (*4*), and gamma (*5*, *6*)]. The structural genes for these proteins may therefore be identical in humans and chimpanzees. In other cases, for example, myoglobin (*7*) and the

delta chain of hemoglobin (*5, 8*), the human polypeptide chain differs from that of the chimpanzee by a single amino acid replacement. The amino acid replacement in each case is consistent with a single base replacement in the corresponding structural gene.

Owing to the limitations of conventional sequencing methods, exactly comparable information is not available for larger proteins. Indeed, the sequence information available for the proteins already mentioned is not yet complete. By applying the microcomplement fixation method to large proteins, however, one can obtain an approximate measure of the degree of amino acid sequence difference between related proteins (*9*). This method indicates that the sequences of human and chimpanzee albumins (*10*), transferrins (*11*), and carbonic anhydrases (*4, 12*) differ slightly, but that lysozyme (*13*) is identical in the two species (Table 1) (*14*). Based on the proteins listed in Table 1, the average degree of difference between human and chimpanzee proteins is

$$\frac{19 \times 1000}{2633} = 7.2 \qquad (1)$$

amino acid sites per 1000 substitutions. That is, the sequences of human and chimpanzee polypeptides examined to date are, on the average, more than 99 percent identical.

Electrophoretic comparison of proteins. Electrophoresis can provide an independent estimate of the average amino acid sequence difference between closely related species. We have compared the human and chimpanzee polypeptide products of 44 different structural genes. Table 2 indicates the allelic frequencies and the estimated probability of identity at each locus. The

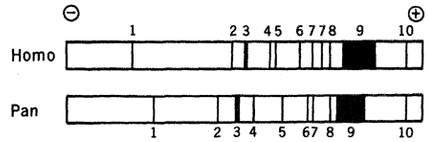


Fig. 1. Separation of human and chimpanzee plasma proteins by acrylamide electrophoresis at *p*H 8.9. The proteins are: 1, α_2-macroglobulin; 2, third component of complement; 3, transferrin; 4, haptoglobin; 5, ceruloplasmin; 6, α_{2HS}-glycoprotein; 7, Gc-globulins; 8, α_1-antitrypsin; 9, albumin; and 10, α_1-acid glycoprotein. The chimpanzee plasma has transferrin genotype *Pan* CC; the human plasma has transferrin genotype *Homo* CC and haptoglobin genotype 1-1. The direction of migration is from left to right.

symbol S_i represents the probability that human and chimpanzee alleles will be electrophoretically identical at a particular locus *i*, or

$$S_i = \sum_{j=1}^{A_i} x_{ij} y_{ij} \qquad (2)$$

where x_{ij} is the frequency of the *j*th allele at the *i*th locus in human populations, and y_{ij} the frequency of the *j*th allele at the *i*th locus in chimpanzee populations for all A_i alleles at that locus. For example, Table 2 indicates the frequencies of the three alleles (AP^a, AP^b, and AP^c) found at the acid phosphatase locus for human and chimpanzee populations. The probability of identity of human and chimpanzee alleles at this locus, that is, S_i is $(0.29 \times 0) + (0.68 \times 1.00) + (0.03 \times 0)$, or 0.68.

Of the loci in Table 2, 31 code for intracellular proteins; 13 code for secreted or extracellular proteins. In general, the intracellular proteins were analyzed by starch gel electrophoresis of red blood cell lysates, with the buffer systems indicated in the table and stains specific for the enzymatic activity of each protein. For a few intracellular proteins (cytochrome c, the hemoglobin chains, and myoglobin), amino acid sequences have been published for both species, so that direct sequence comparison is also possible.

Most of the secreted proteins were compared by acrylamide gel electrophoresis of human and chimpanzee plasma (*15*). The electrode chamber contained tris(hydroxymethyl)aminomethane (tris) borate buffer, *p*H 8.9; acrylamide gel slabs were made with tris-sulfate buffer, *p*H 8.9. Gels were stained with amido black, a general protein dye. The identification of bands on a gel stained with this dye poses a problem, since it is not obvious, particularly for less concentrated proteins, which protein each band represents. We determined the electrophoretic mobilities of the plasma proteins by applying the same sample to several slots of the same gel, staining the outside columns, and cutting horizontal slices across the unstained portion of the gel at the position of each band. The protein was eluted separately from each band in 0.1 to 0.2 milliliter of an appropriate isotonic tris buffer (*9*) and tested for reactivity with a series of rabbit antiserums, each specific for a particular human plasma protein, by means of immunoelectrophoresis and immunodiffusion in agar (*15, 16*). The results of this analysis are shown in Fig. 1.

Some of the secreted proteins were compared by means of other electrophoretic methods as well. Albumin and transferrin were surveyed by cellulose acetate electrophoresis; and α_1-antitrypsin, Gc-globulin (group-specific component), the haptoglobin chains, lysozyme, and plasma cholinesterase were analyzed on starch gels, with the buffers indicated in Table 2.

The results of all electrophoretic comparisons are summarized in Fig. 2. About half of the proteins in this survey are electrophoretically identical for the two species, and about half of them are different. Only a few loci are highly polymorphic in both species (see *17*).

The proportion of alleles at an "average" locus that are electrophoretically identical in human and chimpanzee populations can be calculated from Table 2 and Eq. 3, where *L* is the number of loci observed:

$$\bar{S} = \frac{1}{L}(S_1 + S_2 + \ldots + S_L) = 0.52 \qquad (3)$$

Table 1. Differences in amino acid sequences of human and chimpanzee polypeptides. Lysozyme, carbonic anhydrase, albumin, and transferrin have been compared immunologically by the microcomplement fixation technique. Amino acid sequences have been determined for the other proteins. Numbers in parentheses indicate references for each protein.

Protein	Amino acid differences	Amino acid sites
Fibrinopeptides A and B (*3*)	0	30
Cytochrome c (*4*)	0	104
Lysozyme (*13*)	~0	130
Hemoglobin α (*4*)	0	141
Hemoglobin β (*4*)	0	146
Hemoglobin $^A\gamma$ (*5, 6*)	0	146
Hemoglobin $^G\gamma$ (*5, 6*)	0	146
Hemoglobin δ (*5, 8*)	1	146
Myoglobin (*7*)	1	153
Carbonic anhydrase (*4, 12*)	~3	264
Serum albumin (*10*)	~6	580
Transferrin (*11*)	~8	647
Total	~19	2633

In other words, the probability that human and chimpanzee alleles will be electrophoretically identical at a particular locus is about one-half.

Agreement between electrophoresis and protein sequencing. The results of electrophoretic analysis can be used to estimate the average number of amino acid differences per polypeptide chain for humans and chimpanzees, for comparison with the estimate based on amino acid sequences and immunological data. To calculate the average amino acid sequence difference between human and chimpanzee proteins, we need first an estimate of the proportion ($\bar{c}$) of amino acid substitutions detectable by electrophoresis. Electrophoretic techniques detect only amino acid substitutions that change the net charge of the protein observed. Four amino acid side chains are charged at *p*H 8.6: arginine, lysine, glutamic acid, and aspartic acid. The side chain of histidine is positively charged below approximately *p*H 6. The proportion of accepted point mutations that would be detectable by the buffer

Table 2. Electrophoretic comparison of chimpanzee and human proteins. In the first column, Enzyme Commission numbers are given in parentheses; N is the number of chimpanzees analyzed, both in this study and by other investigators. Abbreviations: MW, molecular weight; aa, amino acids; tris, tris(hydroxymethyl)aninomethane; EDTA, ethylenediaminetetraacetate. Secreted proteins differ more frequently for the two species than intracellular proteins (*93*).

Locus (i) and allele (j)	Allele frequency		Probability of identity† (S_i)	Comments and references‡
	Human* (x_{ij})	Chimpanzee (y_{ij})		
			Intracellular proteins	
Acid phosphatase (3.1.3.2); $N = 86$				
AP^a	0.29	0	0.68	Red cells; 15,000 MW; 110 aa; citrate-phosphate, *p*H 5.9, starch electrophoresis (*54, 55*)
AP^b	0.68	1.00		
AP^c	0.03	0		
Adenosine deaminase (3.5.4.4); $N = 22$				
ADA^1	0.96	0	0	Red cells; 35,000 MW; 300 aa; chimpanzee protein faster on starch electrophoresis (*54*); polymorphism in human populations (*16*)
ADA^2	0.04	0		
$ADA^{ape=5}$	0	1.00		
Adenylate kinase (2.7.4.3); $N = 86$				
AK^1	0.98	1.00	0.98	Red cells; 21,500 MW; 190 aa; well buffer is citrate-NaOH, *p*H 7.0; gel buffer is histidine-NaOH, pH 7.0, starch electrophoresis (*54, 56, 57*)
AK^2	0.02	0		
Carbonic anhydrase I or B (4.2.1.1); $N = 111$	1.00	1.00	1.00	Red cells; 28,000 MW; 264 aa; well buffer is borate-NaOH, *p*H 8.0; gel buffer is borate-NaOH, *p*H 8.6, starch electrophoresis (*56, 58*)
Cytochrome c	1.00	1.00	1.00	Mitochondria; 12,400 MW; 104 aa; sequence identity based on amino acid analysis (*5*); possible heterogeneity in man (*59*)
Esterase A_1 (3.1.1.6); $N = 111$	1.00	0	0	Red cells; well buffer is lithium borate, *p*H 8.2; gel buffer is lithium-borate and tris-citrate, *p*H 7.3, starch electrophoresis (*58, 60*)
	0	1.00		
Esterase A_2§ (3.1.1.6); $N = 111$	1.00	Absent		See esterase A_1
Esterase A_3 (3.1.1.6); $N = 111$				
$EstA_3^a$	1.00	0	0	See esterase A_1
$EstA_3^b$	0	1.00		
Esterase B (3.1.1.1); $N = 111$	1.00	1.00	1.00	See esterase A_1
Glucose-6-phosphate dehydrogenase (1.1.1.49); $N = 86$				
Gd^A	0.01	0	0.99	Red cells; six subunits, each 43,000 MW; ~ 370 aa; phosphate, *p*H 7.0, starch electrophoresis (*56*); A and B variants identical by microcomplement fixation (*61*); sequences differ by one amino acid, aspartic acid in A variant, asparagine in B variant (*61*)
Gd^B	0.99	1.00		
Glutamate-oxalacetate transaminase (soluble form) (2.6.1.1); $N = 63$				
$sGOT^1$	1.00	0	0	Red cells; two subunits, each 50,000 MW; ~ 430 aa; tris-citrate, *p*H 7.0, starch electrophoresis (*62*); chimpanzee protein faster (*63*)
$sGOT^2$	0	1.00		
Glutathione reductase (1.6.4.2); $N = 64$				
GSR^2 and GSR^3	0.97	1.00	0.97	Red cells; tris-EDTA, *p*H 9.6, starch electrophoresis; polymorphism in human populations (*64*), possibly associated with gout; GSR^2 and GSR^3 not distinguishable at *p*H 9.6
GSR^5	0.01	0		
GSR^6	0.02	0		
Hemoglobin α chain; $N = 108$				
Hb_α^A	1.00	0.99	0.99	Red cells; 15,100 MW; 141 aa; tris-glycine, *p*H 8.4, cellulose acetate electrophoresis (*15*); tryptic peptides of human and chimpanzee α chains identical (*65*); chimpanzee α chain variant is electrophoretically identical to human Hb^J (*66*)
Hb_α^J	< 0.01	0.01		

(Table 2 is continued on pages 110 and 111)

Locus (i) and allele (j)	Allele frequency		Probability of identity† (S_i)	Comments and references‡
	Human* (x_{ij})	Chimpanzee (y_{ij})		
Hemoglobin β chain; $N = 108$				
Hb_β^A	0.99	0.99	0.98	Red cells; 16,000 MW; 146 aa; tris-glycine, pH 8.4, cellulose acetate electrophoresis (*15*); amino acid sequences of β^A chains identical (*65*); chimpanzee Hb^B electrophoretically identical to human Hb^S (*66*)
$Hb_\beta^{S=B}$	0.01	0.01		
Hemoglobin $^A\gamma$ chain	1.00	1.00	1.00	Fetal red cells; 16,000 MW; 146 aa; amino acid sequence of human and chimpanzee γ chains identical; $^A\gamma$ and $^G\gamma$ arep roducts of different structural genes, differ at residue 136; A, alanine; G, glycine (*67*)
Hemoglobin $^G\gamma$ chain	1.00	1.00	1.00	See hemoglobin $^A\gamma$
Hemoglobin δ chain	1.00	1.00	1.00	Red cells; 16,000 MW; 146 aa; human and chimpanzee electrophoretic mobilities identical, but one amino acid difference at position 125: humane δ, methionine; chimpanzee δ, valine (*8*)
Lactate dehydrogenase H (1.1.1.27); $N = 74$	1.00	1.00	1.00	Red cells; H and M subunits each 34,000 MW; 330 aa; citrate-phosphate, pH 6.0, starch electrophoresis (*69*); three intermediate bands of five-band, tetrameric electrophoretic pattern have different mobilities for humans and chimpanzees, because of difference in M polypeptide (*70*)
Lactate dehydrogenase M (1.1.1.27); $N = 74$				
ldh M^a	1.00	0	0	See lactate dehydrogenase H
ldh M^b	0	1.00		
Malate dehydrogenase (cytoplasmic) (1.1.1.37); $N = 88$	1.00	1.00	1.00	Red cells; two subunits, each 34,000 MW; 330 aa; see LDH for procedures; polymorphic in some human populations (*71*)
Methemglobin reductase (1.6.99); $N = 86$				
MR^1	1.00	0	0	Red cells; tris-citrate, pH 6.8, starch electrophoresis (*72*) distinguishes human and chimpanzee enzymes, no difference with tris-EDTA, pH 9.3, electrophoresis (*56, 73*)
MR^2	0	1.00		
Myoglobin	1.00	1.00	1.00	Muscle; 16,900 MW; 153 aa; tryptic and chymotryptic peptides of cyanmethemoglobin electrophoretically identical at pH 8.6 (*74*), but at position 116, human has glutamine, chimpanzee has histidine (*7*)
Peptidase A (3.4.3.2); $N = 63$				
$PepA^1$ and $PepA^S$	0.99	1.00	0.99	Red cells; two subunits, each 46,000 MW; ~ 400 aa; tris-maleate, pH 7.4 starch electrophoresis, leucyl-glycine substrate (*65*); $PepA^1$ and $PepA^S$ not distinguishable in red blood cell lysates (*75*)
$PepA^2$	0.01	0		
Peptidase C (3.4.3.2); $N = 63$				
$PepC^1$	0.99	1.00	0.99	Red cells; 65,000 MW; ~ 565 aa; see peptidase A for procedures; polymorphism in human populations (*76*)
$PepC^4$	0.01	0		
Phosphoglucomutase 1 || (2.7.5.1); $N = 168$				
PGM_1^1	0.77	0.26	0.20	Red cells; subunits PGM_1 and PGM_2 each 62,000 MW; ~ 540 aa; tris-maleate-EDTA, pH 7.4, starch electrophoresis (*16, 55, 61, 77*)
PGM_1^2	0.23	0		
PGM_1^{Pan}	0	0.74		
Phosphoglucomutase 2 (2.7.5.1); $N = 168$				
PGM_2^1	1.00	1.00	1.00	See phosphoglucomutase 1
PGM_2^3	$<< 0.01$	< 0.01		
6-Phosphogluconate dehydrogenase (1.1.1.44); $N = 86$				
PGD^A	0.96	0	0.04	Red cells; two subunits, each 40,000 MW; 350 aa; see G6PD for procedures; chimpanzee allele electrophoretically identical to human "Canning" variant (*55*)
PGD^C	0.04	1.00		
Phosphohexose isomerase (5.3.1.9); $N = 86$				
PHI^1	1.00	0	0	Red cells; two subunits, each 66,000 MW; 580 aa; tris-citrate, pH 8.0, starch electrophoresis (*56*); chimpanzee protein has slower mobility, both cathodally migrating (*78*)
PHI^B	0	1.00		
Superoxide dismutase A (indophenol oxidase) (1.15.1.1); $N = 64$	1.00	1.00	1.00	Red blood cells; two subunits, each 16,300 MW; 158 aa (*68*); see phosphoglucomutase for procedure
Triosephosphate isomerase A (5.3.1.1)	1.00	1.00	1.00	Fibroblasts; dimers 48,000 MW; each polypeptide 248 aa (*79*); β polypeptide found only in hominoids.
Triosephosphate isomerase B (5.3.1.1)	1.00	1.00	1.00	See triosephosphate isomerase A

Locus (i) and allele (j)	Allele frequency		Probability of identity† (S_i)	Comments and references‡
	Human* (x_{ij})	Chimpanzee (y_{ij})		
			Secreted proteins	
α_1-Acid glycoprotein (orosomucoid); $N = 123$				
Or^S	0.32	0	0.68	Glycoprotein in plasma; carbohydrate > 50 percent; 44,100 MW; 181 aa; acrylamide electrophoresis, *p*H 8.9 (see text); polymorphism in human populations detectable at *p*H 2.9 (*80*); isoelectric point is 1.82 for human and chimpanzee proteins, but proteins differ by quantitative precipitin analysis (*81*)
Or^F	0.68	1.00		
Albumin; $N = 123$				
Alb^A	0	1.00	0	Plasma; 69,000 MW, ~ 580 aa; tris-citrate, *p*H 5.5, cellulose acetate electrophoresis; acrylamide electrophoresis, *p*H 8.9; chimpanzee protein slower mobility, immunological difference detected by microcomplement fixation (*10, 42*); rare polymorphic alleles in human populations (*82*)
Alb^{Pan}	1.00	0		
α_1-Antitrypsin; $N = 123$				
Pi^M	0.95	0	0	Plasma; 49,000 MW; ~ 380 aa; anodal well buffer is citrate-phosphate, *p*H 4.5; cathodal well buffer is borate-NaOH, *p*H 9.0; gel buffer is tris-citrate, *p*H 4.8; starch electrophoresis (*56*); acrylamide electrophoresis, *p*H 8.9; polymorphism in human populations (*83*)
Pi^F	0.03	0		
Pi^S	0.02	0		
Pi^{Pan}	0	1.00		
Ceruloplasmin; $N = 123$				
Cp^A and Cp^{Pan}	0.01	1.00	0.01	Plasma; eight subunits, each 17,000 MW; ~ 150 aa; acrylamide electrophoresis, *p*H 8.9; possible adaptive significance of polymorphism in human populations (*84*)
Cp^B	0.98	0		
Cp^C	0.01	0		
Third component of complement; $N = 123$				
$C'3^{1=F}$	0.12	0	0	Plasma; total MW 240,000; acrylamide electrophoresis, *p*H 8.9; polymorphism in human populations detectable by high voltage electrophoresis (*85*)
$C'3^{2=S}$	0.87	0		
$C'3^S$	0.01	0		
$C'3^{Pan}$	0	1.00		
Group-specific component; $N = 206$				
Gc^1	0.74	0	0	Plasma; two subunits, each 25,000 MW; ~ 220 aa; acrylamide electrophoresis, *p*H 8.9; human Gc 2-2 and chimpanzee protein similar on acrylamide, chimpanzee slightly faster on starch or immunoelectrophoresis (*86*)
Gc^2	0.26	0		
Gc^{Pan}	0	1.00		
α_{2HS}-Glycoprotein; $N = 123$				
Gly^A	1.00	0	0	Plasma; 49,000 MW; ~ 400 aa; acrylamide electrophoresis, *p*H 8.9 (*15*)
Gly^B	0	1.00		
Haptoglobin α chain; $N = 300$				
$Hp_\alpha{}^1$	0.36	0	0	Plasma; α^1 chain is 8,900 MW, 83 aa; α^2 chain is 16,000 MW, 142 aa; β chain is 36,000 to 40,000 MW; ~ 330 aa; acrylamide electrophoresis, *p*H 8.9; borate-NaOH well buffer and tris-citrate gel buffer, *p*H 8.6, starch electrophoresis (*56*); chimpanzee Hp shares six human Hp 1-1 and eight Hp 2-2 antigenic determinants; Hp^2 evolved since human-chimpanzee divergence (*87*)
$Hp_\alpha{}^2$	0.64	0		
$Hp_\alpha{}^{Pan}$	0	1.00		
Haptoglobin β chain; $N = 300$	1.00	1.00	1.00	See haptoglobin α chain
Lysozyme				
lzm^A	1.00	0	0	Milk; 14,400 MW; 130 aa; starch gel electrophoresis, *p*H 5.3 (*88*)
lzm^B	0	1.00		
α_2-Macroglobulin; $N = 123$				
Xm^A	1.00	0	0	Plasma; four subunits, each 196,000 MW; acrylamide electrophoresis, *p*H 8.9; X-linked antigenic polymorphism observed in human populations (*89*) but not detectable by electrophoresis; human and chimpanzee proteins immunologically indistinguishable (*14*)
Xm^B	0	1.00		
Plasma cholinesterase (3.1.1.8); $N = 111$				
$E_1{}^u$	1.00	0	0	Plasma; four subunits, each ~ 87,000 MW; see esterase A_1 for procedures; chimpanzee protein has four components with faster mobilities than analogous human components (*15*)
$E_1{}^{Pan}$	0	1.00		
Transferrin; $N = 133$				
Homo: Tf^C	0.99	0	0	Plasma; 73,000 to 92,000 MW; ~ 650 aa; acrylamide electrophoresis, *p*H 8.9; tris-glycine, *p*H 8.4, cellulose acetate electrophoresis (*77, 90*)
Tf^{D1}	0.01	0		
Pan: Tf^A	0	0.08		
Tf^B	0	0.06		
Tf^C	0	0.70		
Tf^D	0	0.15		
Tf^E	0	0.02		

* Allelic frequencies for human populations are calculated from data summarized by Nei and Roychoudhury (*28*). Sample sizes generally greater than 1000. Only alleles with frequency > 0.01 are listed. The relative sizes of racial groups were estimated to be Caucasian, 45 percent; Black African, 10 percent; and Mongoloid-Amerind (combined), 45 percent. † See Eq. 2 in text. ‡Given in this column are: the tissue used, polypeptide chain length, electrophoretic conditions, and references to previous studies on people and chimpanzees. Genetic, population, and physiological studies of most human red cell and plasma proteins are summarized by Giblett (*56*) or Harris (*91*); studies of plasma proteins are summarized by Schultze and Heremans (*92*). References are for additional studies of chimpanzee or human proteins. § Not included in identity calculations. || Notation for the chimpanzee alleles at the PGM_1 locus differs in publishhed surveys. Ours is as follows: $PGM_1{}^{Pan}$ (which is chimpanzee $PGM_1{}^1$, of Goodman and co-workers and $PGM_1{}^{Pan}$ of Schmitt and co-workers) is the allele with slowest electrophoretic mobility; $PGM_1{}^1$ (which is human $PGM_1{}^1$, the chimpanzee $PGM_1{}^1$ of Schmitt, and the chimpanzee $PGM_1{}^2$ of Goodman) is intermediate; and $PGM_1{}^2$ (found only in human populations) has the fastest mobility.

systems used in this study is about 0.27 (*18*).

If we assume that, at a particular amino acid site on a given protein, amino acid substitutions have occurred (i) independently and (ii) at random with respect to species since the evolutionary divergence of humans and chimpanzees, then the number of proteins that have accumulated r amino acid substitutions since this divergence approximates a Poisson variate (*19*). That is, the probability that r substitutions have accumulated in a particular polypeptide is

$$P_r = \frac{(mc)^r e^{-mc}}{r!} \quad (4)$$

where m is the expected number of amino acid substitutions per polypeptide (the mean of the Poisson distribution), and c is the proportion of those substitutions that are electrophoretically detectable. The probability that the polypeptides are electrophoretically identical (that is, that no electrophoretically detectable substitutions have occurred) is 0.52. Therefore,

$$P_0 = 0.52 = \frac{(mc)^0 e^{-mc}}{0!} = e^{-mc} \quad (5)$$

Thus $mc = 0.65$ and the expected number of amino acid differences per polypeptide is

$$m = 0.65/0.27 = 2.41 \quad (6)$$

For comparative purposes, this value can also be expressed in terms of the expected number of amino acid differences per 1000 amino acids. The average number of amino acids per polypeptide for all the proteins analyzed electrophoretically is 293 ± 27 (standard error). Therefore the expected degree of amino acid difference between human and chimpanzee is

$$\frac{2.41 \times 1000}{293} = 8.2 \quad (7)$$

substitutions per 1000 sites, with a range (within one standard error) of 7.5 to 9.1 differences per 1000 amino acids. The estimate based on amino acid sequencing and immunological comparisons (Eq. 1) agrees well with this estimate. Both estimates indicate that the average human protein is more than 99 percent identical in amino acid sequence to its chimpanzee homolog (*20*).

Comparison of nucleic acids. Another method of comparing genomes is nucleic acid hybridization. Several workers have compared the thermostability of human-chimpanzee hybrid DNA formed in vitro with the thermostability of DNA from each species separately. By this criterion, human and chimpanzee mitochondrial DNA's appear identical (*21*). Working with "nonrepeated" DNA sequences, Kohne has estimated that human-chimpanzee hybrid DNA dissociates at a temperature (ΔT) 1.5°C lower than the dissociation temperature of reannealed human DNA (*22*). Hoyer *et al.*, on the other hand, have estimated that ΔT equals 0.7°C for human-chimpanzee hybrid DNA (*23*). If ΔT is the difference in dissociation temperature of reannealed human DNA and human-chimpanzee hybrid DNA prepared in vitro, then the percentage of nucleic acid sequence difference is $k \times \Delta T$ where the calibration factor k has been variously estimated as 1.5, 1.0, 0.9, or 0.45 (*22*, *24*). Based on k being 1.0 and ΔT being 1.1°C, the nucleic acid sequence difference of human and chimpanzee DNA is about 1.1 percent. In a length of DNA 3000 bases long (representing 1000 amino acids), there will be about 0.011×3000, or 33 nucleotide sequence differences between the two species.

The evidence from the DNA annealing experiments indicates that there may be more difference at the nucleic acid level than at the protein level in human and chimpanzee genomes. For every amino acid sequence difference observed, about four base differences are observed in the DNA. Li *et al.* (*25*) found the same distinction between amino acid and nucleic acid differences in the tryptophan synthetase of several bacterial species: the nucleic acid sequences were about three times as different as the amino acid sequences. A similar result has been observed in three related RNA bacteriophages, as well as in studies of the relative rates of DNA and protein evolution in cow, pig, and sheep (*26*).

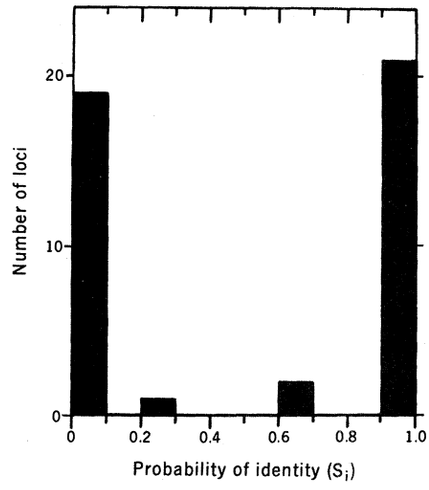


Fig. 2. Electrophoretic comparison of 43 proteins from humans and chimpanzees. The probability of identity (S_i) represents the likelihood that at locus i, human and chimpanzee alleles will appear electrophoretically identical.

There are a number of probable reasons for this discrepancy (*25*, *26*). First, more changes may appear in DNA than in proteins because of the redundancy of the code and consequently the existence of third-position nucleotide changes which do not lead to amino acid substitutions. The nature of the code indicates that if first-, second-, and third-position substitutions were equally likely to persist, then about 30 to 40 percent of potential base replacements in a cistron would not be reflected in the coded protein; that is, 1.4 to 1.7 base substitutions would occur for each amino acid substitution (*27*). However, it is likely that a larger proportion of the actual base substitutions in a cistron are third-position changes, since base substitutions that do not affect amino acid sequence are more likely to spread through a population. In addition, many of the nucleic acid substitutions may have occurred in regions of the DNA that are not transcribed and are therefore not conserved during evolution. Proteins analyzed by electrophoresis, sequencing, or microcomplement fixation techniques, on the other hand, all have definite cellular functions and may therefore have been conserved to a greater extent during evolution.

Genetic Distance and the Evolution of Organisms

The resemblance between human and chimpanzee macromolecules has been measured by protein sequencing, immunology, electrophoresis, and nucleic acid hybridization. From each of these results we can obtain an estimate of the genetic distance between humans and chimpanzees. Some of the same approaches have been used to estimate the genetic distance between other taxa, so that these estimates may be compared to the human-chimpanzee genetic distance.

First, we consider genetic distance estimated from electrophoretic data, using the standard estimate of net codon differences per locus developed by Nei and Roychoudhury (*28*). Other indices have been suggested for handling electrophoretic data (*29*) and give the same

qualitative results, though somewhat different underlying assumptions are required. Nei and Roychoudhury's standard estimate of genetic distance between humans and chimpanzees can be written:

$$D = D_{HC} - \frac{D_C + D_H}{2} \qquad (8)$$

where

$$D_{HC} = -\log_e \bar{S}$$

$$D_H = -\log_e \left(\frac{1}{L} \sum_{i=1}^{L} \sum_{j=1}^{A_i} x_{ij}^2 \right)$$

$$D_C = -\log_e \left(\frac{1}{L} \sum_{i=1}^{L} \sum_{j=1}^{A_i} y_{ij}^2 \right)$$

according to the notation of Table 2 and Eqs. 2 and 3. Therefore, D is an estimate of the variability between human and chimpanzee populations (D_{HC}), corrected for the variability within human populations (D_H) and within chimpanzee populations (D_C). D_C and D_H are also measurements of the degree of heterozygosity in human and chimpanzee populations (*30*). Based on the data of Table 2, D_{HC} is 0.65, D_C is 0.02, and D_H is 0.05, so that:

$$D = 0.62 \qquad (9)$$

In other words, there is an average of 0.62 electrophoretically detectable codon differences per locus between homologous human and chimpanzee proteins.

This distance is 25 to 60 times greater than the genetic distance between human races (*28, 31*). In fact, the genetic distance between Caucasian, Black African, and Japanese populations is less than or equal to that between morphologically and behaviorally identical populations of other species. In addition, these three human populations are equally distant from the chimpanzee lineage (Fig. 3).

However, with respect to genetic distances between species, the human-chimpanzee D value is extraordinarily small, corresponding to the genetic distance between sibling species of *Drosophila* or mammals (Fig. 4). Nonsibling species within a genus (referred to in the figure as congeneric species) generally differ more from each other, by electrophoretic criteria, than humans and chimpanzees. The genetic distances among species from different genera are considerably larger than the human-chimpanzee genetic distance.

The genetic distance between two species measured by DNA hybridization also indicates that human beings and chimpanzees are as similar as sibling species of other organisms. The difference in dissociation temperature, ΔT, between reannealed human DNA and human-chimpanzee hybrid DNA is about 1.1°C. However, for sibling species of *Drosophila*, ΔT is 3°C; for congeneric species of *Drosophila*, ΔT is 19°C; and for congeneric species of mice (*Mus*), ΔT is 5°C (*32*).

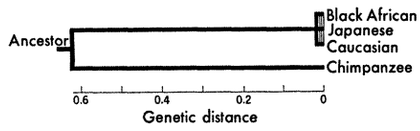
Immunological and amino acid sequence comparisons of proteins lead to the same conclusion. Antigenic differences among the serum proteins of congeneric squirrel species are several times greater than those between humans and chimpanzees (*33*). Moreover, antigenic differences among the albumins of congeneric frog species (*Rana* and *Hyla*) are 20 to 30 times greater than those between the two hominoids (*34, 35*). In addition, the genetic distances among *Hyla* species, estimated electrophoretically, are far larger than the chimpanzee-human genetic distance (*36*). Finally, the human and chimpanzee β chains of hemoglobin appear to have identical sequences (Table 1), while the β chains of two *Rana* species differ by at least 29 amino acid substitutions (*37*). In summary, the genetic distance between humans and chimpanzees is well within the range found for sibling species of other organisms.

The molecular similarity between chimpanzees and humans is extraordinary because they differ far more than sibling species in anatomy and way of life. Although humans and chimpanzees are rather similar in the structure of the thorax and arms, they differ substantially not only in brain size but also in the anatomy of the pelvis, foot, and jaws, as well as in relative lengths of limbs and digits (*38*). Humans and chimpanzees also differ significantly in many other anatomical respects, to the extent that nearly every bone in the body of a chimpanzee is readily distinguishable in shape or size from its human counterpart (*38*). Associated with these anatomical differences there are, of course, major differences in posture (see cover picture), mode of locomotion, methods of procuring food, and means of communication. Because of these major differences in anatomy and way of life, biologists place the two species not just in separate genera but in separate families (*39*). So it appears that molecular and organismal methods of evaluating the chimpanzee-human difference yield quite different conclusions (*40*).

An evolutionary perspective further illustrates the contrast between the results of the molecular and organismal approaches. Since the time that the ancestor of these two species lived, the chimpanzee lineage has evolved slowly relative to the human lineage, in terms of anatomy and adaptive strategy. According to Simpson (*41*):

> *Pan* is the terminus of a conservative lineage, retaining in a general way an anatomical and adaptive facies common to all recent hominoids *except Homo*. *Homo* is both anatomically and adaptively the most radically distinctive of all hominoids, divergent to a degree considered familial by all primatologists.

This concept is illustrated in the left-hand portion of Fig. 5. However, at the macromolecular level, chimpanzees and humans seem to have evolved



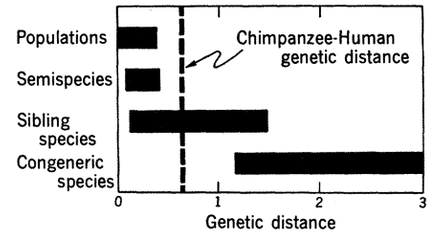


Fig. 3 (left). Phylogenetic relationship between human populations and chimpanzees. The genetic distances are based on electrophoretic comparison of proteins. The genetic distances among the three major human populations ($D = 0.01$ to 0.02) that have been tested are extremely small compared to those between humans and chimpanzees ($D = 0.62$). No human population is significantly closer than another to the chimpanzee lineage. The vertically hatched area between the three human lineages indicates that the populations are not really separate, owing to gene flow. Fig. 4 (right). The genetic distance, D, between humans and chimpanzees (dashed line) compared to the genetic distances between other taxa. Taxa compared include several species of *Drosophila* [*D. willistoni* (*94*), *D. paulistorum* (*95*), and *D. pseudoobscura* (*96*)], the horseshoe crab *Limulus polyphemus* (*97*), salamanders from the genus *Taricha* (*98*), lizards from the genus *Anolis* (*99*), the teleost fish *Astyanax mexicanus* (*100*), bats from the genus *Lasiurus* (*101*), and several genera of rodents [*Mus, Sigmodon, Dipodomys, Peromyscus,* and *Thomomys* (*99*), *Geomys* (*101*), and *Apodemus* (*102*)]. Selander and Johnson (*99*) summarize most of the data used in this figure. The great majority of proteins in these studies are intracellular.

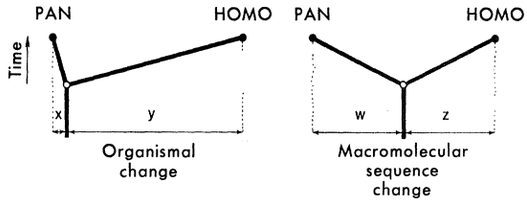


Fig. 5. The contrast between biological evolution and molecular evolution since the divergence of the human and chimpanzee lineages from a common ancestor. As shown on the left, zoological evidence indicates that far more biological change has taken place in the human lineage (y) than in the chimpanzee lineage ($y \gg x$); this illustration is adapted from that of Simpson (*41*). As shown on the right, both protein and nucleic acid evidence indicate that as much change has occurred in chimpanzee genes (w) as in human genes (z).

at similar rates (Fig. 5, right). For example, human and chimpanzee albumins are equally distinct immunologically from the albumins of other hominoids (gorilla, orangutan, and gibbon) (*10, 42, 43*), and human and chimpanzee DNA's differ to the same degree from DNA's of other hominoids (*21, 22*). Construction of a phylogenetic tree for primate myoglobins shows that the single amino acid difference between the sequences of human and chimpanzee myoglobin occurred in the chimpanzee lineage (*7*). Analogous reasoning indicates that the single amino acid difference between the sequences of human and chimpanzee hemoglobin δ chains arose in the human lineage (*8*). It appears that molecular change has accumulated in the two lineages at approximately equal rates, despite a striking difference in rates of organismal evolution. Thus, the major adaptive shift which took place in the human lineage was probably not accompanied by accelerated protein or DNA evolution.

Such an observation is by no means peculiar to the case of hominid evolution. It appears to be a general rule that anatomically conservative lineages, such as frogs, have experienced as much sequence evolution as have lineages that have undergone rapid evolutionary changes in anatomy and way of life (*34, 35, 44*).

Molecular Basis for the Evolution of Organisms

The contrasts between organismal and molecular evolution indicate that the two processes are to a large extent independent of one another. Is it possible, therefore, that species diversity results from molecular changes other than sequence differences in proteins? It has been suggested by Ohno (*45*) and others (*46*) that major anatomical changes usually result from mutations affecting the expression of genes. According to this hypothesis, small differences in the time of activation or in the level of activity of a single gene could in principle influence considerably the systems controlling embryonic development. The organismal differences between chimpanzees and humans would then result chiefly from genetic changes in a few regulatory systems, while amino acid substitutions in general would rarely be a key factor in major adaptive shifts.

Regulatory mutations may be of at least two types. First, point mutations could affect regulatory genes. Nucleotide substitutions in a promoter or operator gene would affect the production, but not the amino acid sequence, of proteins in that operon. Nucleotide substitutions in a structural gene coding for a regulatory protein such as a repressor, hormone, or receptor protein, could bring about amino acid substitutions, altering the regulatory properties of the protein. However, we suspect that only a minor fraction of the substitutions which accumulate in regulatory proteins would be likely to alter their regulatory properties.

Second, the order of genes on a chromosome may change owing to inversion, translocation, addition or deletion of genes, as well as fusion or fission of chromosomes. These gene rearrangements may have important effects on gene expression (*47*), though the biochemical mechanisms involved are obscure. Evolutionary changes in gene order occur frequently. Microscopic studies of *Drosophila* salivary chromosomes show, as a general rule, that no two species have the same gene order and that inversions are the commonest type of gene rearrangement (*48*). Furthermore, there is a parallel between rate of gene rearrangement and rate of anatomical evolution in the three major groups of vertebrates that have been studied in this respect, namely birds, mammals, and frogs (*46*). Hence gene rearrangements may be more important than point mutations as sources for evolutionary changes in gene regulation.

Although humans and chimpanzees have rather similar chromosome numbers, 46 and 48, respectively, the arrangement of genes on chimpanzee chromosomes differs from that on human chromosomes. Only a small proportion of the chromosomes have identical banding patterns in the two species. The banding studies indicate that at least 10 large inversions and translocations and one chromosomal fusion have occurred since the two lineages diverged (*49*). Further evidence for the possibility that chimpanzees and humans differ considerably in gene arrangement is provided by annealing studies with a purified DNA fraction. An RNA which is complementary in sequence to this DNA apparently anneals predominantly at a cluster of sites on a single human chromosome, but at widely dispersed sites on several chimpanzee chromosomes (*50*). The arrangement of chromosomal sites at which ribosomal RNA anneals may also differ between the two species (*50*).

Biologists are still a long way from understanding gene regulation in mammals (*51*), and only a few cases of regulatory mutations are now known (*52*). New techniques for detecting regulatory differences at the molecular level are required in order to test the hypothesis that organismal differences between individuals, populations, or species result mainly from regulatory differences. When the regulation of gene expression during embryonic development is more fully understood, molecular biology will contribute more significantly to our understanding of the evolution of whole organisms. Most important for the future study of human evolution would be the demonstration of differences between apes and humans in the timing of gene expression during development, particularly during the development of adaptively crucial organ systems such as the brain.

Summary and Conclusions

The comparison of human and chimpanzee macromolecules leads to several inferences:

1) Amino acid sequencing, immunological, and electrophoretic methods of protein comparison yield concordant estimates of genetic resemblance. These approaches all indicate that the average human polypeptide is more than 99 per-

cent identical to its chimpanzee counterpart.

2) Nonrepeated DNA sequences differ more than amino acid sequences. A large proportion of the nucleotide differences between the two species may be ascribed to redundancies in the genetic code or to differences in nontranscribed regions.

3) The genetic distance between humans and chimpanzees, based on electrophoretic comparison of proteins encoded by 44 loci is very small, corresponding to the genetic distance between sibling species of fruit flies or mammals. Results obtained with other biochemical methods are consistent with this conclusion. However, the substantial anatomical and behavioral differences between humans and chimpanzees have led to their classification in separate families. This indicates that macromolecules and anatomical or behavioral features of organisms can evolve at independent rates.

4) A relatively small number of genetic changes in systems controlling the expression of genes may account for the major organismal differences between humans and chimpanzees. Some of these changes may result from the rearrangement of genes on chromosomes rather than from point mutations (*53*).

References and Notes

1. S. L. Washburn, Ed., *Classification and Human Evolution* (Aldine, Chicago, 1963). That there were striking similarities in blood proteins between apes and humans was known in 1904 [G. H. F. Nuttall, *Blood Immunity and Blood Relationships* (Cambridge Univ. Press, London, 1904)].
2. Sibling species are virtually identical morphologically.
3. R. F. Doolittle, G. L. Wooding, Y. Lin, M. Riley, *J. Mol. Evol.* **1**, 74 (1971).
4. M. O. Dayhoff, Ed., *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Georgetown Univ. Medical Center, Washington, D.C., 1972), vol. 5.
5. S. H. Boyer, E. F. Crosby, A. N. Noyes, G. F. Fuller, S. E. Leslie, L. J. Donaldson, G. R. Vrablik, E. W. Schaefer, T. F. Thurmon, *Biochem. Genet.* **5**, 405 (1971).
6. W. W. W. DeJong, *Biochim. Biophys. Acta* **251**, 217 (1971).
7. A. E. Romero Herrera and H. Lehmann, *ibid.* **278**, 62 (1972).
8. W. W. W. DeJong, *Nat. New Biol.* **234**, 176 (1971).
9. E. M. Prager and A. C. Wilson, *J. Biol. Chem.* **246**, 5978 and 7010 (1971).
10. V. M. Sarich and A. C. Wilson, *Science* **158**, 1200 (1967).
11. A. C. Wilson and V. M. Sarich, *Proc. Natl. Acad. Sci. U.S.A.* **63**, 1088 (1969); J. E. Cronin and V. M. Sarich, personal communication; R. Palmour, personal communication.
12. L. Nonno, H. Herschman, L. Levine, *Arch. Biochem. Biophys.* **136**, 361 (1970).
13. N. Hanke, E. M. Prager, A. C. Wilson, *J. Biol. Chem.* **248**, 2824 (1973).
14. A variety of immunological techniques have been used to compare chimpanzee proteins with their human counterparts [N. Mohagheghpour and C. A. Leone, *Comp. Biochem. Physiol.* **31**, 437 (1969); M. Goodman and G. W. Moore, *Syst. Zool.* **20**, 19 (1971); K. Bauer, *Humangenetik* **17**, 253 (1973)]. The immunodiffusion techniques employed in these studies are less sensitive to small differences in amino acid sequence than is microcomplement fixation [E. M. Prager and A. C. Wilson, *J. Biol. Chem.* **246**, 5978 (1971)]. Nevertheless, their results are generally consistent with those in Table 1. The few cases of large antigenic differences between human and chimpanzee proteins are probably not indicative of large sequence differences. For example, the haptoglobin difference reported by Mohagheghpour and Leone is due mainly to the fact that the haptoglobin 2 polypeptide is nearly twice the length of the haptoglobin 1 polypeptide [J. A. Black and G. H. Dixon, *Nature (Lond.)* **218**, 736 (1968)]. Human haptoglobin 1 is immunologically very similar to chimpanzee haptoglobin [J. Javid and H. H. Fuhrmann, *Am. J. Hum. Genet.* **207**, 496 (1971)]. The immunoglobulin differences reported by Bauer may be due to comparison of peptide chains that are not strictly homologous. In addition, Bauer's kappa chain results conflict with quantitative studies which detected no immunological difference [A. C. Wang, J. Shuster, A. Epstein, H. H. Fudenberg, *Biochem. Genet.* **1**, 347 (1968)]. Finally, the large Xh factor difference that Bauer reported might result from the fact that the chimpanzees in his studies were not pregnant and thus lacked Xh factor.
15. M.-C. King, thesis, University of California, Berkeley (1973).
16. D. Stollar and L. Levine, in *Methods in Enzymology*, S. P. Colowick and N. O. Kaplan, Eds. (Academic Press, New York, 1967), vol. 6, p. 928.
17. F. J. Ayala, M. L. Tracey, L. G. Barr, J. F. McDonald, S. Perez-Salas, *Genetics* **77**, 343 (1974).
18. See figure 9.3 in Dayhoff (*4*). We determined the proportion of amino acid substitutions causing a charge change during vertebrate evolution for several additional proteins: cytochrome c, lysozyme, myoglobin, α and β hemoglobin chains, triosephosphate dehydrogenase, α-lactalbumin, growth hormone, trypsin, and insulin. The average for these proteins is also 0.27. Our estimate of 0.16 for hemoglobin alone is very similar to that of Boyer *et al.* [S. H. Boyer, A. N. Noyes, C. F. Timmons, R. A. Young, *J. Hum. Evol.* **1**, 515 (1972)], who calculated that the ratio between electrophoretically silent and electrophoretically detectable hemoglobin alleles in primates is about 5.5; that is, about 15 percent of amino acid substitutions in primate hemoglobin evolution would be electrophoretically detectable.

    A change in charge at a single amino acid site may have little or no effect on the net charge of a protein unless the substituted amino acid is on the exposed surface of the protein. Lee and Richa[illegible] [B. Lee and F. M. Richards, *J. Mol. Biol.* **55**, 379 (1971)] determined the degree of exposure of each of the amino acid residues of lysozyme, ribonuclease, and myoglobin, based on the three-dimensional structure of these molecules. Their data indicate that 100 percent of the lysine residues, 100 percent of the arginine residues, 95 percent of the aspartic acid residues, 100 percent of the glutamic acid residues, and 70 percent of the histidine residues are on exposed surfaces of the proteins. Thus more than 90 percent of the substitutions involving charged amino acids would have affected the net charge of the protein and would, therefore, be detectable by electrophoresis.
19. A negative binomial variable may better describe the distribution of amino acid substitutions along lineages, since substitutions occur in proteins which are subject to varying selective pressures. That is, since different proteins evolve at different rates, the probability of a particular protein accepting a mutation varies from protein to protein [T. Uzzell and K. W. Corbin, *Science* **172**, 1089 (1971)]. However, for small values of m the negative binomial distribution agrees substantially with the Poisson distribution [C. I. Bliss and R. A. Fisher, *Biometrics* **9**, 176 (1953)]. Thus, for this calculation, the Poisson distribution should provide a very good estimate of the true m.
20. Further evidence regarding the correlation between electrophoretic and immunological measures of genetic distance has been obtained in studies involving many taxa of mammals, reptiles, and amphibians, by S. M. Case, H. C. Dessauer, G. C. Gorman, P. Haneline, K. H. Keeler, L. R. Maxson, V. M. Sarich, D. Shochat, A. C. Wilson, and J. S. Wyles.
21. W. M. Brown and R. L. Hallberg, *Fed. Proc.* **31** (2), Abstr. 1173 (1972).
22. D. E. Kohne, *Q. Rev. Biophys.* **3**, 327 (1970); ———, J. A. Chiscon, B. H. Hoyer, *J. Hum. Evol.* **1**, 627 (1972).
23. B. H. Hoyer, N. W. van de Velde, M. Goodman, R. B. Roberts, *J. Hum. Evol.* **1**, 645 (1972).
24. N. R. Kallenbach and S. D. Drost, *Biopolymers* **11**, 1613 (1972); J. R. Hutton and J. G. Wetmur, *Biochemistry* **12**, 558 (1973); J. S. Ullman and B. J. McCarthy, *Biochim. Biophys. Acta* **294**, 416 (1973).
25. S. L. Li, R. M. Denney, C. Yanofsky, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1112 (1973).
26. H. D. Robertson and P. G. N. Jeppesen, *J. Mol. Biol.* **68**, 417 (1972); C. Laird, B. L. McConaughy, B. J. McCarthy, *Nature (Lond.)* **224**, 149 (1969).
27. *Cold Spring Harbor Symp. Quant. Biol.* **31**, 1 (1966).
28. M. Nei and A. K. Roychoudhury, *Am. J. Hum. Genet.* **26**, 431 (1974).
29. L. L. Cavalli-Sforza and A. W. F. Edwards, *ibid.* **19**, 233 (1967); N. E. Morton, *Annu. Rev. Genet.* **3**, 53 (1969); V. Balakrishnan and L. D. Sangvhi, *Biometrics* **24**, 859 (1968); T. W. Kurczynski, *ibid.* **26**, 525 (1970); P. W. Hedrick, *Evolution* **25**, 276 (1971); R. R. Sokal and P. H. A. Sneath, *Principles of Numerical Taxonomy* (Freeman, San Francisco, 1973); J. S. Rogers, "Studies in Genetics," *Univ. Texas Publ. No. 7213* (1972), vol. 7, p. 145.
30. The average heterozygosity estimates for the loci in this study are quite low, especially for chimpanzee populations. To obtain comparable heterozygosity estimates for humans and chimpanzees, we included only loci that have been surveyed for both species and only polymorphisms detectable by techniques used for surveying both species. Thus some confirmed polymorphisms in human populations were excluded. There are at least three reasons for the difference between the heterozygosity estimates for human and chimpanzee populations. First, many more humans than chimpanzees have been surveyed at each locus, so that the variability estimate for humans is biased insofar as it is based on alleles present at low frequency in human populations. Second, there are many more humans than chimpanzees alive today, living in a greater variety of environments and with a larger number of gene pools. As a result, more mutants reach appreciable frequencies in human populations. Third, and probably most important, the chimpanzees in colonies available for study are based on even fewer gene pools and are highly inbred in many cases. The discrepancies in real population size and sampling technique between human and chimpanzee populations probably account for the greater number of polymorphic loci, the larger number of alleles at polymorphic loci, and the higher average heterozygosity estimates in human populations.
31. M. Nei and A. K. Roychoudhury, *Science* **177**, 434 (1972).
32. N. R. Rice, *Brookhaven Symp. Biol.* **23**, 44 (1972); C. D. Laird, *Annu. Rev. Genet.* **3**, 177 (1973).
33. M. E. Hight, M. Goodman, W. Prychodko, *Syst. Zool.* **23**, 12 (1974).
34. D. G. Wallace, L. R. Maxson, A. C. Wilson, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 3127 (1971); D. G. Wallace, M.-C. King, A. C. Wilson, *Syst. Zool.* **22**, 1 (1973).
35. L. R. Maxson, and A. C. Wilson, *Syst. Zool.*, in press.
36. R. K. Selander, personal communication; L. R. Maxson and A. C. Wilson, *Science* **185**, 66 (1974).
37. T. O. Baldwin and A. Riggs, *J. Biol. Chem.* **249**, 6110 (1974).
38. G. H. Bourne, Ed., *The Chimpanzee* (Karger, New York, 1970).
39. Th. Dobzhansky, in *Classification and Human Evolution*, S. L. Washburn, Ed. (Aldine, Chicago, 1963), p. 347; E. Mayr, in *ibid.*, p. 332; E. L. Simons, *Primate Evolution* (Mac-

millan, New York, 1972); G. G. Simpson, *Principles of Animal Taxonomy* (Columbia Univ. Press, New York, 1961). L. Van Valen [*Am. J. Phys. Anthropol.* **30**, 295 (1969)] has suggested that, based on differences in their adaptive zones, humans and chimpanzees be placed in separate suborders.
40. On the basis of some protein evidence available in 1970, Goodman and Moore proposed that humans and African apes be placed in the same subfamily [M. Goodman and G. W. Moore, *Syst. Zool.* **20**, 19 (1971)]. By analogy, the protein evidence now available would lead to placement of chimpanzees and humans in the same genus. However, as protein evolution and organismal evolution apparently can proceed independently, it is questionable whether organismal classifications should be revised on the basis of protein evidence alone.
41. G. G. Simpson, in *Classification and Human Evolution*, S. L. Washburn, Ed. (Aldine, Chicago, 1963).
42. V. M. Sarich, in *Old World Monkeys*, J. R. Napier and P. H. Napier, Eds. (Academic Press, New York, 1970), p. 175.
43. ——— and A. C. Wilson, *Proc. Natl. Acad. Sci. U.S.A.* **58**, 142 (1967).
44. V. M. Sarich, *Syst. Zool.* **18**, 286 and 416 (1969); *Nature (Lond.)* **245**, 218 (1973).
45. S. Ohno, *J. Hum. Evol.* **1**, 651 (1972).
46. A. C. Wilson, L. R. Maxson, V. M. Sarich, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 2843 (1974); A. C. Wilson, V. M. Sarich, L. R. Maxson, *ibid.*, p. 2028; E. M. Prager and A. C. Wilson, *ibid.* **72**, 200 (1975).
47. E. Bahn, *Hereditas* **67**, 79 (1971); B. Wallace and T. L. Kass, *Genetics* **77**, 541 (1974).
48. M. J. D. White, *Annu. Rev. Genet.* **3**, 75 (1969).
49. J. de Grouchy, C. Turleau, M. Roubin, F. C. Colin, *Nobel Symp.* **23**, 124 (1973); B. Dutrillaux, M.-O. Rethoré, M. Prieur, J. Lejeune, *Humangenetik* **20**, 343 (1973); D. Warburton, I. L. Firschein, D. A. Miller, F. E. Warburton, *Cytogenet. Cell Genet.* **12**, 453 (1973); C. C. Lin, B. Chiarelli, L. E. M. de Boer, M. M. Cohen, *J. Hum. Evol.* **2**, 311 (1973); J. Ecozcue, M. R. Caballin, C. Goday, *Humangenetik* **18**, 77 (1973); M. Bobrow and K. Madan, *Cytogenet. Cell Genet.* **12**, 107 (1973).
50. In situ annealing studies have been performed with RNA complementary to purified human satellite DNA [K. W. Jones, J. Prosser, G. Carneo, E. Ginelli, M. Bobrow, *Symp. Med. Hoechst* **6**, 45 (1973)] and with human ribosomal RNA [A. Henderson, D. Warburton, K. C. Atwood, *Chromosoma* **46**, 435 (1974)].
51. J. E. Darnell, W. R. Jelinek, G. R. Molloy, *Science* **181**, 1214 (1973); E. H. Davidson and R. J. Britten, *Q. Rev. Biol.* **48**, 565 (1973); C. A. Thomas, Jr., in *Regulation of Transcription and Translation in Eukaryotes*, E. K. F. Bautz, Ed. (Springer-Verlag, Berlin, 1973).
52. D. J. Weatherall and J. B. Clegg, *The Thalassaemia Syndromes* (Blackwell, Oxford, ed. 2, 1972).
53. Additional inferences can be drawn from the comparison of human and chimpanzee macromolecules; some of these will be discussed elsewhere.
54. I. N. H. White and P. J. Butterworth, *Biochim. Biophys. Acta* **229**, 193 (1971).
55. J. Schmitt, K. H. Lichte, W. Fuhrmann, *Humangenetik* **10**, 138 (1970); G. Tariverdian, H. Ritter, J. Schmitt, *ibid.* **11**, 323 (1971).
56. E. R. Giblett, *Genetic Markers in Human Blood* (Blackwell, Oxford, 1969).
57. J. Schmitt, G. Tariverdian, H. Ritter, *Humangenetik* **11**, 100 (1971).
58. R. E. Tashian, *Am. J. Hum. Genet.* **17**, 257 (1965).
59. H. Matsubara and E. L. Smith, *J. Biol. Chem.* **237**, 3575 (1962).
60. R. Schiff and C. Stormont, *Biochem. Genet.* **4**, 11 (1970).
61. A. Yoshida, *Proc. Natl. Acad. Sci. U.S.A.* **57**, 835 (1967); *Biochem. Genet.* **1**, 81 (1967); J. Kömpf, H. Ritter, J. Schmitt, *Humangenetik* **11**, 342 (1971); M. Goodman and M. D. Poulik, "Genetic variations and phylogenetic properties of protein macromolecules of chimpanzees" (6571st Aeromedical Research Laboratory, Rep. ARL-TR-68-3, Holloman Air Force Base, New Mexico, 1968).
62. J. L. Brewbaker, M. D. Upadhya, Y. Mäkinen, T. Macdonald, *Physiol. Plant* **21**, 930 (1968); C. R. Shaw and R. Prasad, *Biochem. Genet.* **4**, 297 (1970).
63. J. Kömpf, H. Ritter, J. Schmitt, *Humangenetik* **13**, 72 (1971).
64. J. C. Kaplan, *Nature (Lond.)* **217**, 256 (1968); W. K. Long, *Science* **155**, 712 (1967).
65. D. Rifkin and W. Konigsberg, *Biochim. Biophys. Acta* **104**, 457 (1965).
66. H. Harris, *J. Med. Genet.* **8**, 444 (1971).
67. R. G. Davidson, J. A. Cortner, M. C. Rattazzi, F. H. Ruddle, H. A. Lubs, *Science* **196**, 391 (1970).
68. G. Beckman, E. Lundgren, A. Tärnvik, *Hum. Hered.* **23**, 338 (1973); B. B. Keele, Jr., J. M. McCord, I. Fridovich, *J. Biol. Chem.* **246**, 2875 (1971); H. M. Steinman and R. L. Hill, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3725 (1973).
69. G. S. Bailey and A. C. Wilson, *J. Biol. Chem.* **243**, 5843 (1968).
70. A. L. Koen and M. Goodman, *Biochem. Genet.* **3**, 457 (1969).
71. R. L. Kirk, E. M. McDermid, N. M. Blake, R. L. Wight, E. H. Yap, M. J. Simons, *Humangenetik* **17**, 345 (1973); G. Tariverdian, H. Ritter, J. Schmitt, *ibid.* **11**, 339 (1971).
72. G. Tariverdian, H. Ritter, G. G. Wendt, *ibid.*, p. 75.
73. J. Schmitt, G. Tariverdian, H. Ritter, *ibid.*, p. 95.
74. P. C. Hudgins, C. M. Whorton, T. Tomoyoshi, A. J. Riopelle, *Nature (Lond.)* **212**, 693 (1966).
75. W. H. P. Lewis, *Ann. Hum. Genet.* **36**, 267 (1973).
76. S. Povey, G. Corney, W. H. P. Lewis, E. B. Robson, J. M. Parrington, H. Harris, *ibid.* **35**, 455 (1972).
77. M. Goodman and R. E. Tashian, *Hum. Biol.* **41**, 237 (1969).
78. G. Tariverdian, H. Ritter, J. Schmitt, *Humangenetik* **12**, 105 (1971).
79. H. Rubinson, M. C. Meienhofer, J. C. Dreyfus, *J. Mol. Evol.* **2**, 243 (1973); P. H. Corran and S. G. Waley, *Biochem. J.* **139**, 1 (1974).
80. W. E. Marshall, *J. Biol. Chem.* **241**, 4731 (1966).
81. Y. T. Li and S. C. Li, *ibid.* **245**, 825 (1970).
82. A. L. Tarnoky, B. Dowding, A. L. Lakin, *Nature (Lond.)* **225**, 742 (1970).
83. G. Kellermann and H. Walter, *Humangenetik* **10**, 145 (1970); G. Kellermann and H. Walter, *ibid.*, p. 191.
84. M. H. K. Shokeir and D. C. Shreffler, *Biochem. Genet.* **4**, 517 (1970).
85. C. A. Alper and F. S. Rosen, *Immunology* **14**, 251 (1971); E. A. Azen, O. Smithies, O. Hiller, *Biochem. Genet.* **3**, 214 (1969).
86. H. Cleve, *Hum. Hered.* **20**, 438 (1970); F. D. Kitchin and A. G. Bearn, *Am. J. Hum. Genet.* **17**, 42 (1965).
87. B. S. Blumberg, *Proc. Soc. Exp. Biol. Med.* **104**, 25 (1960); S. H. Boyer and W. J. Young, *Nature (Lond.)* **187**, 1035 (1960); M. Cresta, *Riv. Antrop. Roma* **47**, 225 (1961); V. Lange and J. Schmitt, *Folia Primatol.* **1**, 208 (1963); O. Mäkelä, O. V. Rekonen, E. Salonen, *Nature (Lond.)* **185**, 852 (1960); W. C. Parker and A. G. Bearn, *Ann. Hum. Genet.* **25**, 227 (1961); J. Javid and M. H. Fuhrman, *Am. J. Hum. Genet.* **23**, 496 (1971); B. S. Shim and A. G. Bearn, *ibid.* **16**, 477 (1964); J. Planas, *Folia Primatol.* **13**, 177 (1970).
88. Although human and chimpanzee lysozymes have been reported to be electrophoretically identical [N. Hanke, E. M. Prager, A. C. Wilson, *J. Biol. Chem.* **248**, 2824 (1973)], more refined techniques indicate that their mobilities in fact differ (E. Prager, personal communication).
89. K. Berg and A. G. Bearn, *Annu. Rev. Genet.* **2**, 341 (1968).
90. M. Goodman, R. McBride, E. Poulik, E. Reklys, *Nature (Lond.)* **197**, 259 (1963); M. Goodman and A. J. Riopelle, *ibid.*, p. 261; M. Goodman, W. G. Wisecup, H. H. Reynolds, C. H. Kratochvil, *Science* **150**, 98 (1967).
91. H. Harris, *The Principles of Human Biochemical Genetics* (Elsevier, New York, 1970).
92. H. E. Schultze and J. F. Heremans, *Molecular Biology of Human Proteins* (Elsevier, New York, 1966), vol. 1.
93. M.-C. King and A. C. Wilson, in preparation.
94. F. J. Ayala, J. R. Powell, M. L. Tracey, C. A. Mourão, S. Pérez-Salas, *Genetics* **70**, 113 (1972).
95. R. C. Richmond, *ibid.*, p. 87.
96. S. Prakash, R. C. Lewontin, J. L. Hubby, *ibid.* **61**, 841 (1969).
97. R. K. Selander, S. Y. Yang, R. C. Lewontin, W. E. Johnson, *Evolution* **24**, 402 (1970).
98. D. Hedgecock and F. J. Ayala, *Copeia* (1974), p. 738.
99. R. K. Selander and W. E. Johnson, *Annu. Rev. Ecol. Syst.* **4**, 75 (1973).
100. C. R. Shaw, *Biochem. Genet.* **4**, 275 (1970).
101. R. K. Selander, D. W. Kaufman, R. J. Baker, S. L. Williams, *Evolution* **28**, 557 (1974).
102. W. Engel, W. Vogel, I. Voiculescu, H. Ropers, M. T. Zenges, K. Bender, *Comp. Biochem. Physiol.* **44B**, 1165 (1973).
103. Samples of chimpanzee blood for this study were obtained from the Laboratory for Experimental Medicine and Surgery in Primates, New York University Medical Center, P.O. Box 575, Tuxedo, N.Y. 10987; M. Goodman, Wayne State University School of Medicine, Detroit, Mich. 48201; and H. Hoffman, National Institutes of Health, Bethesda, Md. 20014. This work was supported by grant GM-18578 from NIH. Many colleagues helped us with this project. We thank S. S. Carlson, R. K. Colwell, L. R. Maxson, J. Maynard Smith, E. M. Prager, V. M. Sarich, and G. S. Sensabaugh for advice and ideas; M. Nei for unpublished data; and E. Bradley, D. Healy, P. Lozar, K. Pippen, and R. Wayner for expert technical assistance.