Jpn. J. Genet. (1993) **68**, *pp.* 539–547

SPECIAL FEATURE

On the Occasion of the 25th Anniversary of the Neutral Theory (II)

Relaxed natural selection in human populations during the Pleistocene

Naoyuki TAKAHATA

Department of Genetics, The Graduate University for Advanced Studies, Mishima 411, Japan

(*Received 29 November 1993*)

ABSTRACT

Available genetic data reveals that the human population is more variable than the chimpanzee population at the protein level, whereas the opposite is the case at the DNA level. The lower level of silent polymorphism in the human population suggests that its long-term breeding size is smaller than the chimpanzee's. The neutral theory suggests that natural selection has been relaxed in the human population under the improved environment. The possibility that the relaxation began with the emergence of *Homo sapiens* is examined, because it is known that *H. habilis* underwent for the first time dramatic changes in brain size, way of life, and culture, and that the childhood of *H. erectus* was already twice as long as that of chimpanzee. The relaxation hypothesis predicts that, relative to chimpanzee, some 20% of deleterious mutations became harmless under the changed environment throughout the Pleistocene. More extensive study of genetic variation in non-human primates is necessary not only to confirm the hypothesis, but also to better understand the human genome itself.

1. ELECTROPHORETIC DATA

Electrophoresis detects about one-third of amino acid changes in proteins, but not silent or synonymous changes. The past 25 years or so witnessed its wide application to various problems in evolutionary biology. A puzzling finding is that the extent of genetic variation in human is highest among primates, or even higher than that in most vertebrates (e.g., Nei and Graur, 1984). The probability that two randomly chosen orthologous proteins are different in charge profile, or the so-called average heterozygosity (H), is 14.3% over 121 loci of human. The value is much higher than that of non-human primates, which is 1.3% over 43 loci of chimpanzee and 4.6% over 22 loci of Gorilla (King and Wilson, 1975; Bruce and Ayala, 1979), and as high as that of some *Drosophila* species (see Nei and Graur, 1984 and references therein). These concordance and discordance may occur simply by chance if different loci are used in different organisms, because the

degree of polymorphism at the amino acid level differs greatly from locus to locus. Another possible cause is different per-year mutation rates among different taxonomic groups. The mutation rate may well depend on reproductive ages and/or physiological conditions (Wu and Li, 1985; Martin and Palumbi, 1993). To avoid these uncertainties, it is sensible to compare H for the same set of loci among relatively closely related primates. Chimpanzee, the closest relative to human (Horai et al., 1992), is clearly the best organism to be compared in the present context. There are 35 loci that were examined commonly to both species with reasonably large sample sizes (King and Wilson, 1975). The average H value is 8.0% for human and 2.3% for chimpanzee (Table 1).

It is possible to evaluate the fraction of neutral mutations, f_E, which contribute to H. One method developed by Kimura (1983) uses rare variant alleles. They are alleles whose frequencies are smaller than a pre-assigned value q, usually

Table 1. Genetic diversity of human and chimpannzee

	human	chimpanzee	ratio*
	35 loci (commonly surveyed by electrophoresis)[§]		
H	0.080 (0.143[†])	0.023 (0.013[†])	3.2 (12.6)
f_E	0.34	0.21	
	mtDNA		
π_S (noncoding)	0.017	0.078	0.22
π_R (restriction)	0.004	0.013	0.31
π_R/π_S	0.24	0.17	
	Nuclear loci[††]		
π_N (45,656 bp)	0.0002	NA	-
π_S (19,324 bp)	0.0008	NA	-
$f_N = \pi_N/\pi_S$	0.31	-	-
	ABO glycosyltransferases[¶]		
π_N (270 bp)	0.015	0.015	-
π_S (135 bp)	0.007	0.030	-
$f_N = \pi_N/\pi_S$	2	0.5	-

* The ratio of human genetic diversity H or π to that of chimpanzee (in the same row). [§]Loci at which human exhibits more variation than chimpanzee are: acid phosphatase ($H = 0.453$ vs. 0), α_1-acid glycoprotein (0.435 vs. 0), third component of compliment (0.229 vs. 0), group-specific component (0.385 vs. 0), and haptoglobin α chain (0.461 vs. 0). The only exception is the transferrin locus at which chimpanzee H is 0.477. Excluded are esterase A2 and eight other loci listed in Table 2 in King and Wilson (1975). [†]The H value in parenthesis is taken from Nei and Graur (1984). NA, not available. [††]Taken from Li and Sadler (1991). [¶]After Martinko et al. (1992). The values in the row "chimpanzee" are actually obtained by comparing chimpanzee A and gorilla B allele sequences. The subscript stands for electrophoresis (E), restriction analysis (R), sequence differences at silent or synonymous sites (S) and nonsynonymous sites (N).

taken as 0.01 or less. Because of their rarity, they are less likely to be subject to natural selection. The average sample size of individuals for the 35 loci is 111. The sample size is relatively small and the value of q chosen as 0.05 is relatively large. Nonetheless, these values will not much affect the conclusion because the number of rare variant alleles of $q \leq 0.01$ is also much larger in human than in chimpanzee (11 vs. 3; see King and Wilson, 1975). The estimated f_E is 43% for human and 21% for chimpanzee (Table 1). The rare-allele method was previously applied to much larger data sets for two ethnic groups, giving that $f_E = 14\%$ for European and 21% for American (Kimura, 1983). Both estimates are smaller than 43%, but this is again probably due to different loci examined. In any event, all this indicates that the fraction of neutral alleles is larger in human than in chimpanzee and that the remaining 57% and 79% or more of mutations are definitely deleterious in human and chimpanzee, respectively. Hence, both H and f_E show that the human population is more variable than the chimpanzee population at the protein level.

2. DNA SEQUENCE DATA

The DNA sequences of alleles at 49 loci in human were compared (Li and Sadler, 1991). A commonly used statistic is the nucleotide diversity, π (the number of differences per nucleotide site between two randomly chosen DNA sequences from a population). The π is 0.04% in the whole coding region and 0.11% in the noncoding region. In the coding region, there are 15 nonsynonymous (amino acid replacement) changes in total and 11 synonymous changes, corresponding to $\pi_N = 0.02\%$ per nonsynonymous site and $\pi_S = 0.08\%$ per synonymous site. The ratio of π_N/π_S, denoted by f_N, is a measure of the average degree of neutrality for nonsynonymous changes. The estimated f_N is 31%, indicating that about 70% of nonsynonymous changes at these 49 loci are selected against, relative to synonymous changes. The f_N value is smaller than $f_E = 43\%$, but larger than 14% and 21% for the ethnic groups. If electrophoresis detects one-third of nonsynonymous changes as mentioned earlier, the expected relationship may be given approximately by $f_N = 3f_E$. However, the extent of nonsynonymous polymorphism again depends strongly on genes examined, so that the expected relationship between f_N and f_E may not be necessarily warranted.

There is a striking difference in the value of f_N between human and *Drosophila*. Except human, *Drosophila* is one of a few genera whose genomes have been extensively examined for the DNA polymorphism. Eleven *D. melanogaster Adh* sequences (Kreitman, 1983) show that there is only one nonsynonymous change but the extent of synonymous diversity ($\pi_S = 0.6\%$) is much higher than that of human. A large scale survey of DNA polymorphism by restriction enzyme analysis shows that the π_R around 20 loci on the *D. melanogaster* genome ranges from 0% to 0.9%, the average being 0.34% (Aquadro and Begun, 1992). A

general feature is that the π_N and f_N in *Drosophila* is fairly low. These low values are often interpreted by negative selection against slightly deleterious mutations which is effective in a large population (see next section).

Unfortunately, there are no such comparable data for chimpanzee and any other non-human primates. Exceptions may be *ABO glycosyltransferases* (Martinko et al., 1993) and mitochondrial (mt) DNA (Aquadro and Greenberg, 1983). Partial DNA sequences 405 bp long are available for human *A* and *B* alleles and there are four nonsynonymous and one synonymous differences ($f_N = 2$ in Table 1). Three chimpanzee *A* and two gorilla *B* alleles are sequenced. Because the B antigen is absent in chimpanzee and the A antigen is absent in gorilla, we may compare chimpanzee *A* and gorilla *B* alleles. There are four synonymous and four nonsynonymous changes so that $f_N = 0.5$. Although information is limited and it is likely that some of these nonsynonymous changes are subject to positive selection (Martinko et al., 1993), the large f_N between human *A* and *B* allele is along the line of expectation. Secondly, restriction enzyme analysis of primate mtDNA shows that $\pi_R = 0.4\%$ in human, 1.3% in chimpanzee, 0.6% in gorilla (Aquadro and Greenberg, 1983). These π_R values are largely due to silent changes. For the noncoding region of 1135 bp length, DNA sequence data can be used to estimate π_S; 1.7% averaged over 14 human mtDNAs and 7.8% averaged over three chimpanzee mtDNAs (Kocher and Wilson, 1991). Because this region is known to evolve much faster than the coding region, the large π_S value is expected. The point is that both restriction enzyme and DNA sequence analyses consistently reveal that the extent of silent polymorphism of mtDNA in human is one-fourth to one-third of that in chimpanzee. This conclusion would be strengthened if the mutation rate per generation has become high in the human lineage, because of the prolonged generation time since the evolution of *H. erectus*.

Since most of silent polymorphism, if not all, is free from natural selection, the neutral theory (Kimura, 1968) suggests that the breeding size, N, of human has been one-fourth to one-third of that of chimpanzee for the period during which the currently observed mtDNA variations were generated. This time period probably corresponds to 200,000–400,000 years, if the long-term female breeding size is 5,000 and the generation time is 20 years (Takahata, 1993). In the case of nuclear DNA, the expected persistence time of polymorphism is four times longer and therefore amounts to the period of the Pleistocene epoch (lasting from about 2 million years ago to 10,000 years ago). This is because there is an approximately four-fold difference in the number of nuclear and mitochondrial genomes in a population.

3. DELETERIOUS MUTATIONS

One hypothesis of molecular evolution assumes that the majority of nonsynonymous changes are slightly deleterious rather than strictly neutral (Ohta, 1975).

To explain, we denote by s the selection intensity against deleterious mutations. These are operationally defined as slightly deleterious, if the population dynamics is similar to neutral in a small population of size N_h ($N_h s < 1$). In a large population of size N_c ($N_c s > 1$), these mutations have little or no contribution to molecular evolution. However, in order for a large H value in a smaller population to be accounted for, the inequality $N_h s < 1 < N_c s$ must be met (subsequently, the subscript h and c stand for human and chimpanzee). Let v_d be the rate of slightly deleterious mutations whose s values are specified by the above inequality. Thus, these mutations are neutral in the human but deleterious in the chimpanzee population. Let v_n be the neutral mutation rate with $N_c s < 1$ and v_T be $v_d + v_n$.

Under neutrality, the expected H in a population of size N and with neutral mutation rate v is given by

$$H = \frac{4Nv}{1+4Nv} \tag{1}$$

(Kimura and Crow, 1964), or $4Nv = H/(1-H)$. From the two equations corresponding to human and chimpanzee, we define the ratio of

$$\frac{N_c v_n}{N_h v_T} = \frac{(1-H_h)H_c}{H_h(1-H_c)} . \tag{2}$$

Since $H_h = 0.080$ and $H_c = 0.023$, the ratio becomes 0.27. If $N_c = 3 - 4 \times N_h$, as suggested by the silent polymorphism, the expected ratio of v_n/v_T becomes about 8%. In other words, 92% of electrophoretically detectable mutations are required to be slightly deleterious. The value of s must range from $1/(4N_h)$ to $1/N_h$. If N_h is about 10^4 (Nei and Graur, 1984; Takahata, 1993), most electrophoretic variants must occur with s being in the range of $2.5 \times 10^{-5} < s < 10^{-4}$.

We can evaluate how many such mutations have accumulated in a gene lineage over the past 1 million years. Assume that the average gene locus is 900 bp long and the mutation rate is 10^{-9} per year per nucleotide site (Klein et al., 1993), to be conservative. Assume also that 2/3 of the nucleotide sites are nonsynonymous and there are 600 such sites per locus. If $f_E = 43\%$ of the total mutations are neutral in the human population, of which 92% are actually slightly deleterious, the total number of slightly deleterious mutations is $10^{-9} \times 600 \times 0.43 \times 0.92 \times 106 = 0.24$ per gene lineage per 1 million years. An index of deterioration of this gene is defined as $I = (1-s)^{0.24} \approx e^{-0.24s}$, which is small. However, if there are 10^5 loci in the human genome, the index of genome deterioration as a whole is raised to 10^5 power, the total number of slightly deleterious mutations being 24,000 per genome per 1 million years. For $s = 5 \times 10^{-5}$, the value is $e^{-1.2} = 0.30$ per genome. The fitness of human individuals carrying two such genomes is $e^{-2.4} = 0.09$ as compared with that of chimpanzees, or the average fitness of modern humans has reduced to less than 10%. This figure appears to be

intolerably low for the reproductive capacity of human females, and if this were the case, the population expansion after the late Ice Age would not have occurred.

4. RELAXED NATURAL SELECTION

The original and present form (Kimura, 1968, 1991) of the neutral theory assumes that there are only two major classes of mutations; neutral and definitely deleterious. Deleterious mutations do not contribute at all to polymorphism and molecular evolution. In this view, changes in population size are irrelevant to the issue, and changes in selection pressure are the sole possibility. It should be noted that the theory does not necessarily assume that the degree of selective constraint is constant throughout evolutionary times. For instance, the relaxation in recent times is conspicuous particularly in countries with advanced technology. It means that deleterious mutations are accumulating faster than they are being eliminated, because these mutations no longer are deleterious under the *improved* environment.

A question is whether or not the relaxation responsible for the protein polymorphism can be supported by paleo-anthropological data. It is known that the human lineage has experienced dramatic changes in brain size, economic and social complexity, and culture over the past 2 million years (Tobias, 1991). Although neither learning nor the mother-infant relationship is specific to human, the period of *Homo habilis* was an incomparable stage of the development of human characters fostered by the enlarged brain size. It meant that infants were born at an earlier stage of mental and physiological maturity. For example, like the brain and digestive systems, the immune system that protects our body from pathogens is not yet fully developed in the newborn baby. Extensive parental care at home bases was needed to nurture such infants and the period of parental care became considerably longer. Thus, it is believed that associated with the enlarged brain size were anatomical and social changes.

During the Pleistocene, repeated glaciations of the Northern Hemisphere occurred and tropical forests were replaced by savanna woodlands. It was essential to solve dietary problems in the face of changing environmental conditions. Changed diet together with innovation of tool and fire must have altered the selection pressure as well, particularly for the enzymatic system, the point epitomized as *we are what we ate* (Milton, 1993).

Therefore, it seems reasonable to postulate that the enlarged brain had improved the environment of *H. habilis* and *H. erectus*. Previously deleterious mutations might have become neutral, some of which were incorporated into the population as polymorphism (Fig. 1). In all respects, we are constrained by what our ancestor was challenged to survive. The protein polymorphism, too, could not have escaped from this constraint.

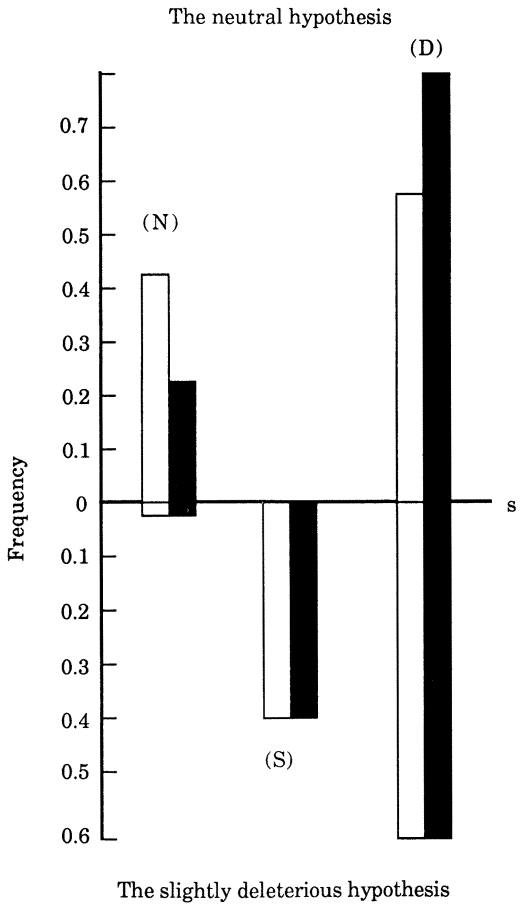

Fig. 1. Relative frequency of three classes of mutations; (N) neutral, (S) slightly deleterious, and (D) deleterious mutations. Those under the neutral theory (Kimura, 1968, 1991) and the slightly deleterious hypothesis (Ohta, 1975) are depicted in upper and lower part, respectively. The frequency in human is represented by open bars and that in chimpanzee by solid bars. In the neutral theory, some 20% of deleterious mutations (D) are converted as neutral and there is no intermediate class of mutations (S). The slightly deleterious hypothesis assumes that whereas the fraction D is the same in both species, the intermediate class (S) is effectively neutral in the human population with a smaller size and is driven by genetic drift. Since this hypothesis does not consider relaxed natural selection, the accumulation of slightly deleterious mutations gradually deteriorates the human genome.

5. CONCLUSION

One can still argue that the currently available data are insufficient to distinguish between the two alternative hypotheses discussed here (or that there are many other explanations). Theoretically, however, the slightly deleterious hypothesis must assume a particular frequency distribution of mutations against the fitness effect (Fig. 1). It is this restrictive assumption that makes the hypothesis less attractive even if the precise figure may change under different assumptions (Ohta and Tachida, 1990). The hypothesis also implies continual deterioration of the genome unless the optimum of alleles changes in a short time scale, less than a few million years. This genome deterioration has to be distinguished from that in endangered species such as the cheetah (O'Brien et al., 1987), which is likely to result from recessively deleterious mutations manifested by a sudden reduction in the population size or inbreeding depression.

In contrast, the relaxation hypothesis does not imply that the human genome has been deteriorating as long as we live under the improved environment. It is true, however, that if we for some reason had to go back and live the way our ancestor (e.g., *Australopithecus*) did, we would suffer the weakening effects of mutations that have accumulated in the intervening years as selection was relaxing (Crow, 1994).

In any case, comparative study of the genome between human and non-human primates is required to further characterize the human genome and trace its evolution in it. It is hoped that more systematic study will soon become feasible, including regulatory regions (King and Wilson, 1975). It may then provide some important clues to understand genetic bases of morphological changes during the development of humanity.

I thank Dr. J. F. Crow for his interests and comments on an early version of this communication. This work is supported by grants from the Ministry of Education, Science and Culture of Japan.

REFERENCES

Aquadro, C. F. and Greenberg, B. D. (1983). Human mitochondrial DNA variation and evolution: Analysis of nucleotide sequences from seven individuals. *Genetics* **103**, 287–312.

Aquadro, C. F. and Begun, D. J. (1992). Evidence for and implication of genetic hitchhiking in the *Drosophila* genome. In: *Mechanisms of Molecular Evolution* (eds.: N. Takahata and A. G. Clark), pp. 159–178. Japan Sci. Soc. Press, Tokyo/ Sinauer Assoc. INC. Sunderland.

Bruce, E. J. and Ayala, F. J. (1979). Phylogenetic relationships between man and the apes: Electrophoretic evidence. *Evolution* **33**, 1040–1056.

Crow, J. F. (1994). The role of mutations in evolution. In: *Principles of Medical Biology* (eds.: E. E. Bittar and N. Bittar). JAI Press, Inc., Greenwich, (in press).

Horai, S., Satta, Y., Hayasaka, K., Kondo, R., Inoue, T., Ishida, T., Hayashi, S. and Takahata, N. (1992). Man's place in Hominoidea revealed by mitochondrial DNA genealogy. *J. Mol. Evol.* **35**, 32–43.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature* **217**, 624–626.

Kimura, M. (1983). Rare variant alleles in the light of the neutral theory. *Mol. Evol. Biol.* **1**, 84–93.

Kimura, M. (1991). The neutral theory of molecular evolution: A review of recent evidence. *Jpn. J. Genet.* **66**, 367–386.

Kimura, M. and Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725–738.

King, M.-C. and Wilson, A. C. (1975). Evolution at two levels: Molecular similarities and biological differences between human and chimpanzees. *Science* **118**, 107–188.

Klein, J., Satta, Y., O'hUigin, C. and Takahata, N. (1993). The molecular descent of the major histocompatibility complex. *Annu. Rev. Immunol.* **11**, 269–295.

Kocher, T. D. and Wilson, A. C. (1991). Sequence evolution of mitochondrial DNA in humans and chimpanzees: Control region and a protein-coding region. In: *Evolution of Life* (eds.: S. Osawa and T. Honjo), pp. 391–413. Springer-Verlag, Tokyo.

Kreitman, M. (1983). Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* **304**, 412–417.

Li, W.-H. and Sadler, L. A. (1991). Low nucleotide diversity in man. *Genetics* **129**, 513–523.

Martin, A. P. and Palumbi, S. R. (1993). Body size, metabolic rate, generation time, and the

molecular clock. *Proc. Natl. Acad. Sci. USA* **90**, 4087–4091.

Martinko, J. M., Vincek, V., Klein, D. and Klein, J. (1993). Primate ABO glycosyltransferases: Evidence for trans-species evolution. *Immunogenetics* **37**, 274–278.

Milton, K. (1993). Diet and primate evolution. *Sci. Amer.* **269**, 70–77.

Nei, M. and Graur, D. (1984). Extent of protein polymorphism and the neutral mutation theory. *Evol. Biol.* **17**, 73–118.

Takahata, N. (1993). Allelic genealogy and human evolution. *Mol. Evol. Biol.* **10**, 2–22.

Tobias, P. V. (1991). Man, culture, and environment. In: *Evolution of Life* (eds.: S. Osawa and T. Honjo), pp. 363–378. Springer-Verlag, Tokyo.

O'Brien, S. J., Wildt, D., Bush, M., Caro, T. M., FitzGibbon, C., Aggundey, I. and Leakey, R. E. (1987). East African cheetahs: Evidence for two population bottlenecks? *Proc. Natl. Acad. Sci. USA* **84**, 508–511.

Ohta, T. (1975). Statistical analyses of *Drosophila* and human protein polymorphisms. *Proc. Natl. Acad. Sci. USA* **72**, 3194–3196.

Ohta, T. and Tachida, H. (1990). Theoretical study of near neutrality. I. Heterozygosity and rate of mutant substitution. *Genetics* **126**, 219–229.

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