# Human Mitochondrial DNA: A Clue to the Development and Dispersion of Asian Populations

Satoshi Horai

National Institute of Genetics
Mishima, Japan

**ABSTRACT**

Using restriction fragment length polymorphism (RFLP) of mitochondrial DNA (mtDNA), my colleagues and I first showed that the Japanese population could be separated into two distinct groups: A group with smaller frequency (group I) first diverged from the other group with larger frequency (group II). By a more exhaustive survey, the existence of the two groups was confirmed in two other Japanese populations, though the frequencies of the groups are different among populations. A phylogenetic analysis among three major racial groups indicated that the group I Japanese and the majority of Negroids first diverged from the rest of Japanese (group II) and Caucasoids. Later, we extended our analysis to various ethnic groups of humans to study their evolutionary relationships by the sequence analysis of the major noncoding region of mtDNA. Based on the sequence comparison of over a hundred of individuals, remarkable features of nucleotide substitutions and insertion/deletion events have been revealed. The uncleotide diversity among the sequences is estimated as 1.45%, which is three to six-fold higher than the corresponding value estimated from restriction enzyme analysis of the whole mtDNA genome. More recently we applied polymerase chain reaction (PCR) to molecular evolutionary studies on not only contemporary but archaeological samples. We have succeeded to amplify mtDNA extracted from an ancient Japanese bone, whose age is estimated at about 6000 years B. P., and determined the nucleotide sequence of part of the major noncoding region. Sequence comparison shows that the ancient individual has a close phylogenetic affiliation to Southeast Asians.

## I. INTRODUCTION

During the last decade, considerable progress has been made in both the molecular and the population genetics of human mitochondrial DNA (mt DNA). The human mtDNA is a circular genome of approximately 16.5 kilobase pairs (kb) in length, and the complete nucleotide sequence and gene organization of this genome have been determined (Anderson *et al.*, 1981). The genome can be divided into two domains—a coding region that constitutes over 90% of the genome, and one major noncoding region plus several small noncoding segments. In the coding region, the genes for 13 proteins, two ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs) have been identified (Anderson *et al.*, 1981; Attardi, 1985; Chomyn *et al.*, 1985). The major noncoding region contains the origin of replication of the H-strand (Anderson *et al.*, 1981), the origins of transcription of both strands (Cantatore and Attardi, 1980), and the displacement loop (D-loop). The origin of replication of the L-strand is located in a small noncoding segment, about 5.7 kb away from the origin of replication of the H-strand. Numerous reports have indicated that the nucleotide sequence of mtDNA is evolving much more rapidly than that of single-copy nuclear genes (Brown *et al.*, 1979; Ferris *et al.*, 1981).

Since there are substantial variations in nucleotide sequence among individuals (Brown and Goodman, 1979), restriction enzyme analysis of mtDNAs has become a powerful tool in attempts to elucidate evolutionary relationships among human ethnic groups (Brown, 1980; Denaro *et al.*, 1981; Blanc *et al.*, 1983; Johnson *et al.*, 1983; Horai *et al.*, 1984; Cann *et al.*, 1984; Wallace *et al.*, 1985; Horai and Matsunaga, 1986; Brega *et al.*, 1986; Cann *et al.*, 1987; Harihara *et al.*, 1988). Results from such studies suggest that there is a high correlation be-

tween mtDNA restriction types of mtDNAs and the ethnic origins of individuals.

Elucidation of the phylogenetic affiliation of the Japanese, as it developed during the course of human evolution is of particular interest to us. Although the Japanese have been isolated on the Japanese islands for a long time and are regarded as a seemingly homogeneous population, there are two Japanese populations that are ethnically different from the mainland Japanese in japan, namely, the Ainu, who inhabit the northern island (Hokkaido) and the Ryukyuans, who live on the most southern islands (Okinawa). The mainland Japanese are regarded as descendants mainly of postneolithic immigrants from the Asian Continent. By contrast, the Ainu and the Ryukyuans are considered to be descendants of native populations that existed during the neolithic Jomon period, even though they have undergone subsequent admixing with the mainland Japanese. Anthropological studies (Ikeda, 1982; Hanihara, 1984) and genetic data on classical genetic markers, such as blood groups, enzymes and serum proteins (Omoto *et al.*, 1973; Omoto and Misawa, 1976; Yuasa *et al.*, 1985) have indicated that the Ainu and the Ryukyuans have characteristics different from the mainland Japanese.

In a previous study, using restriction fragment length polymorphism (RFLP) of mtDNA, we first showed that the mainland Japanese population can be separated into two distinct groups: one group, with members found at lower frequency (group I) appears to have diverged from another group, with members found at higher frequency (group II) (Horai and Matsunaga 1986). As the result of a more exhaustive survey, the existence of both of these two groups was confirmed in the Okinawan population, even though the frequency of member of the individual groups is quite different from that in the mainland population (Horai *et al.*, 1987). This grouping can also be applied to other Mongoloid populations since a 9-bp deletion in region V (Cann and Wilson, 1983), which characterizes the group I Japanese, has also been observed in non-Japanese Mongoloid populations (Stoneking and Wilson, 1989; Stoneking *et al.*, 1992; Hertzberg *et al.*, 1989; Horai *et al.*, 1992; Horai, 1991a).

This report provide a summary of our recent studies on mtDNA polymorphisms in human populations. These studies include restriction enzyme analysis, nucleotide sequence analysis and analysis by the polymerase chain reaction (PCR). The results are discussed with reference to the evolution of mtDNA sequences at the gene level, as well as at the population level.

## II. METHODS FOR ANALYZING mtDNA

### 1. Restriction enzyme analysis

Highly purified mtDNAs, extracted from 259 placentae obtained from three different Japanese populations, were analyzed with the following nine enzymes: *Hae*III, *Hin*fI, *Sau*3AI, *Hha*I, *Rsa*I, *Taq*I, *Hpa*II, *Ava*II, and *Acc*II. Closed circular forms of mtDNAs were prepared as described previously (Horai *et al.*, 1984). The fragments of mtDNA produced by digestion with the enzymes were analyzed by electrophoresis on either agarose or polyacrylamide gels. After staining of fragments with ethidium bromide, restriction patterns were visualized under UV light.

2. Sequence analysis

A total of 115 individuals of three different racial origins were analyzed in terms of their mtDNA sequences in the D-loop region. Sixty-two of these individuals were from the central part of Japan (Horai *et al.*, 1984; Horai and Matsunaga, 1986); twenty-nine were non-Japanese Mongoloids, namely, three Koreans, four Chinese, eleven Malaysians, four Indonesians, four Thailanders, one Philippino, one Sri Lankan and one Papua New Guinean. We also collected samples of placenta from 17 Europeans (16 from Europe and America and one from India) and seven Africans from seven different regions of Africa, whose mothers gave birth to their babies in Tokyo, Japan.

Purified mtDNAs, doubly digested with *Kpn*I and *SacI*, were cloned into plasmid vectors. The double digestions with *Kpn*I and *Sac*I generated four or five fragments from each sample as a result of the polymorphism of one of the *Kpn*I sites. The smallest fragment with both *Sac*I and *Kpn*I recognition sites is either 482 or 563 bp in length and is derived from part of the D-loop region of mtDNA. Nucleotide sequences of 482- or 563-bp fragments, inserted into a vector, were determined by the dideoxy-chain termination method (Sanger *et al.*, 1977) using the Sequenase Kit (U. S. Biochemical) in accordance with the manufacturer's directions.

3. Polymerase chain reaction

The method for extraction of DNA from ancient bones was described elsewhere (Horai *et al.*, 1989).

A fragment of mtDNA was amplified by the method described by Saiki *et al.* (1988). For amplification of the major noncoding region of mtDNA, the set of primers had the following sequences: primer A, 16190-5'-CCCCATGCTTACAAGCAAG-3'-16208; and primer B, 16422-5'-ATTGATTTCACGGAGGATGG-3'-16403 [the notation of Anderson *et al.* (1981) is used for numbering of bases]. These oligonucleotide primers were synthesized on a Applied Biosystem model 380B DNA synthesizer. The polymerase chain reaction (PCR) was carried out for a total of 30 cycles with the use of a Thermal Cycler (Perkin Elmer Cetus). The cycle times were as follows: denaturation, 10 sec at 94℃; annealing, 10 sec at 45℃; primer extension, 15 sec at 72℃. The amplified fragments were separated by eletrophoresis on 1.5% agarose gels and detected fluorographically after staining with ethidium bromide. For the second PCR, in the case of ancient DNAs, one-tenth of the products of the first PCR was used as template and subjected to the PCR in exactly the same manner as the first PCR. The direct sequencing from the agarose extracts of the products of the second PCR (Gyllensten and Erlich, 1988) was performed with the Sequenase Kit.

4. Analysis of data

For restriction enzyme analysis, mtDNAs of all individuals were classified into restriction fragment types by combining the nine enzyme morphs. Genetic distances among types were calculated by the method of Nei and Li (1979). Phylogenetic trees were constructed by applying the unweighted pair-group (UPG) method (Sokal and Sneath, 1963; Nei, 1987) to the genetic distance data.

For analysis of sequences, we estimated the number of nucleotide substitutions per site between individual sequences using the six-parameter model of nucleotide substitution (Go-

jobori *et al.*, 1982). On the basis of the estimated number of nucleotide substitutions, phylogenetic trees were constructed by the UPG method.

## III. RESULTS AND DISCUSSION

### 1. Restriction enzyme analysis of mtDNA in the Japanese

The restriction enzyme analysis of mtDNA from 259 Japanese from three different regions, which are located at some distance from one another, has revealed remarkable features of mtDNA polymorphism within and between populations (Horai and Matsunaga, 1986; Horai *et al.*, 1987; Horai, 1991a). First, the average value of the nucleotide diversity for mtDNA is 0.26% both in Shizuoka and Aomori, and 0.16% in Okinawa. Second, we found seven common types, which accounted for 33-39% of individuals in each local population, while we observed 27 specific types in Aomori (46%), 45 in Shizuoka (42%) and 25 in Okinawa (35%) (see Table I in Horai, 1991a). These regional specificities may be attributed to the maternal inheritance of mtDNA. It seems reasonable to assume that women have retained their regional founder types of mtDNA for years since, in the past, once settled in one district women had fewer chances to move to some distant location than did men. If this assumption is valid, it is natural that we should observe a large number of specific types of mtDNA in each region. These regional specificities in mtDNAs at the present time contrast sharply with polymorphism of nuclear DNA which shows no regional differences (Horai *et al.*, 1992; Horai, 1991a). Third, we found regional specificities in the phylogenetic trees constructed for each population. The phylogenetic trees reveal that respective regions have two major clusters - group I and group II - and that the frequencies of members of groups I and II vary among populations. The frequency of members of group I is 5% in Okinawa, 18% in Shizuoka and 28% in Aomori (see Table II in Horai, 1991a). There seems to be a tendency for the frequency of members of group I to increase from the south to the north of Japan. This difference in the frequencies of members of groups I and II may also reflect the founder effect of maternal ancestors.

It is of some considerable interest that the Japanese have two major maternal lineages. We have attempted to study processes by which the groups have been formed and to identify inherited characteristics in each lineage. Individuals belonging to group I possess several variations in common, *Hae*III-morph 35, *Hin*fI-morphs 27 and 28, *Rsa*I-morph 13 or 17, *Taq*I-morph 15, while those in group II never show these variations (Horai and Matsunaga, 1986). Moreover, the majority of the members of group I have the *Hae*III-morph 2, a length polymorphism that we have omitted from consideration in the construction of the phylogenetic trees. In this morph, a fragment of 132-bp is observed instead of the fragment of 141-bp that is usually observed in most individuals when the mtDNA is digested by *Hae*III. This variation is due to a deletion of one copy of a 9-bp tandem repeat sequence (CCCCCTCTA) (Wrischnik *et al.*, 1987). The deletion of 9-bp is a variation found only in Mongoloids (Asians and Oceanians) and not in Europeans or Africans.

### 2. Relationships among Japanese, Europeans and Africans

As mentioned above, phylogenetic analysis of the Japanese populations indicates that there

exist at least two distinct groups of mtDNAs in Japan. To evaluate the evolutionary significance of this finding, we constructed a phylogenetic tree of mtDNAs from the three major races, combining our data for Japanese with those for Europeans and Africans analyzed by Cann (1982). A total of 117 restriction types were observed among the three racial groups. It is to be noted that the three groups do not share any restriction types in common. The average number of nucleotide substitutions for the Japanese is almost the same as that for the European, whereas the value for the Africans is about twice that for the other two groups. This difference implies that the mtDNAs of the Africans are much more diverse than those of the Europeans or Japanese (Horai *et al.*, 1986; 1987).

We also estimated the number of nucleotide substitutions for all pairs of individuals in the three racial groups and constructed a phylogenetic tree for the 117 types by the UPG method (see Fig. 2 in Horai *et al.*, 1987). From the clustering patterns on the tree, all lineages were classified into eight clusters. While three of them represent clusters that contain interminglings of individuals from different racial groups, the remaining five show distinct clusters within the same races. On the tree, the African lineages and some of the Japanese ones (equivalent to the Japanese group I) appears to have diverged first from the rest of the clusters. The first branching point in the tree goes back to about 170,000 - 340,000 years ago, if we assume the rate of nucleotide substitution for mtDNA to be 1 - $2x10^{-8}$ per site per year (Wilson *et al.*, 1985).

3. Properties of intraspecific differences in mtDNA sequences

In a previous study (Horai and Hayasaka, 1990), we determined nucleotide sequences of the major noncoding region of human mtDNA from 95 individuals from the three major races. These sequences include a region of at least 482-bp that encompasses most region that forms the D-loop. Comparisons of these sequences with those previously determined by Anderson *et al.* (1981) and Greenberg *et al.* (1983) have revealed remarkable features of nucleotide substitutions and insertion/deletion events.

A total of 100 sequences were compared with the reference sequence reported by Anderson *et al.* (1981). Base substitutions, deletions and insertions detected in at least one individual were also scored. We found mutations at 113 sites in total. At four sites, two different kinds of nucleotide substitution were observed. Nucleotide changes were observed more frequently in the 5' half of the region than in the 3' half. The observed numbers of mutations, which were classified according to the type of mutation, indicate that transition types of substitution are more frequent than transversions, namely, 97% were transitions while only 3% were transversions. Moreover, transitions between pyrimidines were more prevalent than those between purines. When we counted the nucleotide substitutions for each site, a 10-fold bias in favor of transitions over transversions was still observed. The results confirmed several features of the noncoding region of mtDNA that were previously reported by Greenberg *et al.* (1983) and Aquadro and Greenberg (1983), although only eight or seven sequences were compared by these authors. We also observed the deletion of adenine at two sites and insertion of cytosine at five sites. These deletions and insertions were mainly found in a particular domain that contained serially repeated stretches of adenines and cytosines. In this 14-bp region (bp 16180-16193), 17 different sequences were detected (see Table 2 in Horai and Hayasaka, 1990). About 60% of individuals exhibited the sequence that contains a 14-base

stretch of four adenines, five cytosines, one thymine and four cytosines. However, the remainder showed a variety of sequences in this domain. In eight variant sequences, there were base substitutions in the stretches of cytosines. In another eight variant sequences, differences were the result of elongation of the stretch of cytosines and shortening of the stretch of adenines. Once the thymine at bp 16189 is replaced by cytosine, the number of adenines and cytosines becomes flexible, probably as a result of errors in replication. This thymine-to-cytosine transition has occurred independently several times in different lineages, as confirmed by the phylogenetic analysis that is discussed below.

A significantly nonrandom distribution of nucleotide substitutions and variations in sequence length was also noted. Analysis of the distribution of mutations in the D-loop region showed that sites of most of the highly variable bases lie in near the gene for proline tRNA, while one highly polymorphic site was observed near the origin of replication of the H-strand. This site is polymorphic for all three major races and the polymorphism is probably derived from an ancient polymorphism at this site. Furthermore, the nonrandom distribution of mutations can be explained by the neutral theory (Kimura, 1983) which points that some selective constraint is operative even in the major noncoding region of mtDNA.

4. Phylogeny of mtDNA sequences

The nucleotide sequence analysis of the major noncoding region has been extended by the addition of 20 individuals, mainly from Southeast Asia (Horai, 1991b). The 482-bp sequences from a total of 121 individuals were aligned and compared, and the number of nucleotide substitutions between each pair of sequences was also estimated. The nucleotide diversity among the 121 individuals is estimated as 1.44%, which is three-to six-fold higher than estimates based on restriction analysis of mtDNA reported in human populations (Brown, 1980; Horai *et al.*, 1986; Cann *et al.*, 1987). Based on the estimated number of nucleotide substitutions between individual sequences, a phylogenetic tree was constructed by the UPG method, as shown in Fig. 1. On the basis of clustering patterns on the tree, all lineages were classified into five clusters, which are shown by brackets with cluster numbers (C1-C5).

The phylogenetic analysis indicates that diversity among the Africans is much larger than that among the Europeans or the Asians. In fact, the majority of Africans diverged first from the rest of the individuals in the phylogenetic tree. A striking finding in the phylogenetic analysis is that the Asians can be separated into two distinct groups. Divergence of part of the Asian population (C2 in Fig. 1) followed the earliest divergence of part of the population of Africans. The remainder of the Asians (C3 through C5 in Fig. 1) diverged subsequently, together with the Europeans. This observation confirms the result of our earlier study which clearly demonstrated the exsitence of two distinct groups of Japanese by restriction enzyme analysis (Horai and Matsunaga, 1986).

5. Phylogenetic analysis within and between racial groups

To analyze the nucleotide diversities within and between the racial groups, the nucleotide sequences were compared quantitatively on the basis of estimates of the number of nucleotide substitutions between individual sequences. Table I shows that the nucleotide diversity within Africans is much larger than that within Europeans or Asians. Furthermore, the nucleotide diversity within Africans is larger than all of the interracial diversity. Thus, earlier findings

that Africans have highly diversified mtDNAs, as compared with Europeans and Asians, which were deduced from restriction enzyme analysis (Horai *et al.*, 1986; Cann *et al.*, 1987) were confirmed by the quantitative analysis of the nucleotide sequences.

In Table 1 the nucleotide differences between Asians-1 and Asians-2 ($d_{XY} = 0.0169$) are seen to be much larger than those between Europeans and Asians-2 ($d_{XY} = 0.0124$). This Table also shows that the nucleotide diversity among Asians-1 ($d_X = 0.0167$) is much larger than that among Asians-2 ($d_X = 0.0114$). These two subpopulations of Asians correspond roughly to groups I and II of the Japanese, respectively, which were inferred from the restriction enzyme analysis (Horai and Matsunaga, 1986). The restriction enzyme analysis revealed that nucleotide diversity within group I is larger than that within group II, a result that agrees well with the results of the present study. Thus, it is evident that the Asian population can be separated into two subpopulations.

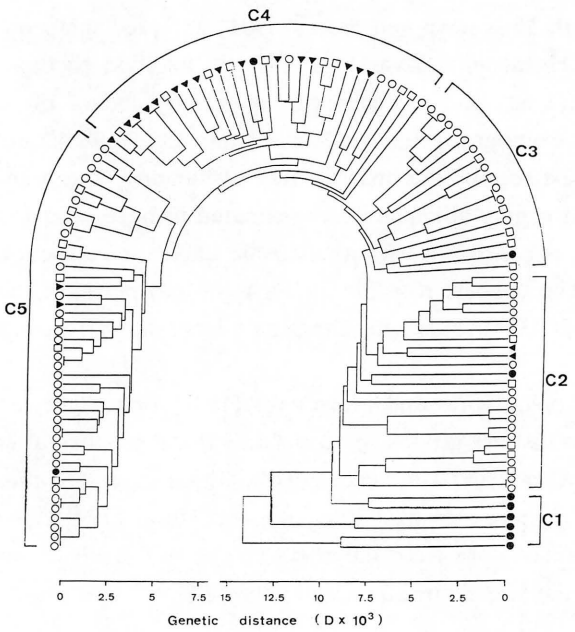


Fig. 1. Phylogenetic tree showing the 121 mtDNA lineages from the three racial groups, based on the sequence data from the major noncoding region. All lineages are classified into 5 clusters which are designated Cl through C5 according to the clustering patterns. Distances (D) are expressed by the number of nucleotide substitutions per site per year. ○, Japanese; □, non-Japanese Asian; ▲, European; ●, African.

Table 1. Estimates of the average number of nucleotide substitutions per site between races ($d_{XY}$) and within races ($d_X$ or $d_Y$)

| | European N=20 | Asian-2 N=69 | Asian-1 N=22 | African N=10 |
|---|---|---|---|---|
| European | 0.0096 | | | |
| Asian-2 | 0.0124 | 0.0114 | | |
| Asian-1 | 0.0146 | 0.0169 | 0.0167 | |
| African | 0.0194 | 0.0195 | 0.0226 | 0.0235 |

Figures on the diagonal represents $d_X$ or $d_Y$. Those below the diagonal represent $d_{XY}$. Asian-2 includes individuals that belong to C3 - C5 in Fig. 1, while the other Asians that belong to C2 are Asian-1.

6. Dating the origin of polymorphisms in human mtDNA

From restriction enzyme analysis of mtDNA from the major human races, several authors have suggested that the deepest root of the mtDNA tree of humans can be dated back to about 200,000 years ago (Brown, 1980; Johnson *et al.*, 1983; Horai *et al.*, 1986; Cann *et al.*, 1987). Recently, Vigilant *et al.* (1989) attacked this problem by directly sequencing 740 nucleotides in the major noncoding region of mtDNAs of 83 individuals from the major geographic populations. They estimated that deepest root of the mtDNA tree of humans dated back to about 238,000 years ago, consistent with previous estimates from restriction enzyme analysis. The human/chimpanzee divergence was taken as a reference to calibrate the clock, which suggested that a large number of multiple substitutions had occurred between these two species. The way in which Vigilant *et al.* (1989) corrected for multiple substitutions was, however, only approximate and it seems that it might be worthwhile to reanalyze their data by a more rigorous statistical method.

More recently, using a generalized least-squares method developed by Hasegawa *et al.* (1985) and by Kishino and Hasegawa (1990), Hasegawa and Horai (1991) analyzed the sequence data from Vigilant *et al.* (1989) and Horai and Hayasaka (1990), in addition to those from previous reports by Anderson *et al.* (1981) and Greenberg *et al.* (1983) and to those from the common chimpanzee and pygmy chimpanzee reported by Foran *et al.* (1988), and they have estimated the timing of the deepest root of the mtDNA tree of humans. The branching between the common chimpanzee and pygmy chimpanzee is estimated to have occurred about 2 million years ago, and this estimate is consistent with an estimate based on restriction analysis of mtDNA (Wilson *et al.*, 1985). The deepest root for mtDNA polymorphism in the human population is estimated to be some 280,000 years ago, although a large standard error is still attached to this estimate.

From restriction enzyme analysis of mtDNA, Horai and Matsunaga (1986) first suggested that the Japanese population consists of two distinct groups (groups I and II), a proposal that was confirmed by the sequence analysis of Horai and Hayasaka (1990), which suggested that the grouping can be also applied to non-Japanese Asians. Hasegawa and Horai (1991) also estimated the data of branching using sequence data from members of the two groups, and they suggested that the separation between the two groups occurred about 200,000 years ago.

7. Amplification of mtDNA from ancient Japanese bones.

The development of PCR makes it possible to amplify DNA in targeted regions from a very small amount of template DNA. By applying this technique to archaeological or ancient samples, we have successfully amplified DNA and analyzed a part of the nucleotide sequence. However, such analysis was possible only when we used soft tissues, such as frozen, mummified or stuffed tissue, haired skin and muscle tissue (Johnson *et al.*, 1985; Higuchi *et al.*, 1984; Pääbo, 1986; Pääbo,1989; Pääbo *et al.*, 1988). Such suitable samples have generally been preserved artificially or accidentally. Most of the human remains preserved to data are hard tissues, such as bones. If we could amplify DNA extracted from bones and analyze it, we would probably get some quite important information relevant to the process of human evolution, the divergence of human races, the restoration of details of human populations in the past, and their migration or dispersal.

We attempted to amplify mtDNA from 22 samples of human bones derived from different

eras. In the first PCR we were able only to amplify DNA in a sample of contemporary DNA as a positive control. However, when we took a portion of the products of the first PCR and performed a second 30-cycle PCR, DNA in some samples of bone extracts could be amplified. Although the amplification of DNA by the second PCR from the majority of samples was successful, it was impossible to amplify DNA in some samples. Success in this regard is probably dependent on the condition of preservation of bones. However, we succeeded in amplifying mtDNA from the oldest sample, a skull designated Urawa-1, excavated from five meters below the ground in Urawa City, in the central part of Japan, in 1988.

Collagen, isolated simultaneously from this skull, was used to estimate the absolute age of Urawa-1 as at 5790±120 years B. P. by the direct detection of $^{14}C$ by accelerator mass spectrometry (Nakai *et al.*, 1984; Nakamura *et al.*, 1985). This age corresponds to the first part of the Jomon period in Japanese prehistory. We determined the nucleotide sequence of the product the second PCR of Urawa-1 by the direct-sequencing method (Gyllensten and Erlich, 1988). We found nucleotide substitutions at three positions when we compared a 190-bp region sequence from Urawa-1 with that originally reported by Anderson *et al.* (1981).

As mentioned above, we determined and compared nucleotide sequences of the major noncoding region of mtDNA from 115 humans from three racial groups (Horai, 1991b). These sequences include a region of 482 bp that encompasses most of the D-loop region of mrDNA. The PCR-amplified region of mtDNA in the present study was completely included within this region.

The nucleotide sequence from Urawa-1 was compared with those from 121 modern humans. Table II summarizes the number of nucleotid differences between the sequence from Urawa-1 and those from each of the 121 individuals in the 190-bp region. In comparing the sequence from Urawa-1, we found complete identity with sequences from two Southeast Asians (a Malaysian and an Indonesian) out of 29 non-Japanese Asians. However, we never observed a sequence identical to that from Urawa-1 among the 62 contemporary Japanese. Fifteen Japanese differ at one position and eight Japanese differ at two positions. Furthermore, 39 Japanese differ at three to eight positions even within this short region. These observations indicate that the ancestor of the Japanese who, presumably, lived in the central part of Japan about 6000 years ago shared a common origin with some contemporary Southeast Asians. It is striking that the 62 contemporary Japanese do not have a sequence that is completely identical to that of Urawa-1, even though about one-third of the individuals examined exhibit only one or two nucleotide differences in the 190-bp region.

Table II. Sequence differences between Urawa-1 and each of 121 contemporary humans in the 190-bp region (bp 16190-16379)

| | Number of nucleotide differences | | | |
|---|---|---|---|---|
| Race | 0 | 1 | 2 | 3 - 8 |
| Japanese | 0 | 15 | 8 | 39 |
| Asian* | 2 | 1 | 1 | 25 |
| European | 0 | 0 | 1 | 19 |
| African | 0 | 0 | 0 | 10 |

*denotes non-Japanese Asians.

Restriction enzyme analysis revealed that the Japanese population can be separated into at least two distinct groups. This classificaion is confirmed by the sequence analysis of the major noncoding regions from Mongoloids, including non-Japanese Asians. In this respect, the ancient Japanese (Urawa-1) belongs to one of the two groups (group II) of modern Japanese. Some of the people who migrated from the continent during a period of two to three thousand years ago may be representatives of the other group (group I) of modern Japanese, because the sequences from individuals in group I show 3 to 8 nucleotide differences in the 190-bp region when compared with sequence from Urawa-1. Although our findings were made by determining the nucleotide sequence from only one ancient individual, this archaeological and molecular genetic study adds a new perspective to the evolutionary history of human populations.

REFERENCES

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981). Nature. *290*, 457.

Aquadro, C. F., and Greenberg, B. D. (1983). Genetics. *103*, 287.

Attardi, G. (1985). International Review of Cytology. *93*, 93.

Blanc, H., Chen, K. H., D'Amore, M. A., and Wallace, D. C. (1983). Am. J. Hum. Genet. *35*, 167.

Brega, A., Gardella, R., Semino, O., Morpurgo, G., Astaldi Ricotti, G. B., Wallace, D. C., and Santachiara-Benerecetti, A. S. (1986). Am. J. Hum. Genet. *39*, 502.

Brown, W. M. (1980). Proc. Natl. Acad. Sci. USA. *77*, 3605.

Brown, W. M., and Goodman, H. M. (1979). *In* "Extrachromosomal DNA" (D. Cummings, P. Borst, I. Dawid, S. Weissman, and C. F. Fox, eds.) , p. 485. Academic Press, New York.

Brown, W. M., George, JR, M., and Wilson, A. C. (1979). Proc. Natl. Acad. Sci. USA. *76*, 1967.

Cann, R. L. (1982). "The Evolution of Human Mitochondrial DNA" Ph. D. thesis, University of California, Berkeley, California.

Cann, R. L., and Wilson, A. C. (1983). Generics. *104*, 699.

Cann, R. L., Brown, W. M., and Wilson, A. C. (1984). Genetics. *106*, 479.

Cann, R. L., Stoneking, M., and Wilson, A. C. (1987). Nature. *325*, 31.

Cantatore, P., and Attardi, G. (1980). Nucleic Acids Res. *8*, 2605.

Chomyn, A., Mariottini, P., Cleeter, M. W. J., Ragan, C. I., Matsuno-Yagi, A., Hatefi, Y., Doolittle, R. F., and Attardi, G. (1985). Nature. *314*, 592.

Denaro, M., Blanc, H., Johnson, M. J., Chen, K. H., Wilmsen, E., Cavalli - Sforza, L. L., and Wallace, D. C. (1981). Proc. Natl. Acad. Sci. USA. *78*, 5768.

Ferris, S., Brown, W. M., Davidson, W. S., and Wilson, A. C. (1981). Proc. Natl. Acad. Sic. USA. *78*, 6319.

Foran, D. R., Hixon, J. E., and Brown, W. M. (1988). Nucleic Acids Res. *16*, 5841.

Gojobori, T., Ishii, K., and Nei, M. (1982). J. Mol. Evol. *18*, 414.

Greenberg, B. D., Newbold, J. E., and Sugino, A. (1983). Gene. *21*, 33.

Gyllensten, U. B., and Erlich, H. A. (1988). Proc. Natl. Acad. Sci. USA. *85*, 7652.

Hanihara, K. (1984). Acta Anthropogenet. *8*, 149.

Harihara, S., Saitou, N., Hirai, M., Gojobori, T., Park, K. S., Misawa, S., Ellepola, S. B., Ishida, Y., and Omoto, K. (1988). Am. J. Hum. Genet. *43*, 134.

Hasegawa, M., and Horai, S. (1991). J. Mol. Evol. 32, 37.

Hasegawa, M., Kishino, H., and Yano, T. (1985). J. Mol. Evol. *22*, 160.

Hertzberg, M., Mickleson, K. N. P., Serjeantson, S. W., Prior, J. F., and Trent, R. J. (1989). Am. J. Hum. Genet. *44*, 504.

Higuchi, R. G., Bowman, B., Freiberger, M., Ryder, O. A., and Wilson, A. C. (1984). Nature. *312*, 282.

Horai, S. (1991a). *In* "New Aspects of the Genetics in Molecular Evolution" (M. Kimura, and N. Takahata, eds.), p.135. Japan Sci. Soc. Press, Tokyo/Springer-Varlag, Berlin.

Horai, S. (1991b). *In* "New Era of Bioenergetics" (Y. Mukohata, eds.), p.273. Academic Press. Tokyo.

Horai, S., and Hayasaka, K. (1990). Am. J. Hum. Genet. *46*, 828.

Horai, S., and Matsunaga, E. (1986). Hum. Genet. *72*, 105.

Horai, S., Gojobori, T., and Matsunaga, E. (1984). Hum. Genet. *68*, 324.

Horai, S., Gojobori, T., and Matsunaga, E. (1986). Jpn. J. Genet. *61*, 271.

Horai, S., Gojobori, T., and Matsunaga, E. (1987). *In* "Human Genetics, Preceedings of the 7th International Congress" (F. Vogel, and K. Sperling, eds.) p. 177. Springer-Verlag, Heidelberg.

Horai, S., Hayasaka, K., Hirayama, K., Takenaka, S., and Pan, I-H. (1992). *In* "Proceedings of Circum-Pacific Prehistory Conference, Seattle 1989" (in press)

Horai, S., Hayasaka, K., Murayama, K., Wate, N., Koike, H., and Nakai, N. (1989). Proc. Jpn. Acad. *65, Ser. B*, 229.

Ikeda, J. (1982). J. Anthrop. Soc. Nippon. *90* (Suppl.), 91.

Johnson, P. H., Olson, C. B., and Goodman, M. (1985). Comp. Biochem. Physiol. *81. B*, 1045.

Johnson, M. J., Wallace, D. C., Ferris, S. D., Rattazi, M. C., and Cavalli - Sforza, L. L. (1983). J. Mol. Evol. *19*, 255.

Kimura, M. (1983). "The Neutral Theory of Molecular Evolution." Cambridge University Press, Cambridge.

Kishino, H., and Hasegawa, M. (1990). Methods Enzymol. *183*, 550.

Nakai, N., Nakamura, T., Kimura, M., Sakase, T., Sato, S., and Sakai, A. (1984). Instr. & Meth. *233*, 171.

Nakamura, T., Nakai, N., Sakase, T., Kimura, M., Ohishi, S., Taniguchi, M., and Yoshioka, S. (1985). Japan J. Appl. Phys. *24*, 1716.

Nei, M. (1987). "Molecular Evolutionary Genetics. " Columbia University Press, New York.

Nei, M., and Li, W-H. (1979). Proc. Natl. Acad. Sci. USA. *76*, 5269.

Omoto, K., and Misawa, S. (1976). *In* "Biological Origins of the Australians" (R. K. Kirk, and A. G. Thorne, eds.) , p. 365. Australian Institute Aboriginal Studies, Canberra.

Omoto, K., Ishizaki, K., Harada, S., Akaishi, S., Kudo, T., and Takahashi, K. (1973). J. Anthrop. Soc. Nippon. *81*, 159.

Pääbo, S. (1986). Nature. *314*, 644.
Pääbo, S. (1989). Proc. Natl. Acad. Sic. USA. *86*, 1939.
Pääbo, S., Gifford, J. A., and Willson, A. C. (1988). Nucl. Acids Res. *16*, 9775.
Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R. G., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Science. *239*, 487.
Sanger, F., Nicklen, S., and Coulson, A. R. (1977). Proc. Natl. Acad. Sci. USA. *74*, 5463.
Sokal, R. R., and Sneath, P. H. A. (1963). *In* "Principles of Numerical Taxonomy"P. X. W. H. Freeman, San Francisco.
Stoneking, M., and Wilson, A. C. (1989). *In* "The Colonization of the Pacific: a Genetic Trail" (A. Hill, and S. Serjeantson, eds.), p. 215. Oxford University Press, Oxford.
Stoneking, M., Sofro, A. S. M., and Wilson, A. C. (1992). *In* "Proceedings of the Circum-Pacific Prehistory Conference, Seattle 1989. " (in press)
Vigilant, L., Pennington, R., Harpending, H., Kocher, T. D., and Wilson, A. C. (1989). Proc. Natl. Acad. Sci. USA. *86*, 9350.
Wallace, D. C., Garrison, K., and Knowler, W. C. (1985). Am. J. Phys. Anthropol. *68*, 149.
Willson, A. C., Cann, R. L., Carr, S. M., George, JR. M., Gyllensten, U. B., Helm-Bychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prager, E. M., Sage R. D., and Stoneking, M. (1985). Biol. J. Linn. Soc. Lond. *26*, 375.
Wrischnik, L. A., Higuchi, R. G., Stoneking, M., Erlich, H. A., Arnheim, N., and Wilson, A. C. (1987). Nucl. Acids Res. *15*, 529.
Yuasa, I., Taira, T., Suenaga, K., Ito, K., and Okada, K. (1985). Hum. Genet. 70, 32.

## ミトコンドリアDNAを指標とした日本人を含むアジア人集団の進化

宝来　聰

日本国内３地域（青森・静岡・沖縄）より収集した計259検体の胎盤より、ミトコンドリアDNAを精製し、９種類の制限酵素による多型分析を行った。各制限酵素の切断型の違いによって、タイプ分けを行うと、各地域の約３割の個体は、３地域に共通したタイプを示したが、４割以上は、それぞれの地域で特有に観察されるタイプであった。さらに各地域集団で観察されたタイプの間で系統樹を作成したところ、いずれの地域でも、大きく２つのクラスター（グループⅠとⅡ）に分かれ、それらの割合は地域によって異なることが明らかになった。さらに、３大人種を含めた系統分析によって日本人に観察されるグループⅠの系統は、アフリカ人のクラスターに引続いて分岐する古い起源をもつことが明らかになった。

日本人に見られる２大クラスターの人類進化学的意義を明らかにするため、日本人以外のアジア人、欧米人、アフリカ人から胎盤を収集し、日本人も含めた計121人の３大人種に関して、ミトコンドリアDNAの遺伝子をコードしていない主要領域の塩基配列を直接決定し、その比較系統分析を行った。この領域における塩基の多様性の度合は、制限酵素切断型でミトコンドリアゲノム全体を対象とした場合の３倍以上であり、個体間の変異を固定するのに極めて有効であることが示された。さらに系統樹作成による分析で、日本人に観察される２大グループ分けは、アジア人集団にも当てはまることが明らかとなった。すなわち日本人を含むアジア人集団は、アフリカ人の主要な系統に引き続いて分岐するグループと、その後、欧米人のグループと共に分岐するグループに明瞭に分かれることが明らかとなった。さらに系統樹の詳細な解析

から、人類集団におけるミトコンドリアDNAの多型の起源は約28万年前に遡ることが推定され、アジア人に観察される2大グループも約20万年前に分岐したとの値が得られた。

Polymerase Chain Reaction (PCR) の開発によって、微量のDNAから目的とする領域のDNAの増幅が可能となった。この方法は、考古学試料あるいは陳旧試料からのDNA分析にも応用できることが報告されたが、いずれも、軟組織を用いた場合に、DNAの増幅およびその一部の塩基配列の解析に成功している。しかし、これらの試料は人為的あるいは偶然の結果、保存されていたものが多い。ヒトの遺物として現在まで保存されているものの大部分は、硬組織である骨である。従って骨からDNAの増幅と分析が可能になれば、人類の進化、人種の成立、過去の人類集団の復元および人類諸集団の移住と拡散の過程の解明に、きわめて貴重な情報を与えるものと思われる。我々は、現在に近い過去の試料（昭和初期）から順次、時代を遡り、最終的には、縄文時代の人骨（浦和1号）よりDNAの増幅と塩基配列の決定に成功した。増幅した領域は、ミトコンドリアDNAの中の遺伝子をコードしていない主要領域である。コラーゲンを用いて浦和1号の絶対年代を測定したところ、B.P.5790±120年の値が得られ、縄文前期と考えられる。さてミトコンドリアDNAは、Andersonら（1981）によってその全塩基配列が決定されているが、浦和1号より決定した190塩基対の配列と比較したところ3カ所で塩基置換が観察された。さらに上記の分析で決定した現代人121人の配列と比較したところ、これらの変異をもつ個体は、現代日本人62人の中には見つからなかった。しかし日本人以外のアジア人29人のうち、東南アジア出身の2人（マレーシア人・インドネシア人）と全く同じ箇所で変異を共有していた。このことは、約6000年前に日本列島の中心部（現在の埼玉県）に住んでいたこの縄文人の祖先は、現代東南アジア人の一部と共通の起源をもつことを示している。