Mitochondrial 16S rRNA Sequence Diversity of Hominoids

R. Noda, C. G. Kim, O. Takenaka, R. E. Ferrell, T. Tanoue, I. Hayasaka, S. Ueda, T. Ishida, and N. Saitou

We determined nucleotide sequences of the 16S rRNA gene of mitochondrial DNA (mtDNA) (about 1.6 kb) for 35 chimpanzee, 13 bonobo, 10 gorilla, 16 orangutan, and 23 gibbon individuals. We compared those data with published sequences and estimated nucleotide diversity for each species. All the ape species showed higher diversity than human. We also constructed phylogenetic trees and networks. The two orangutan subspecies were clearly separated from each other, and Sumatran orangutans showed much higher nucleotide diversity than Bornean orangutans. Some gibbon species did not form monophyletic clusters, and variation within species was not much different from that among species in the subgenus *Hylobates*.

From the Laboratory of Evolutionary Genetics, National Institute of Genetics, Mishima, 411-8540 Japan (Noda, Kim, and Saitou), Primate Research Institute, Kyoto University, Inuyama, Japan (Takenaka), Department of Human Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania (Ferrell), SanwaKagaku Kenkyusho Kumamoto Primate Park (Tanoue and Hayasaka), and Department of Biological Science, University of Tokyo (Ueda and Ishida). This study was partially supported by grants-in-aid for scientific studies from the Ministry of Education, Culture, Sport, Science, and Technology, Japan and by the Cooperation Research Program of Primate Research Institute, Kyoto University (to N.S.). We thank Aoshiko Akiko, Suzuki Mayumi, and Kondo Masayo for their experimental assistance. We also thank Suzuki Mayumi for assistance in submitting sequence data to DDBJ via SAKURA. We thank the Ueno Zoo of Tokyo for providing us two gorilla samples. This study was planned by N. Saitou. DNA samples of 3 orangutan and all gibbons were provided by O. Takenaka; those of 14 chimpanzee, 7 gorilla, 8 orangutan, and all bonobo were provided by R. E. Ferrell; 2 chimpanzee, 1 gorilla, and 1 orangutan were provided by S. Ueda; 4 orangutan were provided by T. Ishida; and those of 19 chimpanzee individuals were provided by T. Tanoue and I. Hayasaka. Sequences of gibbons, orangutan, gorilla, and bonobo were determined by R. Noda, while C. G. Kim determined those of all chimpanzees. Sequence analyses were conducted by R. Noda and N. Saitou, and the manuscript was written by N. Saitou and R. Noda. Address correspondence to N. Saitou at the address above or e-mail: nsaitou@genes.nih.ac.jp. This paper was delivered at a symposium entitled “Primate Evolutionary Genetics” sponsored by the American Genetic Association at Town and Country Resort and Convention Center, San Diego, CA, USA, May 19–20, 2001.

© 2001 The American Genetic Association 92:490–496

Materials and Methods

Genomic DNA samples for 35 chimpanzee (*Pan troglodytes*), 14 bonobo (*Pan paniscus*), 10 gorilla (*Gorilla gorilla*), 16 orangutan (*Pongo pygmaeus*), and 23 gibbon (*Hylobates*) individuals were used (see Table 1). Polymerase chain reaction (PCR) was carried out using specific primers based on primate mitochondrial sequences: SN-MtSd, 5'-CACACCCGGCTACGCG-CYTCAA-3', and SN-MtSr, 5'-ACCGGGCTC TGCCATCTTACSN-3′ for PCR. This region includes part of the 12S rRNA, valine transfer RNA, and most of the 16S rRNA gene. Each PCR mixture contained 160 mM each of dATP, dTTP, dCTP, and dGTP, 150 mM MgCl₂, reaction buffer containing no Mg²⁺, 1 mM of each primer, and 2 units Taq polymerase (TOYOBO and NIPPON GENE). About 200 ng of genomic DNA was used as template. The PCR program consisted of 20–45 cycles of 1-min denaturation at 94°C followed by 1-min primer annealing at 65°C or 68°C, and a 2-min extension at 72°C (PE GeneAmp PCR system 2400 and 9700). Immediately preced-
Table 1. Ape DNA samples used in this study

<table>
<thead>
<tr>
<th>Species</th>
<th>individual ID</th>
<th>Accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chimpanzee</td>
<td>chimp_#1–#35</td>
<td>AB062513–AB062547</td>
</tr>
<tr>
<td>Bonobo</td>
<td>bonobo_#1–#5</td>
<td>AB050147–AB050151</td>
</tr>
<tr>
<td></td>
<td>bonobo_#6–#13</td>
<td>AB065137–AB065144</td>
</tr>
<tr>
<td>Gorilla</td>
<td>gorilla_#1–#4</td>
<td>AB050152–AB050155</td>
</tr>
<tr>
<td></td>
<td>gorilla_#5–#10</td>
<td>AB064490–AB064495</td>
</tr>
<tr>
<td>Orangutan</td>
<td>orang_#1–#16</td>
<td>AB050156–AB050172</td>
</tr>
<tr>
<td>H. agilis</td>
<td></td>
<td>AB050172–AB050174</td>
</tr>
<tr>
<td>H. lar</td>
<td></td>
<td>AB050175–AB050177</td>
</tr>
<tr>
<td>H. muelleri</td>
<td></td>
<td>AB050178–AB050179</td>
</tr>
<tr>
<td>H. klossii</td>
<td></td>
<td>AB050180–AB050182</td>
</tr>
<tr>
<td>H. pileatus</td>
<td></td>
<td>AB050183–AB050185</td>
</tr>
<tr>
<td>H. moloch</td>
<td></td>
<td>AB050186–AB050187</td>
</tr>
<tr>
<td>H. syndactylus</td>
<td></td>
<td>AB050188–AB050192</td>
</tr>
<tr>
<td>H. concolor</td>
<td></td>
<td>AB050193</td>
</tr>
</tbody>
</table>

Results

Because there were considerable nucleotide differences among hominoid mtDNA 16S rRNA sequences, we divided hominoids into three groups: African apes, orangutan, and gibbons.

African Apes

We determined mtDNA sequences for 35 chimpanzee, 13 bonobo, and 10 gorilla. The following sequence data retrieved from the DDBJ/EMBL/GenBank International Sequence Database were also used for comparison: two chimpanzee sequences (D38113 and X93335), one bonobo sequence (D38113), and one gorilla sequence (X93335), and two gorilla sequences (D38116 and X93335).
Table 2A. Variant sites of chimpanzee

<table>
<thead>
<tr>
<th>Site</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11111111</td>
</tr>
<tr>
<td></td>
<td>11111212345678990833377888</td>
</tr>
<tr>
<td></td>
<td>22452863832806663603608234</td>
</tr>
<tr>
<td>Sequence</td>
<td>4570452556079027878808685607</td>
</tr>
</tbody>
</table>

- Chimps 13, 16, 18, 19, 20, 22, 25, and 29 were identical with this sequence.
- Chimps 6, 7, 10, 12, and 21 were identical with this sequence.
- Chimps 5 was identical with this sequence.
- Chimpanzee 17 was identical with this sequence.
- Chimps 23, 27, and 30 were identical with this sequence.
- Chimps 24 and 28 were identical with this sequence.
- Chimps 33 and 34 were identical with this sequence.

Table 2B. Variant sites of bonobo

<table>
<thead>
<tr>
<th>Site</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11111111</td>
</tr>
<tr>
<td></td>
<td>03734975889079545778</td>
</tr>
</tbody>
</table>

- The numbers in parentheses show the sequences retrieved from the DDBJ/EMBL/GenBank database.
- Standard error (S.E.) was calculated by the bootstrap method of MEGA2.
- Number of nucleotide sites compared.

Table 2C. Variant sites of gorilla

<table>
<thead>
<tr>
<th>Site</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>11111</td>
</tr>
<tr>
<td></td>
<td>111112346778900144</td>
</tr>
<tr>
<td></td>
<td>1223349278188937811</td>
</tr>
</tbody>
</table>

- The nucleotide diversity within species was calculated using the Jukes–Cantor correction, and the result is shown in Table 3.
- Phylogenetic relationships of the haplotypes belonging to the same species were more or less compatible with the phylogenetic networks shown in Figure 1.
- The nucleotide diversities of chimpanzee, bonobo, and gorilla (0.0027 ± 0.0007, 0.0033 ± 0.0008, and 0.0024 ± 0.0007, respectively) were greater than that for human (0.0016 ± 0.0004), in which sequence data of the 53 complete mitochondrial genomes (Ingman et al. 2001) were used.

Phylogenetic networks for each species were constructed based on the multiply aligned sequence data. We determined mtDNA sequences for 16 orangutans. Orangutan sequence data retrieved from the DDBJ/EMBL/GenBank International Sequence Database were also

The nucleotide diversity of chimpanzee sequences (D38114 and X93347). Nucleotide sequences with accession numbers starting with D were determined by Horai et al. (1995), while those with accession numbers starting with X were determined by Xu and Arnason (1996). Those sequence data were multiply aligned for each species, and variant nucleotide sites are presented in Table 2A, B, and C for chimpanzee, bonobo, and gorilla, respectively.

Phylogenetic networks for each species were constructed based on the multiply aligned sequence data (Figure 1). Four chimpanzee individuals (nos. 2, 4, 5, and 9) were clearly separated from the remaining individuals. Arrows indicate the position of the roots based on the phylogenetic tree of four hominid species shown in Figure 2. Networks for chimpanzee (Figure 1A) and for gorilla (Figure 1C) happened to be unrooted trees, for all the nucleotide sites were compatible with each other. One rectangle in the bonobo network (Figure 1B) denotes the existence of incompatible nucleotide configurations; the configuration at site 1788 is incompatible with those at sites 225, 687, and 1368. Subspecies identification is not clear for all the chimpanzees used in this study, but most of them seem to be Pan troglodytes verus according to their mtDNA control region sequences (Shinoda K, personal communication). Therefore it is possible that the small cluster with four chimpanzees may correspond to another subspecies, probably Pan troglodytes troglodytes.

We also constructed a phylogenetic tree for human, chimpanzee, bonobo, and gorilla (Figure 2). Two human sequences that showed the greatest divergence out of the 53 complete mitochondrial genomes (Ingman et al. 2001) were included in this tree. The tree was rooted by using the midpoint rooting method. The phylogenetic relationship of the four species is compatible with that based on the complete mtDNA sequences (Horai et al. 1995). Phylogenetic relationships of the haplotypes belonging to the same species were more or less compatible with the phylogenetic networks shown in Figure 1.

The nucleotide diversity within species was calculated using the Jukes–Cantor correction, and the result is shown in Table 3. The nucleotide diversities of chimpanzee, bonobo, and gorilla (0.0027 ± 0.0007, 0.0033 ± 0.0008, and 0.0024 ± 0.0007, respectively) were greater than that for human (0.0016 ± 0.0004), in which sequence data of the 53 complete mitochondrial genomes (Ingman et al. 2001) were used. It should be noted that individuals sampled from the global scale were used for human.

**Orangutan**

We determined mtDNA sequences for 16 orangutans. Orangutan sequence data retrieved from the DDBJ/EMBL/GenBank International Sequence Database were also
used for comparison: D38115 determined by Horai et al. (1995) for a Bornean orangutan (P. p. pygmaeus) and X97707 determined by Xu and Arnason (1996) for a Sumatran orangutan (Pongo p. abelii). Sequences for ora#001, ora#002, ora#003, ora#004, ora#006, ora#007, ora#014, and ora#015 determined in the present study were closely related with the Bornean orangutan sequence, while those for ora#005, ora#008, ora#009, ora#010, ora#011, ora#012, ora#013, and ora#016 were similar to the Sumatran orangutan. These sequence features were compatible with the source of the individuals.

The nucleotide difference between these two groups is substantial. There are 54 nucleotide differences in the 1.7 kb region that contains the complete 16S rRNA gene. We therefore concluded that those two sequence groups belong to two subspecies of orangutan. These two groups of sequences were multiply aligned, as shown in Table 4. Phylogenetic networks for each subspecies were constructed based on the multiply aligned sequence data (Figure 3). Both networks turned out to be unrooted trees because of no reticulation. We also constructed a phylogenetic tree for the two orangutan subspecies (Figure 4). As expected, individuals belonging to the two subspecies formed clear monophyletic groups, separated by a branch with 100% bootstrap probability.

Nucleotide diversities within subspecies are 0.0034 ± 0.0008 and 0.0066 ± 0.0013 for Bornean and Sumatran orangutans, respectively (see Table 3). Therefore Sumatran orangutan was shown to have about two times higher mtDNA diversity than Bornean orangutan. It is interesting to note that nucleotide diversities for those two subspecies are greater than those for African great apes and for human, even if we treated the two subspecies separately.

**Gibbons**

We determined the mtDNA 16S rRNA sequences for 22 individuals of gibbons. Because the number of individuals examined was small for most of the species (see Table 1), we present variant sites only for
siamangs (Table 5). Figure 5 shows the phylogenetic network for those five siamang sequences. We observed one pair of incompatible configurations: one for site 1199 and the other for sites 191 and 668. This is why there is one rectangle.

Figure 5 shows the phylogenetic tree of the 22 gibbon sequences and a Hylobates lar sequence determined by Xu and Arnason (1996; DDBJ/EMBL/GenBank accession number X99256). The root was determined by the outgroup sequences (human, chimpanzee, bonobo, gorilla, and orangutan). We covered three subgenera of gibbons in this study (see Table 1), and the subgenus Symphalangus made a clear monophyletic cluster. Subgenus Hylobates also formed a monophyletic group, except for one individual (no. 3) of H. pileatus. This H. pileatus individual was very close to a H. concolor individual (see Figure 6). Because H. concolor belongs to the subgenus Nomascus (Gloves 1997), this clustering was unexpected. This kind of species intermingling was also observed within the subgenus Hylobates. H. agilis individual no. 1 clustered with H. klossi no. 2, while H. agilis no. 3 clustered with two H. muelleri individuals. H. klossi also showed nonmonophyly.

We therefore estimated net nucleotide substitutions or evolutionary distance between species by subtracting within-species variation (Nei 1987). UPGMA was used for reconstructing the species phylogeny. This is because the branch lengths of a species tree should be proportional to evolutionary time. Figure 7A shows a UPGMA tree for eight gibbon species. Because one individual of H. pileatus clustered with a H. concolor individual in the phylogenetic tree of mtDNA genes (see Figure 6), we also constructed a UPGMA tree after eliminating this H. pileatus individual (tree in Figure 7B). Only the position of H. pileatus differs between the two trees.

**Discussion**

**Higher Nucleotide Diversity of Aps Than Human**

All the ape species used in the present study, except for H. concolor in which only one individual sample was available, showed higher nucleotide diversity than for human. Although a limited number of individuals were used in the present study, this finding is consistent with that of previous studies for mtDNA (e.g., Gagneux et al. 1999; Goldberg and Ruvolo 1997; Wise et al. 1997). This tendency toward higher nucleotide diversity in apes than human was also observed for nuclear DNA (Kaessmann et al. 1999; Kitano et al. 2000; Sumiyama et al. 2000). This is indirect support for the recent expansion model on modern humans. We compared the same set of 53 human individuals for the whole D-loop region of mtDNA of about 1.1 kb using Ingman et al.’s (2001) data. Nucleotide diversity for the control region was estimated to be 0.0158 ± 0.0017, which is 10 times higher than that (0.0016 ± 0.0004) for the 16S rRNA region. Because the same set of individuals was used for these two computations, the source of the difference comes from that of the evolutionary rate. If we assume that the same relationship holds for ape mtDNA sequence evolution, mtDNA D-loop diversity values for our sample are expected to be about 10 times those shown in Table 3.

**Relationship of Bornean and Sumatran Orangutan**

The two orangutan subspecies (Pongo pygmaeus pygmaeus and P. p. abelii) differ cytogenetically by a pericentric inversion in chromosome 2, a specific inversion in chromosome Y, and Sumatran Y carries a distally located nuclear organizing region that is absent in Bornean orangutan, but they interbreed in captivity and produce fertile offspring (Xu and Arnason 1996). Genetic differences between and within Bornean and Sumatran orangutans have been studied by various authors. Zhi et al. (1996) examined mtDNA restriction fragment length polymorphisms (RFLPs), partial sequences of the 16S rRNA gene, and nuclear minisatellite polymorphism. Bornean orangutan (and P. p. pygmaeus) clearly separated from Sumatran orangutan (P. p. abelii) in the mtDNA genealogy. This dichotomous pattern was also observed by Warren et al. (2000), who used mtDNA control region sequences. We also observed two distinct lineages in orangutan (Figure 4). However, Muir et al. (2000) presented a somewhat different picture using NADH dehydrogenase subunit 3 and cytochrome b gene sequences. Bornean

---

**Table 5. Variant sites of siamang (Hylobates syndactylus)**

<table>
<thead>
<tr>
<th>Consensus</th>
<th>Site position</th>
</tr>
</thead>
<tbody>
<tr>
<td>H. syndactylus</td>
<td>H1111</td>
</tr>
<tr>
<td>H. syndactylus</td>
<td>11122346688911257</td>
</tr>
<tr>
<td>H. syndactylus</td>
<td>12889905075612369082</td>
</tr>
<tr>
<td>H. syndactylus</td>
<td>1234783951978447868069703</td>
</tr>
</tbody>
</table>

---

**Figure 5.** Phylogenetic network of the mtDNA 16S rRNA gene region for H. syndactylus (siamang). Edge lengths are proportional to the number of nucleotide differences. Arrows indicate the edges in which roots are probably located.
haplotypes formed a compact cluster, while Sumatran haplotypes showed much larger heterogeneity, and the root of the orangutan gene tree is within the Sumatran group (Figure 4 of Muir et al. 2000). Because sampled individuals differ from study to study, in the future it may be necessary to study more individuals as well as nuclear sequence data to clarify the relationship between Bornean and Sumatran orangutans.

The same problem applies to within-subspecies sequence variation. Warren et al. (2001) studied many Bornean orangutan individuals and found a considerable amount of mtDNA variation in Bornean orangutan. On the other hand, Muir et al. (2000) found much greater genetic variation among Sumatran orangutan than Bornean orangutan. Our present results are intermediate between the results of Warren et al. (2000) and those of Muir et al. (2000); nucleotide diversity (0.0066 ± 0.0013) for Sumatran orangutan was about twice as high as that (0.0034 ± 0.0008) for Bornean orangutan. Zhi et al. (1996) estimated nucleotide diversity based on their mtDNA RFLP data, and Sumatran orangutan showed more than five times higher diversity (0.0175) than Bornean orangutan (0.0033). However, their mtDNA 16S rRNA sequence data seem to show higher variation for Bornean orangutan than Sumatran orangutan. A slightly higher heterozygosity was observed for Bornean orangutan from DNA fingerprint data (Zhi et al. 1996).

**Phylogenetic Relationship of Gibbon Species**

Hayashi et al. (1995) used a single individual for six gibbon species and estimated the phylogenetic tree of those gibbon species. The five species of subgenus *Hylobates* (*H. lar*, *H. klossii*, *H. agilis*, *H. moloch*, and *H. pileatus*) formed a clear monophyletic cluster in their tree. We also observed a clear clustering of the *Hylobates* species subgenus, except for one individual of *H. pileatus* (Figure 6). However, the relationship within this subgenus is unclear. *H. lar* and *H. klossii* formed a monophyletic group with 100 bootstrap probability in Hayashi et al. (1995). Two individuals of *H. klossii* clustered with *H. lar* in our results, but one *H. klossii* individual clustered with two *H. agilis* individuals (Figure 6). There are several competing hypotheses on the phylogenetic relationship of species within subgenus *Hylobates* (Geissmann 1995). Because these gibbon species are closely related, hybridization is possible, and this may be the source of incongruent phylogenetic relationships, depending on the individuals used from the same species. Another problem is species identification. It is not so easy to identify gibbon species (Geissmann 1995).

It is known that a mating of siamang and common gibbon produced a viable offspring (Myers and Shafer 1979). Therefore there is a possibility of introgression, even between subgenera. However, a close similarity between one individual of *H. pileatus* and *H. concolor* is problematic, for the geographic distributions of *H. concolor* and *H. pileatus* does not overlap (Geissmann 1995). Further study will be necessary on this matter.

The phylogenetic relationship of the subgenera of Hylobatidae is also not yet established. *H. concolor* of *Nomascus* was located as basal in Hayashi et al. (1995),
and *H. gabriellae* and *H. leucogenys*, both belonging to subgenus *Nomascus*, were also basal in Roos and Geissmann (2001). However, our gene tree for gibbons (Figure 6) showed that *H. syndactylus*, belonging to subgenus *Symphalangus*, was basal. Because of the shorter branch lengths for the *H. syndactylus* lineage, *H. concolor* became basal in the species tree estimated by using UPGMA (Figure 7). More sequence data from nuclear DNA as well as mtDNA are necessary to clarify this higher order phylogenetic relationship of gibbons.

**References**


Corresponding Editor: Oliver A. Ryder

496 The Journal of Heredity 2001:92(6)