Molecular Phylogeny of the Major Hylobatid Divisions

Christian Roos* and Thomas Geissmann†

*AG Primatengenetik, Deutsches Primatenzentrum, Kellnerweg 4, 37077 Göttingen, Germany; and †Institut für Zoologie, Tierärztliche Hochschule Hannover, Bünteweg 17, 30559 Hannover, Germany

Received October 17, 2000; revised January 8, 2001

INTRODUCTION

Although the monophyly of the gibbons (family Hylobatidae) is widely accepted, this is not the case for the taxonomy adopted within the family. In early studies on gibbon systematics, the Hylobatidae were grouped into two distinct genera, including the siamang (Symphalangus) on the one hand and all the remaining gibbons (Hylobates) on the other (e.g., Napier and Napier, 1967; Schultz, 1933; Simonetta, 1957). When gibbons were studied in more detail, however, it became clear that four, not two, major hylobatid divisions needed to be recognized. These groups are generally accepted now as four distinct subgenera (i.e., Symphalangus, Nomascus, Bunopithecus, and Hylobates) (Geissmann, 1994, 1995; Marshall and Sugardjito, 1986; Nowak, 1999; Prouty et al., 1983; Rowe, 1996).

The phylogenetic relationships among these four major groups are still unknown. Most previous studies have been based on morphology, vocalization, electrophoretic protein evidence, and karyotyping and have differed in their conclusions (Bruce and Ayala, 1979; Creel and Preuschoft, 1984; Geissmann, 1993, 2001; Groves, 1972; Haimoff et al., 1982; Liu et al., 1987; Shafer, 1986). Even the use of molecular techniques based mainly on mitochondrial DNA sequences was not able to resolve the evolutionary relationships among the gibbon subgenera (Garza and Woodruff, 1992; Hall et al., 1998; Hayashi et al., 1995; Zehr, 1999; Zhang, 1997). Furthermore, most molecular studies did not include the subgenus Bunopithecus and therefore presented an incomplete view on gibbon evolution (Garza and Woodruff, 1992; Hayashi et al., 1995).

The mitochondrial control region is known to evolve faster than other parts of mtDNA and may therefore be more suited to resolve a radiation which evolved over a short time span than sequences used in previous studies (Garza and Woodruff, 1992; Hall et al., 1998; Hayashi et al., 1995; Zehr, 1999; Zhang, 1997). We have therefore determined the DNA sequence of the complete mitochondrial control region and adjacent phenylalanine-tRNA (Phe-tRNA) of the four gibbon subgenera with the intention of (1) resolving the evolutionary relationships between the subgenera and (2) comparing the distances between them with those between the great ape genera Homo and Pan.

MATERIALS AND METHODS

Five gibbon species representing the four major groups within the Hylobatidae clade, Hylobates, Bunopithecus, Symphalangus, and Nomascus, were studied. All study animals or their parents were identified by us using fur coloration and vocal characteristics as described in Geissmann (1995).

Sequence Determination

DNA was extracted from peripheral blood lymphocytes and hair samples (H. hoolock) by the standard methods outlined in Sambrook et al. (1989) and Walsh et al. (1991), respectively. The complete mitochondrial control region and adjacent Phe-tRNA from one individual each of H. (Bunopithecus) hoolock, H. (Nomascus) leucogenys leucogenys, H. (Nomascus) gabriellae, and H. (Symphalangus) syndactylus and two individuals of H. (Hylobates) lar were PCR-amplified (Saiki et al., 1988) with the oligonucleotide primers L16007 (5’-CCCAAGCTAAAATTCTAA-3’) and H00651 (5’-TA-ACTGCAGAGGCTAGGACCAAACCT-3’) according to the methods of Saiki et al. (1988).
to Kocher et al. (1989), with H and L designating the heavy- and light-strand sequences of the mitochondrial genome and the numbers indicating the 3’ end of the primers according to the human reference sequence (Anderson et al., 1981). The amplifications were carried out for 35 cycles. Each cycle consisted of a 40-s denaturation at 92°C, 40 s at 50°C for annealing, and 90 s at 72°C for extension.

The resulting PCR products were separated on 1% agarose gels and visualized by ethidium bromide staining.

The fragments of a size of about 1.2 kb were excised from the gel and the DNA was extracted with the Qiagen Gel Extraction Kit.

Direct sequencing reactions were performed with the same primers as indicated above with the Big Dye Terminator Cycle Sequencing Kit (Perkin–Elmer) following the manufacturer’s recommendations. All sequencing reactions were run on an automated ABI377 sequencer (Perkin–Elmer). The sequences determined and details of the individuals in the study presented herein are available in GenBank under the Accession Nos. AF193804 and AF311721–AF311725.

Sequence Comparisons

Considering the fast pace of sequence evolution pertaining to the mitochondrial control region, only great ape sequences were taken into account for the sequence comparisons and phylogenetic analyses. The gorilla (Gorilla gorilla, NC001645) and orang-utan (Pongo pygmaeus, NC001646) (Horai et al., 1995) sequences were excluded from the analyses since both of the sequences exhibit a major deletion in the mitochondrial control region. Thus, the herein determined sequences were compared with the homologous sequences obtained from human (Homo sapiens, NC001807) (Anderson et al., 1981), pygmy chimpanzee (Pan paniscus, NC001644) (Horai et al., 1995), and common chimpanzee (Pan troglodytes, X93335) (Arason et al., 1996).

The sequences were aligned with the Clustal W program, version 1.7 (Thompson et al., 1994), with a gap-opening penalty of 10.00 and a gap-extension penalty of 0.05. The alignments obtained in this way were afterward optimized manually. To remove poorly aligned positions and regions with a too-high divergence we applied the Gblocks software (Castresana, 2000) on the Clustal W-generated alignment. Therefore, default settings for both protein and rDNA alignments were applied.

Distance Calculations

Distances between taxa were estimated by two measures of sequence divergence in the Gblocks alignment. First, the observed proportion of base differences between taxa was calculated by SeqPup (version 0.7). Second, a maximum-likelihood (ML) estimate was obtained with the PUZZLE software (Strimmer and von Haeseler, 1996) with estimated base frequencies and transition:transversion ratios.

Phylogenetic Analyses

A priori tests of the data for the presence of a phylogenetic signal were carried out with the likelihood-mapping option included in PUZZLE. Phylogenetic trees were constructed based on three algorithms: maximum-parsimony (MP) (Fitch, 1971) and neighbor-joining (NJ) (Saitou and Nei, 1987), included in PHYLIP, version 3.5c (Felsenstein, 1993) and maximum-likelihood as implemented in PUZZLE, version 4.0.2 (Strimmer and von Haeseler, 1996).

Distance corrections for the NJ analysis were carried out with the ML distance correction and transition:transversion ratios as estimated in PUZZLE. For ML reconstructions the HKY (Hasegawa et al., 1985) and the TN (Tamura and Nei, 1993) models both assuming uniform rate of sequence evolution and rate heterogeneity across sites, were used.

Support of internal branch length was either determined by bootstrap analyses (MP and NJ) performed with 1000 replicates or indicated by the ML quartet puzzling support values (1000 puzzling steps). The 50%-majority rule consensus trees were calculated with CONSENSE of the PHYLIP package. To check the significance of the differences between the log likelihoods of alternative trees we used the Kishino–Hasegawa test in PUZZLE giving alternative intrees and manually calculated the significance level of each tree.

RESULTS

The lengths of the mitochondrial control regions are 1028 bp (H. lar 1 and H. tar 2), 1059 bp (H. hoolock), 1064 bp (H. syndactylus), 1011 bp (H. gabriellae), and 1031 bp (H. leucogenys).

To determine the phylogenetic affiliations among the gibbon subgenera and the hominids, sequences of all the gibbons, human, and common and pygmy chimpanzees were aligned. The complete alignment of the control region and adjacent Phe-tRNA generated by Clustal W is 1238 bp in length. After the exclusion of gaps and poorly aligned positions by Gblocks with the more stringent defaults for proteins alignments, the final alignment was 857 bp in length with 574 (67%) constant sites (Fig. 1).

The observed base differences between the four gibbon subgenera in the Gblocks alignment range from 8.98 to 13.83% and from 9.57 to 9.68% between Pan and Homo (Table 1).

Initially we checked for the presence of a phylogenetic signal by conducting a likelihood-mapping analysis as implemented in PUZZLE. A starlike evolution was excluded by a strong phylogenetic signal indicat-
FIG. 1. Complete alignment of the mitochondrial control region (upper case) and adjacent phenylalanine-tRNA (lower case) with dashes (-) indicating deletions. In the row "Gblocks" dots (.) indicate included and asterisks (*) indicate excluded positions from the final alignment.
<table>
<thead>
<tr>
<th>Genus</th>
<th>DNA Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Homo sapiens</em></td>
<td>TCTCGTCC-CT-AGTGAAGACCAATTCCTGTGTACTTTTATGAGGAACTTTGGTACCTACCGCAACAGAGGATCTACCTTTGCGCTTCGGCCATTACACCGTCTGGG</td>
</tr>
<tr>
<td><em>Pan paniscus</em></td>
<td>AGCTGAGGCAGGCTCTACACAGGCTGTTCTTTTATGAGGAACTTTGGTACCTACCGCAACAGAGGATCTACCTTTGCGCTTCGGCCATTACACCGTCTGGG</td>
</tr>
<tr>
<td><em>P. troglodytes</em></td>
<td>CCGTGGCCAGCGGATCCACAGGCTGTTCTTTTATGAGGAACTTTGGTACCTACCGCAACAGAGGATCTACCTTTGCGCTTCGGCCATTACACCGTCTGGG</td>
</tr>
<tr>
<td><em>Hylobates (H. lar)</em></td>
<td>TCACTTTTGAGTGTATTTTATGAGGAACTTTGGTACCTACCGCAACAGAGGATCTACCTTTGCGCTTCGGCCATTACACCGTCTGGG</td>
</tr>
<tr>
<td><em>H. (H. lar)</em></td>
<td>TCACTTTTGAGTGTATTTTATGAGGAACTTTGGTACCTACCGCAACAGAGGATCTACCTTTGCGCTTCGGCCATTACACCGTCTGGG</td>
</tr>
<tr>
<td><em>H. (Banopithecus)</em></td>
<td>TCACTTTTGAGTGTATTTTATGAGGAACTTTGGTACCTACCGCAACAGAGGATCTACCTTTGCGCTTCGGCCATTACACCGTCTGGG</td>
</tr>
<tr>
<td><em>H. (Sympalangus)</em></td>
<td>TCACTTTTGAGTGTATTTTATGAGGAACTTTGGTACCTACCGCAACAGAGGATCTACCTTTGCGCTTCGGCCATTACACCGTCTGGG</td>
</tr>
<tr>
<td><em>H. (Nomascus)</em></td>
<td>TCACTTTTGAGTGTATTTTATGAGGAACTTTGGTACCTACCGCAACAGAGGATCTACCTTTGCGCTTCGGCCATTACACCGTCTGGG</td>
</tr>
</tbody>
</table>

**Fig. 1—Continued**

**MOLECULAR PHYLOGENY OF THE GIBBONS**
FIG. 1—Continued
ing the suitability of the data for phylogenetic reconstructions.

Tree reconstructions were carried out by the maximum-parsimony, neighbor-joining, and maximum-likelihood methods (Fig. 2). All three algorithms separate the Homo–Pan clade from the Hylobatidae with a bootstrap value of 100%. Within each of the subgenera Hylobates and Nomascus, two representatives were analyzed. H. lar 1 and H. lar 2, which represent Hylobates, form a sister clade, as do the representatives of Nomascus, H. gabriellae and H. leucogenys. Each clade is supported by bootstrap values of 100% (for the ML, MP, and NJ methods). Furthermore, all three algorithms display Nomascus as the deepest split, with a support of 100, 95, and 88% for ML, NJ, and MP trees, respectively. Discordant phylogenetic relationships among Hylobates, Bunopithecus, and Symphalangus were obtained by the three applied algorithms: whereas maximum-parsimony links Bunopithecus with Symphalangus as sister taxa, neighbor-joining and maximum-likelihood result in a monophyly of Bunopithecus and Hylobates. Bootstrap values for these relationships obtained from the maximum-likelihood and neighbor-joining algorithms are 99 and 66%, respectively, whereas the Bunopithecus–Symphalangus clade displayed by maximum-parsimony is supported by a bootstrap value of only 43%.

To determine the log likelihood differences between alternative ML trees with one representative of each subgenus, the 15 possible patterns of relationship among the subgenera were tested. The log likelihood ranged from $-2633.65$ to $-2667.67$ (Table 2).

The above-described trees are based on the Gblocks alignment with defaults for protein alignments, and in the case of maximum-likelihood they are based on the TN model, with the assumption of uniform rate of sequence evolution. Furthermore, trees were reconstructed with all three algorithms with a Gblocks alignment generated with less stringent parameters. In addition to the described ML tree, maximum-likelihood reconstructions were carried out with the HKY model with the assumptions of uniform rate of sequence evolution and rate heterogeneity across sites. All these alternative trees revealed the same topology and differed only by branch lengths and bootstrap values.

**DISCUSSION**

Previous molecular studies failed to resolve the relationships among the four gibbon subgenera. In contrast, sequences of the mitochondrial control region and adjacent Phe-tRNA represented in our study depict evolutionary relationships among the gibbon subgenera that are supported by high bootstrap values. Even if trees are reconstructed with other models or a Gblocks alignment obtained from less stringent parameters, the resulting trees always show the same
topology and differ only in their branch lengths and bootstrap values. Nevertheless, it should be noted that statistical tests of the 15 possible patterns of evolutionary relationship by the method of Kishino and Hasegawa (1989) revealed two trees that did not differ significantly at \( P < 0.07 \) from the best ML tree (Table 2).

However, the expected monophilies of both the Homo–Pan and the gibbon clades are supported by bootstrap values of 100%. Within the Hylobatidae, Nomascus is the most basal group, followed by Symphalangus, whereas Bunopithecus and Hylobates were the last to diverge. Nomascus as the deepest split is supported by high bootstrap values in all three tree reconstruction methods used. Results about the relationships among Bunopithecus, Hylobates, and Symphalangus are contradictive; whereas maximum-likelihood and neighbor-joining link Bunopithecus and Hylobates, maximum-parsimony groups Bunopithecus with Symphalangus.

The branches in the ML and NJ trees separating Symphalangus from the Bunopithecus–Hylobates clade are supported by high bootstrap values, at least for ML. In contrast, the different phylogeny obtained from maximum-parsimony is supported by a bootstrap value of only 43% and, hence, the relationship among Bunopithecus, Hylobates, and Symphalangus in the MP tree is probably best described as an unresolved trichotomy.

Because of the fast pace and peculiar mode of sequence evolution in the mitochondrial D-loop, we expect a considerable amount of homoplasious exchanges (Stewart, 1993; Wiens and Servedio, 1998). We, therefore, believe that the model of sequence evolution is better approximated in the NJ and ML than in the MP analyses. Hence, the monophyly of Bunopithecus and Hylobates obtained from the maximum-likelihood and neighbor-joining methods is more likely to reflect the real relationship.

Before we can try to evaluate the taxonomic implications of our findings, two difficulties need to be considered. First, an attempt to accurately date the divergence of the four gibbon subgenera from our data is hampered by the fact that the mitochondrial control region does not evolve in a clock-like manner, as could be seen by comparison of the likelihoods of different trees reconstructed under the assumption of an existing or absent molecular clock (data not shown). Second, the proposal to adopt a time-standardized taxonomic classification with taxa of equivalent age holding the same taxonomic rank (Goodman et al., 1998; Avise and Johns, 1999) is difficult to apply to our results because of the significantly different generation times that ex-

**TABLE 1**

Distances in the Gblocks Alignment among Great Ape Species and Gibbon Subgenera\(^ a \)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homo sapiens (1)</td>
<td>—</td>
<td>0.0957</td>
<td>0.0968</td>
<td>0.1966</td>
<td>0.1762</td>
<td>0.1832</td>
<td>0.1779</td>
</tr>
<tr>
<td>Pan troglodytes (2)</td>
<td>0.1035</td>
<td>—</td>
<td>0.0723</td>
<td>0.2025</td>
<td>0.1809</td>
<td>0.1902</td>
<td>0.1908</td>
</tr>
<tr>
<td>Pan paniscus (3)</td>
<td>0.1049</td>
<td>0.0768</td>
<td>—</td>
<td>0.2176</td>
<td>0.1949</td>
<td>0.2042</td>
<td>0.2013</td>
</tr>
<tr>
<td>Hylobates (4)</td>
<td>0.2385</td>
<td>0.2481</td>
<td>0.2714</td>
<td>—</td>
<td>0.1033</td>
<td>0.1237</td>
<td>0.1383</td>
</tr>
<tr>
<td>Bunopithecus (5)</td>
<td>0.2088</td>
<td>0.2160</td>
<td>0.2372</td>
<td>0.1120</td>
<td>—</td>
<td>0.0898</td>
<td>0.1243</td>
</tr>
<tr>
<td>Symphalangus (6)</td>
<td>0.2183</td>
<td>0.2293</td>
<td>0.2511</td>
<td>0.1372</td>
<td>0.0960</td>
<td>—</td>
<td>0.1214</td>
</tr>
<tr>
<td>Nomascus (7)</td>
<td>0.2119</td>
<td>0.2308</td>
<td>0.2471</td>
<td>0.1570</td>
<td>0.1380</td>
<td>0.1343</td>
<td>—</td>
</tr>
</tbody>
</table>

\(^ a \) Values represent substitutions per position. Above the diagonal are observed distances; below the diagonal are ML distances corrected with the Tamura–Nei model and an estimated transition:transversion ratio of 2.34. Distances for gibbons are an average for each subgenus.

**FIG. 2.** The 50%-majority rule consensus trees for the maximum-parsimony (a), neighbor-joining (b), and maximum-likelihood (c) methods. The maximum-likelihood tree is unrooted. Branch lengths are drawn according to the number of substitutions per position, with the bar indicating 0.1 substitutions per site.
ist among the groups under comparison (gibbons at about 8 years vs Hominidae at about 12 years; e.g., Rowe, 1996).

However, by considering the pure observed sequence distances and not taking different generation times into account, it is obvious that the distances among the four gibbon subgenera are in the same range as those between Homo and Pan, or even higher. The uncorrected average distances are 10.3% between Hylobates and Bunopithecus, 10.6% between Symphalangus and the Bunopithecus–Hylobates clade, and 12.8% between Nomascus and the other three subgenera. In contrast, the distance between Homo and Pan is only 9.6%. Based on these findings, it would be justified to elevate all four gibbon subgenera to genus rank.

Our data depict evolutionary relationships among the gibbon genera that are supported by high bootstrap values. However, more extensive stretches of mitochondrial and nuclear DNA may need to be sequenced to definitively establish the branching order of the four gibbon clades.

ACKNOWLEDGMENTS

We thank the staff and officials of the zoos in Leipzig, Munich, Besançon, Twycross, Wuppertal, and Perth for providing the gibbon samples for this study. We are grateful to H. Zischler and T. Metzinger for helpful discussions and support.

REFERENCES


