Letter to the Editor

Phylogenetic Position of Guinea Pigs Revisited

Ying Cao,∗† Norihiro Okada,† and Masami Hasegawa∗

∗The Institute of Statistical Mathematics, Tokyo, Japan; and †Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Japan

Although the traditional taxonomy, mainly based on comparative morphology, classifies the guinea pig into the order Rodentia (Luckett and Hartenberger 1985; Novacek 1992), the molecular phylogenetics has been challenging this view. The first such attempt was made by Graur, Hide, and Li (1991), who suggested that the order Rodentia may not be monophyletic, and that the guinea-pig-like rodents (Caviomorpha) may be an outgroup to the clade formed by the rat-like rodents (Myomorpha) and Primates. Their suggestion contradicts the traditional classification of the guinea pig into Rodentia (the Myomorpha/Caviomorpha clade). These authors used a maximum-parsimony (MP) method in estimating the tree, but it is known that the MP method is sometimes misleading, particularly when the evolutionary rate differs among lineages (Felsenstein 1978). Therefore, Hasegawa et al. (1992) reexamined the data by a maximum-likelihood (ML) method for protein sequences (Kishino, Miyata, and Hasegawa 1990) that is robust against the violation of rate constancy (e.g., Hasegawa, Kishino, and Saitou 1991; Hasegawa and Fujiwara 1993). They showed that, although the MP analysis supported Graur, Hide, and Li’s tree of the rodent polyphyly with as high as 96% bootstrap probability (BP), the result of ML analysis was equivocal, and therefore they suggested that Graur, Hide, and Li’s tree might represent an example of the fact that unequal evolutionary rates can mislead MP analysis. More extensive ML analyses, mainly of nuclear DNA-encoded proteins, strongly supported the traditional tree (Cao et al. 1994b; Kuma and Miyata 1994).

On the basis of phylogenetic analyses of the complete mitochondrial genome from the guinea pig and 14 other eutherian species with opossum as an outgroup, D’Erchia et al. (1996) recently challenged the traditional view again, but proposed the third alternative hypothesis from their NJ analysis and the traditional tree could not be dismissed. They also applied the protein ML to the mitochondrial proteins, and, although they obtained a ML tree consistent with those yielded by MP and NJ, the statistical significance of the ML tree was not evaluated. Since this issue is critical to the higher-level evolution of mammals, we here present a detailed analysis of D’Erchia et al.’s data by the ML and show that the support of their hypothesis is not as strong as they claim.

Two data sets were analyzed in this study. First, we analyzed the aligned data of D’Erchia et al. (1996), which consist of 15 eutherian species with opossum as an outgroup. The insertion/deletion sites were excluded from the analyses. For the rRNAs, regions where alignment was ambiguous were also excluded. By applying the ProtML (protein ML) and NucML (nucleic acid ML) programs in MOLPHY package (Adachi and Hasegawa 1996a) to the mitochondrial proteins and to the mitochondrial ribosomal RNAs (tRNAs), respectively, we estimated the statistical significance of their tree. Since different genotypes evolve at different rates and with different modes, the ML analyses were carried out separately for each gene, and then the estimated log-likelihoods for each gene were summed up in evaluating the total evidence. ProtML with the mREV24-F model (Adachi and Hasegawa 1996a, 1996b) was applied to the proteins, and NucML with the HKY85 model (Hasegawa, Kishino, and Yano 1985) was applied to the tRNAs. The TotalML program in MOLPHY was used to sum up the log-likelihood.

Secondly, since phylogenetic inference generally depends on the species sampling (e.g., Philippe and Douzery 1994), ProtML analysis using additional data of complete mtDNA sequences from cat (Lopez et al. 1996: database accession number U20753), blue whale (Arnason and Gullberg 1993: X72204), and platypus (Ianke et al. 1996: X83427) was also carried out to examine whether the inclusion of the additional species affects the result. This data set consists of 17 eutherian species with opossum and platypus as an outgroup.

Figure 1 gives the ML tree of the tandemly aligned amino acid sequences of the 13 mitochondrial proteins for the 17 eutherian species with opossum and platypus as an outgroup. This was estimated by using the local rearrangement option of the ProtML with several alternative initial tree topologies, and the figure on each branch represents a local bootstrap probability (LBP; Adachi and Hasegawa 1996a), which is a relative bootstrap frequency among the three alternative relationships relevant to the branch when the relationships within the three subtrees attached to the particular branch are fixed as shown in the figure.

It might be noteworthy that, although D’Erchia et al.’s (1996) hypothesis with respect to the relationships among guinea pig, Myomorpha, and the other eutherians (except hedgehog) is preferred also in figure 1, the LBP...

Key words: molecular phylogeny, maximum likelihood, guinea pig, rodents, mitochondrial DNA.

Address for correspondence and reprints: Masami Hasegawa, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan. E-mail: hasegawa@ism.ac.jp.

© 1997 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038
Table 1
Maximum-Likelihood (ML) Analyses of the Mitochondrial DNA Data

<table>
<thead>
<tr>
<th>Tree</th>
<th>ATP6</th>
<th>ATP8</th>
<th>COX1</th>
<th>COX2</th>
<th>COX3</th>
<th>Cyt-b</th>
<th>ND1</th>
<th>ND2</th>
<th>ND3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I...</td>
<td>ML</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II...</td>
<td></td>
<td>-0.9 ± 3.3</td>
<td>-7.2 ± 8.4</td>
<td>-3.5 ± 3.1</td>
<td>-0.2 ± 5.9</td>
<td>-5.0 ± 4.4</td>
<td>-4.3 ± 5.9</td>
<td>-8.1 ± 5.5</td>
<td>-2.8 ± 5.9</td>
</tr>
<tr>
<td>III.</td>
<td>-3.6 ± 3.6</td>
<td>ML</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: — Trees I, II, and III refer, respectively, to the traditional tree (the Myomorpha/Caviomorpha clade), D’Erchia et al.’s (1996) tree (the Caviomorpha/Primates clade), and Graur, Hide, and Liv’s (1991) tree (the Myomorpha/Primates clade) among Caviomorpha, Myomorpha, and Primates. The highest-likelihood tree is indicated as “ML,” and the differences of log-likelihood of alternative trees from that of the ML tree are shown with their SEs estimated by Kishino and Hasegawa’s (1989) formula. The BPS (%) were estimated by the RELL method (Kishino, Miyata, and Hasegawa 1990) with 10^4 replications. For D’Erchia et al.’s (1996) data, “subtotal” and “total” refer, respectively, to total of proteins and of proteins + rRNAs. Transition/transversion (o/B) ratio of the HKY85 model was determined so as to maximize the likelihood (3.9 and 2.7 for 12S and 16S rRNAs). For the data set of the 17 eutherian species with oppossum and platypus as outgroup, “total” refers to total of proteins. The aligned data are available by anonymous ftp in sumnh.ism.ac.jp(133.58.12.20):/pub/data/.

Our reanalysis of D’Erchia et al.’s (1996) data showed that tree II (D’Erchia et al.)’s is preferred but with BPS of only 56% and 60% from the proteins and from the total of proteins and rRNAs, respectively, while the BPs for tree I (traditional) are 13% and 8% from the respective analyses (table 1a).

From the ProtML analysis of the second data set (with the three additional species), we obtained BPs of 32%, 58%, and 10% BP for trees I, II, and III, respectively, for this relationship is only 56%. The association of Cetacea and Artiodactyla and the existence of the Carnivora/Perissodactyla/Artiodactyla/Cetacea clade are consistently strongly supported by the previous works (e.g., Milinkovitch, Orti, and Meyer 1993; Cao et al. 1994a; D’Erchia et al. 1996; Stanhope et al. 1996). Furthermore, the sister group status of hedgehog to the other eutherians is strongly supported in accord with Krettek, Gullberg, and Arnason (1995). Since the LBPs for the placing of Lagomorpha relative to Primates and the Carnivora/Perissodactyla/Artiodactyla/Cetacea clade and for the placing of Perissodactyla relative to Carnivora and the Artiodactyla/Cetacea clade are not high enough to be conclusive (45% and 67%, respectively), and therefore trifurcations with respect to these relationships, that is, (Primates, (Carnivora, Perissodactyla, (Artiodactyla, Cetacea)), Lagomorpha), are assumed in the following analyses, which take into account the differences of tempo and mode among different genes.

Our reanalysis of D’Erchia et al.’s (1996) data showed that tree II (D’Erchia et al.)’s is preferred but with BPs of only 56% and 60% from the proteins and from the total of proteins and rRNAs, respectively, while the BPs for tree I (traditional) are 13% and 8% from the respective analyses (table 1a).

From the ProtML analysis of the second data set (with the three additional species), we obtained BPs of 32%, 58%, and 10% BP for trees I, II, and III, respectively.
Extended 

Table 1

<table>
<thead>
<tr>
<th>ND4</th>
<th>ND4L</th>
<th>ND5</th>
<th>ND6</th>
<th>Subtotal</th>
<th>BP</th>
<th>12S</th>
<th>16S</th>
<th>Total</th>
<th>BP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. For D’Erchia et al.’s (1996) Data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.5 ± 7.3</td>
<td>ML</td>
<td>-4.2 ± 5.0</td>
<td>ML</td>
<td>-13.8 ± 19.8</td>
<td>13</td>
<td>ML</td>
<td>15.3 ± 9.0</td>
<td>-21.7 ± 21.7</td>
<td>8</td>
</tr>
<tr>
<td>10.1 ± 7.4</td>
<td>ML</td>
<td>0.5 ± 5.5</td>
<td>ML</td>
<td>-2.5 ± 4.6</td>
<td>56</td>
<td>ML</td>
<td>-7.3 ± 6.5</td>
<td>ML</td>
<td>60</td>
</tr>
</tbody>
</table>

b. For the Data Set of the 17 Eutherian species with Opossum and Platypus as Outgroup |
| -6.3 ± 6.0 | ML | 0.1 ± 4.9 | ML | 5.3 ± 13.9 | 32 |
| -3.5 ± 6.8 | ML | -1.2 ± 5.0 | ML | -3.3 ± 4.4 | 58 |

| ML | 7.0 ± 7.0 | -3.1 ± 4.4 | 14.6 ± 14.7 | 10 |

The log-likelihood of tree I is lower than that of tree II only by 6.1 ± 13.9 (SE). The addition of the data from the three species increases the support of the traditional tree (although it is still not the ML tree). Thus, the available mitochondrial data do not provide resolution of the controversy with any statistical significance, and even the complete mtDNA sequences do not always resolve the phylogenetic problems at hand.

Previous ProtML analyses of nuclear DNA-encoded proteins supported rodent monophyly (Cao et al. 1994b; Kuma and Miyata 1994). In Table 1 of Kuma and Miyata’s paper, the ProtML analyses of 19 proteins gave only 3% BP support for tree II, while tree I (rodent monophyly) was supported by 88%. The table shows that the log-likelihood of tree I is higher than that of tree II by 37.0 ± 22.4. It is possible, as D’Erchia et al. (1996) mentioned, that paralogous comparisons were made in these studies for some of the proteins, but without the demonstration of that claim, it is inappropriate to dismiss these counter examples of D’Erchia et al.’s hypothesis. The total support for tree I from the mtDNA-encoded and nuclear DNA-encoded proteins can be evaluated by summing up the estimated log-likelihoods for both data sets (Adachi and Hasegawa 1996a), and it turned out that the log-likelihood of tree I is higher than that of tree II by 15.3 ± 31.2 using Table 1a and 30.9 ± 26.4 using Table 1b.

A retrogressed sequence (BC1 RNA), which is believed to have been adapted to a functional role of the rodent nervous system, is present both in Myomorpha and in the guinea pig, but is absent in other mammalian orders (Martignetti and Brosius 1993). This result also supports the traditional tree. Considering all the data available to date, one can only conclude that the evidence presented by D’Erchia et al. is too weak to exclude the rodent monophyly hypothesis and that this hypothesis remains most likely.

Acknowledgments

We thank Michel C. Milinkovitch for helpful comments on an earlier version of the manuscript. This work was supported by grants from the Ministry of Education, Science, Sports, and Culture of Japan.

LITERATURE CITED


NARUYA SAITOU, reviewing editor

Accepted December 24, 1996