Using Novel Phylogenetic Methods to Evaluate Mammalian mtDNA, Including Amino Acid-Invariant Sites-LogDet plus Site Stripping, to Detect Internal Conflicts in the Data, with Special Reference to the Positions of Hedgehog, Armadillo, and Elephant

Peter J. Waddell,[1,3] Ying Cao,[1] Jöerg Hauf,[2] and Masami Hasegawa[1]

[1] *Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan; E-mail: waddell@ism.ac.jp (P.J.W.),*
[2] *Scientific Research & Development GmbH, Oberurseler Str. 43, D-61440 Oberursel, Germany. Email: j.hauf@em.uni-frankfurt.de*

Abstract.—We look at the higher-order phylogeny of mammals, analyzing in detail the complete mtDNA sequences of more than 40 species. We test the support for several proposed superordinal relationships. To this end, we apply a number of recently programmed methods and approaches, plus better-established methods. New pairwise tests show highly significant evidence that amino acid frequencies are changing among nearly all the genomes studied when unvaried sites are ignored. LogDet amino acid distances, with modifications to take into account invariant sites, are combined with bootstrapping and the Neighbor Joining algorithm to account for these violations of standard models. To weight the more slowly evolving sites, we exclude the more rapidly evolving sites from the data by using "site stripping". This leads to changing optimal trees with nearly all methods. The bootstrap support for many hypotheses varies widely between methods, and few hypotheses can claim unanimous support from these data. Rather, we uncover good evidence that many of the earlier branching patterns in the placental subtree could be incorrect, including the placement of the root. The tRNA genes, for example, favor a split between the group hedgehog, rodents, and primates versus all other sequenced placentals. Such a grouping is not ruled out by the amino acid sequence data. A grouping of all rodents plus rabbit, the old Glires hypothesis, is also feasible with stripped amino acid data, and rodent monophyly is also common. The elephant sequence allows confident rejection of the older taxon Ferungulata (Simpson, 1945). In its place, the new taxa Scrotifera and Fereuungulata are defined. A new likelihood ratio test is used to detect differences between the optimal tree for tRNA versus that for amino acids. While not clearly significant as made, some results indicate the test is tending towards significance with more general models of evolution. Individual placement tests suggest alternative positions for hedgehog and elephant. Congruence arguments to support elephant and armadillo together are striking, suggesting a superordinal group composed of Xenarthra and African endemic mammals, which in turn may be near the root of the placental subtree. Thus, while casting doubt on some recent conclusions, the analyses are also unveiling some interesting new possibilities. [amino acid composition; invariant sites; LogDeterminant distances; mammal phylogeny; mitochondrial DNA genomes; Proboscidea; statistical tests; tRNA.]

Generally, the molecular trees of mammalian interordinal relationships now seem to be in closer agreement with one another than with the morphological data (e.g., Springer et al., 1997). However, there are some possible conflicts between mitochondrial (mt) DNA and nuclear data, and perhaps even within the mtDNA itself. Identifying, then resolving, phylogenetic conflicts is a major path for molecular evolutionary studies to advance along.

An important question in this regard is whether a major data partition within the mtDNA, the protein-coding genes, support the same tree as the tRNA genes, within statistical error. If the answer is no, then the models we are using for one or both of these types of data are inadequate to prevent inconsistency of tree selection—given that mtDNA genomes are expected not to undergo recombination. It is also interesting to see how the tRNA genes resolve the tree, in light of the suggestion that they can be very reliable data, not least because of their slow rate of evolution (Kumazawa and Nishida, 1993).

Evaluating Data Set Structure

A useful way to test whether two data sets could have evolved on the same tree

---

[3] Present address: Institute of Molecular Biosciences, Massey University, Palmerston North, New Zealand. Email: waddell@onyx.si.edu

is by a specific type of likelihood ratio test (Waddell, 1995:465). After making a joint estimate (using all the data) of the maximum likelihood (ML) tree, we can determine if the likelihood of the data improves significantly when we find a separate tree that best supports each data partition. A simple but conservative form of this test is to take the joint tree and test it against the individual ML tree for each data set. The use of this test is illustrated.

Herein we make critical evaluations of intriguing hypotheses suggested by mtDNA. Our findings suggest that for nearly all of these, support is much more ambiguous than previously supposed. We pioneer new methods such as the Constant Site Removal or CSR-LogDet distance correction (e.g., Waddell, 1995:ch. 3; Swofford et al., 1996:459–461; Waddell and Steel, 1997; Waddell et al., in press) applied to amino acid (AA) sequences. Use of this distance is suited to these data because it helps adjust for (1) nonstationarity—AA base composition unequal in different taxa, (2) unequal site rates, and (3) the distinctly different base compositions of the slower versus the faster evolving sites. This correction is combined with Neighbor-Joining (NJ; Saitou and Nei, 1987; Swofford et al., 1996) and with Fitch-Margoliash weighted least squares, with the constraint of all edges positive (FM+; Swofford et al., 1996; Bryant and Waddell, 1998; Waddell et al., in press). To help convince ourselves and the reader that the mtDNA trees for mammals could be seriously wrong in parts, we highlight some similar cases in early vertebrates, even when external branch lengths are much shorter (i.e., within fish).

However, we are not only knocking down. We also find good support for new phylogenetic hypotheses: including the Atlantogenata (Waddell et al., 1999), and the Fereuungulata (pronounced fer-you-ung-u-la-ta), which we define in the last section of this paper.

Evaluating Specific Hypotheses

A second aim of this work is to make updated assessments of a variety of phylogenetic hypotheses, especially those previously suggested by the mtDNA sequences. (Please note, the full scientific names of species used in these analyses appear in the materials and methods). We consider facets of the following hypotheses:

1. Marsupionta.—The grouping of marsupials and monotremes, put forward by Gregory (1947) and recently resurrected on the basis of analysis of mtDNA sequences (Janke et al., 1996). While the still quite sparse nuclear data sets are ambivalent on this point, all molecular data do point to a strong conclusion: The divergence of placentals, marsupials, and monotremes was a near trichotomy. The morphological data strongly contradict this conclusion by (a) regarding monotremes as an early offshoot of the living mammals (the Theria hypothesis) and (b) regarding paleontologists' claim that many extinct groups can confidently be placed in the period after the monotreme divergence but before the split of placentals and marsupials (suggesting this time period was considerable; e.g., Szalay, 1993; Rowe, 1993).

2. The grouping of Perissodactyls with Carnivores.—The grouping suggested by most whole mtDNA analyses of selected taxa (e.g., Xu et al., 1996, and onwards). This group comes into conflict with Cetungulata (Perissodactyla, Cetacea, and Artiodactyla; Irwin and Wilson, 1993). Cetungulata recently received support in an analysis including both mitochondrial and nuclear genes (Graur et al., 1997), but the taxon sampling was sparse.

3. The placement of Xenarthra after the divergence of hedgehogs, rodents, rabbits, and primates.—A novel hypothesis, put forward by Arnason et al. (1997) and based on mtDNA, this hypothesis conflicts with αA-crystallin data (e.g., de Jong et al., 1993). Morphologists (e.g., Gregory, 1910; McKenna, 1975; Novacek, 1993) have long supported the claim that Xenarthra is the sister taxon to all other placentals (except possibly Phoeidota). Recently, other morphologists have critiqued Novacek's analysis (Gaudin et al., 1996), concluding there is no clear-cut support for this hypothesis after correcting Novacek's data for errors.

4. The primary split of a hedgehog, Erinaceomorpha, from all other placentals.—Suggested by one complete, but idiosyncratic (in terms of high evolutionary rate and base composition), mtDNA sequence (Krettek et al., 1995). *Erinaceomorpha* has long been considered an archetypical "Insectivore" (e.g., Gregory, 1910). A recent reanalysis of the data (Sullivan and Swofford, 1997) suggests that the present hedgehog sequence can be placed all over the eutherian tree with only a minor difference in likelihood score (However, they used only 1st and 2nd positions, not AAs). A similar problem is apparent with nuclear sequences such as those for interphotoreceptor retinoid-binding protein (IRBP) and von Willebrand factor (vWF), which will, on occasion, place the hedgehog with taxa such as Chiroptera, Perissodactyla, Cetartiodactyla, and Carnivora (Waddell, unpubl.). Earlier nuclear sequences, such as globins, sometimes placed shrews, hedgehogs, or moles in association with taxa such as carnivores and pangolin (e.g., Miyamoto and Goodman, 1986). Early morphologists (e.g., Gregory, 1910, and references therein, including Huxley) considered insectivores as primitive, sometimes placing them near primates and perhaps near rodents and lagomorphs (e.g., Fig. 31 in Gregory, 1910). Some recent morphologists have associated insectivores with carnivores or aardvarks (e.g., Novacek, 1993).

5. The pairing of primates and rabbits.—Perhaps another point of divergence between the mtDNA and the nuclear data, especially in light of the analysis by Graur et al. (1996; but see Halanych, 1998, for a severe critique of this work). So far, the published mtDNA trees have tended to put the rabbit closer to the "crown taxa" of Perissodactyla, Cetartiodactyla, and Carnivora, than to Primates (e.g., Arnason et al., 1997). The traditional alternative to this group is Glires. Before the Glires question can be answered categorically, one must be assured that rodent monophyly is consistent with these data (e.g., D'Erchia et al., 1996; Sullivan and Swofford, 1997), which is what we test here.

6. A group composed of primates, rabbits, and rodents.— Suggested by early nuclear data (Miyamoto and Goodman, 1986). Something like this group, plus insectivores, appears occasionally in the analysis of nuclear genes such as vWF and IRBP (Waddell, unpubl.), so we must check the mtDNA data for any sign of it. Its uncertain appearance could be due to a rooting problem with either the nuclear genes or the mtDNA.

7. The position of Proboscidea (elephants) on the tree.— The only previous mtDNA protein sequence for paenungulates (cytochrome *b*) suggested a position for elephant outside of Cetungulata (Irwin and Wilson, 1993) and perhaps a near-basal position among placentals (e.g., Adachi and Hasegawa, 1996b, using NJ, but not ProtML, discussed later). A similar position is suggested for Paenungulata by the αA-crystallin data analysis (e.g., de Jong et al., 1993). The mtDNA 12S-rRNA tree is also consistent with a deep placement (Waddell, unpubl.). Locating elephant would be very interesting, because this would in turn locate the Paenungulate group of three orders, and possibly also other endemic African orders, including elephant shrews, aardvarks, golden moles, and possibly tenrecs (Springer et al., 1997; Stanhope et al., 1998).

MATERIALS AND METHODS

Sequence Data

The data are all of the published mtDNA sequences for vertebrates as of April, 1998, complete for all protein genes. Following a convention suggested by Waddell and Hasegawa (unpubl.), we directly cite papers with sequences less than 2 years old; otherwise, we indicate where the data may be obtained (GenBank numbers or ftp sites). The outgroups are shark *Mustelus manazo* GenBank Accession Number #AB015962 (Cao et al., 1998), coelacanth *Latimeria chalumnae* #Y12025 (Zardoya and Meyer, 1997), cod *Gadus morhua* #X99772 (Johansen and Bakke, 1996), trout *Oncorhynchus mykiss* #L29771, loach *Crossostoma lacustre* #M91245, carp *Cyprinus carpio* #X61010, bichir *Polypterus ornatipinnis* #U62532, lungfish *Protopterus dolloi* #Y12025 (Zardoya and Meyer, 1996), frog *Xenopus laevis* #Y12025, lamprey *Petromyzon marinus* #U11880 (Lee and Kocher, 1995),

alligator *Alligator mississippiensis* #Y13113 (Janke and Arnason, 1997), ostrich *Struthio camelus* #Y12025 (Harlid et al., 1997), and chicken *Gallus gallus* #X52392.

The mammals are as follows: Monotremata—platypus *Ornithorhynchus anatinus* #X83427, Marsupialia—opossum *Didelphis virginiana* #Z29573, wallaroo *Macropus robustus* #Y10524 (Janke et al., 1997), Placentalia, Insectivora—hedgehog *Erinaceus europaeus* #X88898 (Krettek et al., 1995), Rodentia—mouse *Mus musculus* #J01420, rat *Rattus norvegicus* #X14848, guinea pig *Cavia porcellus* (D'Erchia et al., 1996), Proboscidea—African elephant *Loxodonta africana* (Hauf, unpubl.), Lagomorpha—rabbit *Oryctolagus cuniculus* (D'Erchia et al., 1996), Xenarthra—armadillo *Dasypus novemcinctus* #Y11832 (Amason et al., 1997), Cetartiodactyla (Cetacea + Artiodactyla)—blue whale *Balaenoptera musculus* #X72204, fin whale *Balaenoptera physalus* #X61145, cow *Bos taurus* #J01394, Perissodactyla—Indian rhino *Rhinoceros unicornis* #X97336 (Xu et al., 1996), white rhino *Ceratotherium simum* #Y07726, donkey *Equus asinus* #X97337, horse *Equus caballus* #X79547, Carnivora—cat *Felis catus* #U20753 (Lopez et al., 1996), gray seal *Halichoerus grypus* #X72004, harbor seal *Phoca vitulina* #X63726, Primates—gibbon *Hylobates lar* #X99256 (Arnason et al., 1996), Sumatran orang *Pongo pygmaeus abelii* #X97707, Borneo orang *P. p. pygmaeus* #D38115, gorilla *Gorilla gorilla* #D38114, pygmy chimp *Homo paniscus* #D38116, chimp *H. troglodytes* #D38113, and human *H. sapiens* #D38112. Also included as outgroups are the four mtDNA protein regions for birds (rhea, duck, falcon, passerine) from Mindell et al. (1997), kindly provided by these authors and marked MB. LogDet analyses of these sequences (plus accession numbers), can be found in Waddell et al. (1999).

The DNA sequences are converted to AA sequences by using the vertebrate mtDNA code. The sequences were carefully aligned by eye, and any regions of ambiguity for amniotes, frogs, coelacanths, and ray-finned fish, were excluded. This data set, called SS-BAA, has 3362 sites and clearly shows the regions we chose to use. The gene ND6 is excluded; it is the only gene coded on the light strand and consequently has quite different evolutionary properties from those of the other 12 protein genes, making it an inappropriate mix with model-based methods. Additionally, the tRNA genes of the mtDNA were aligned by eye, with their secondary structure used as a reference, for those taxa for which all tRNAs were sequenced (the alignment SSBtRNA)—that is, all the species in Figure 1, except the rabbit and guinea pig, (for which tRNA sequences were unavailable). All regions of ambiguous alignment and anti-codons were excluded (as indicated in the alignment), leaving 1240 sites. Because of the distinct base composition bias between L and H strands, only L-strand sequences were used, irrespective of the coding strand for each tRNA (Kumazawa and Nishida, 1993). These data sets are available from the official home page of this society, www.utexas.edu/ftp/depts/systbiol.

Data Preparation

Following the general guidelines of Waddell (1995:ch. 3), after making the alignments and excluding areas with ambiguous homology, we prepared the data in consideration of four specific factors:

1. Strong evidence of nonstationarity in base composition.
2. Identification and removal of invariant sites.
3. Choice as to which taxa to include.
4. Exclusion of the more rapidly changing sites.

Often, these four steps in editing data can have a more profound effect on results than the type of reconstruction algorithms used.

Programs

To estimate Poisson and LogDet distances (Lockhart et al., 1994; Swofford et al., 1996), we used an unpublished program (AALogDet.exe) by P.J.W. and S. Day. A standard routine calculated the determinant of the normalized 20×20 **F** matrices for amino acids (LU decomposition, Press et al., 1995:43 which runs very quickly; for
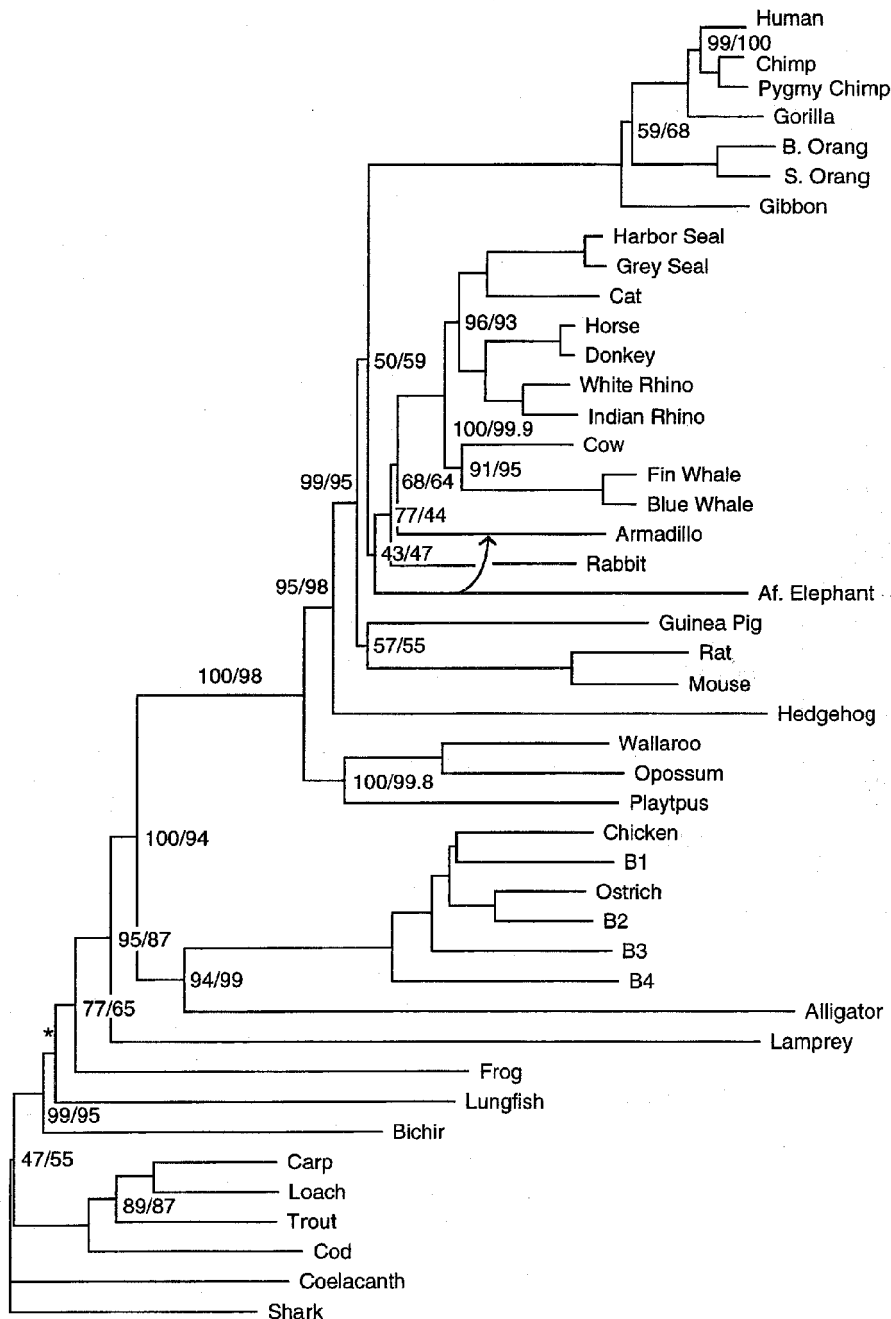
FIGURE 1.   Amino acid LogDet NJ tree. The pairs of numbers are the bootstrap support with all sites in-cluded/invariant sites removed. Edges without numbers occurred in all 1000 replicates. The asterisk indicates that in the bootstrapped trees bichir is placed closer to the tetrapods with support of 56% (52% with invariant sites removed). The arrow indicates the most common rearrangement among mammals, as seen with nearly all other methods. Included are all published vertebrate mtDNA sequences (as of May 1998, plus the elephant). All tree figures are to scale. Here 1cm is ≈ 0.03 weighted substitutions per site.

example, 1000 bootstraps of this large data set were performed in less than 10 minutes on a 300 MHz Pentium II). If either of two sequences being compared is missing the ith amino acid, then the determinant is zero and its log is undefined. This happens especially when constant sites are removed, and all the more so when bootstrapping. To avoid this, we follow Waddell (1995:103) and Swofford et al. (1996:460) by putting the value $1/2$ into entry F_{ii}, before the F matrix is normalized to sum to 1.

One program (Capture.exe) was written to make estimates of the number of invariant sites, using capture–recapture methods, while a second program (AAfreq.exe) was used to perform tests of base composition. These were developed by P.J.W. and H. Mine.

Trees were built and data bootstrapped with Neighbor.exe in PHYLIP 3.5 (Felsenstein, 1993), and PAUP*4.0 d 61-64 (Swofford, 1998) for FM and parsimony trees. For all protein ML analyses, ProtML 2.3 was used (Adachi and Hasegawa, 1996b), along with the rate matrix of Adachi and Hasegawa (1996a); all nucleotide ML analyses used PAUP*4.0.

RESULTS

Evaluating Stationarity

To measure AA composition fluctuations, we performed a G^2 (likelihood ratio) test on the base composition of just the sites that have changed between a pair of species (that is, we take the divergence matrix, F, remove the diagonal elements, and then compare the sums of the columns with that of the rows). The test has degrees of freedom nominally equal to 19 and so, assuming a chi-square distribution of G^2, values above 36.19 are significant at the 99% level.

As discussed in Waddell (1995:ch. 3), only the sites that vary matter, because both the invariant and the by chance unvaried sites are not directly relevant to the question of whether the process of evolution is stationary. The results are shown in Table 1 for selected taxa (generally one species per mammalian order). Nearly all values are significant, leaving little doubt that the process of evolution is nonstationary. Interestingly, the group of cat, cow, armadillo, and rabbit seems reasonably homogeneous. These turn out to be the same placental taxa that do not fail a clock-like rate test (Waddell et al., 1999).

A test similar to this is the test of the symmetry of the divergence matrix, F (Waddell and Steel, 1997) (results not shown), which gives very similar answers to those reported here, for the two tests generally detect the same shifts in the AA substitution process. A standard test of base composition, such as that in PAUP*4.0 (Swofford, 1998), either with or without columns of invariant sites removed, tends to be very conservative (i.e., does not easily detect nonstationarity).

The clear evidence for the violation of stationarity is surprising in its extent; whereas previous authors had noted it for the hedgehog, this more sensitive test shows that it applies to nearly all the studied species (including relatively close species pairs such as whale and cow). The largest differences tend to be to the hedgehog and to the outgroup (the latter understandable). Primates are also a bit unusual, and perhaps also whales. The number of large differences between platypus, marsupials, hedgehog, and murid rodents is a concern. In theory, nonstationarity violates the assumption of all methods, except the LogDet, whose assumptions are violated by the nonuniformity of site rates (see below). However, although the effects are significant, this may also be at least partly due to long sequences evolving at a moderately high rate, so that differences can be easily detected. Moreover, the considerable similarity of the bootstrap results for NJ, with and without the AA LogDet distance, suggests that NJ at least is fairly unaffected by the differences detected here.

Removing Invariant Sites

It is important to remove invariant sites when they are present, to avoid the risk of inconsistency (Waddell, 1995:369–385; Lockhart et al., 1996), but it is also im-

TABLE 1. G^2 pair-wise test of AA composition stationarity ignoring constant sites. Overall sum of values = 3912.9; total degrees of freedom = 1482, therefore highly significant (correlation between pairs not considered).

Species	2	3	4	5	6	7	8	9	10	11	12	13
1 Human	59.8*	60.1*	23.4	49.1*	32.0	37.9*	53.8*	65.8*	110.8*	74.0*	45.4*	51.3*
2 Cat	–	23.2	41.6*	12.8	22.1	29.5	15.1	43.6*	69.9*	39.9*	33.4	99.8*
3 Cow	–	–	46.0*	18.5	23.9	21.9	19.4	30.4	62.6*	41.7*	40.6*	113.3*
4 Whale	–	–	–	35.4	25.5	15.0	35.5	41.0*	99.7*	62.3*	28.0	61.1*
5 Armadillo	–	–	–	–	20.2	35.2	25.9	30.8	62.6*	29.2	42.8*	95.8*
6 Elephant	–	–	–	–	–	33.2	23.0	34.0	60.4*	39.0*	36.4*	90.7*
7 Rabbit	–	–	–	–	–	–	28.4	32.0	78.9*	49.1*	24.7	72.8*
8 Guinea pig	–	–	–	–	–	–	–	29.9	43.3*	32.1	32.7	93.0*
9 Mouse	–	–	–	–	–	–	–	–	38.5*	23.8	51.9*	105.9*
10 Hedgehog	–	–	–	–	–	–	–	–	–	35.9	99.9*	186.1*
11 Opposum	–	–	–	–	–	–	–	–	–	–	74.1*	133.7*
12 Platypus	–	–	–	–	–	–	–	–	–	–	–	71.1*
13 Chicken	–	–	–	–	–	–	–	–	–	–	–	–

*$P < 0.01$.

portant not to remove too many constant sites (or overadjust for unequal site rates) and so proportionately overestimate the larger distances. Doing so can also cause inconsistency (the "anti-Felsenstein zone" problem, or "long edges will repel"; Waddell,1995:385–398). Thus, we need to estimate how many sites may actually be invariant.

Two types of capture–recapture estimates of invariant (unable to change) sites were made, based on codon positions (Sidow et al., 1992) and based on character states in different groups (Waddell, 1995:130). The unbiased form of these predictors (Seber, 1982; Waddell, 1995:133) made very little difference (less than 0.1%) with this number of sites but is calculated here along with its standard error. For the data set used in Figure 1, the number of invariant sites estimated by the method of Sidow et al. (1993) was 943.6 ± 12.3 (1 SE).

The capture–recapture method of Waddell (1995:130) uses two distinct groups, clearly separated by an internal edge on the unrooted tree. Thus, one group is monophyletic while the other may be monophyletic or paraphyletic, although the method seems robust to violations of this condition (i.e., mixing groups up). Here the two groups were placentals versus nonplacentals (it is statistically preferable to make both groups similar in size and sum of edge lengths). Next, one counts how many sites

show any change in the first group (here 1672), how many show change in the second group (2040), and how many sites show change in both groups (1501). The estimate of variable sites is then [(1672 + 1)(2040 + 1)/(1501 + 1)] − 1, so the number of invariant sites is 3362 − 2272.4 = 1089.6. (SE 9.6). That is, 1089.6 out of 1151, or 94.7%, of the constant sites are estimated to be invariant.

Clearly the two different capture–recapture methods are giving distinct results. Because of the possibility that the first method is sensitive to neutral C to T transitions in the first position of leucine (especially common in mtDNA), we will use the second estimate. The data set of amino acids minus these invariant sites (where the invariant sites are removed in proportion to the constant sites: Waddell, 1995:120; Waddell and Steel,1997) is called SBAA1090.

The Tree from Proteins

We now digress to consider the overall tree before considering further data editing and taxon selection. Shown in Figure 1 is the NJ Amino Acid (AA) LogDet or Paralinear distance (Barry and Hartigan, 1987; Lake, 1994; Lockhart et al., 1994; Waddell et al., in press) tree for all published complete vertebrate coding mtDNA sequences. The tree is definitely unusual in the point at which the lamprey joins. Recently, Rasmussen et al. (1998), without the shark sequence, suggested that the lamprey is rooting correctly

and presented this as a bold new hypothesis. Thus, the present analysis is suggesting that the lamprey is a highly degenerate amphibian—although we still prefer the explanation in Cao, Waddell et al. (1998) afforded by the alternative rooting of shark.

Other taxa that pose problems in their mode of evolution, their alignment, or placement are bichir (which traditionally should be sister to ray-fins) and lungfish (which rates high evolutionarily). The position of bichir really needs further consideration with other data (perhaps with additional fish such as sturgeon added; though might this discrepancy be an error on the part of morphological interpretation?). Excluding these taxa, we get a fairly orthodox tree (Cao et al., 1998b; although shark still jumps between the coelacanth and ray-finned fish edge). The effect of outgroups on mammalian hypotheses are studied in detail below.

The exclusion of invariant sites has brought no change to the NJ AA LogDet tree, other than changing the bootstrap values (Fig. 1). In some places, it appears to improve the support for well-established hypotheses, such as the great apes (including humans), Cetartiodactyla, Placentalia, and Archosauria. At the same time, it reduces support for other near-certainties, such as Amniota, and Mammalia. What is perhaps most interesting is that exclusion of invariant sites makes very little difference to some recently advanced controversial hypotheses, especially Marsupionta, the Carnivora plus Perissodactyl clade, and to these taxa plus Cetartiodactyla (whales and cow).

This last grouping is sometimes called the ferungulate clade, but Ferungulata (Simpson, 1945) is defined as the group with living members Proboscidea, Sirenea, Hyracoidea, Artiodactyla, Perissodactyla, and Carnivora. From the analyses below, this group is almost certainly polyphyletic. However, just mentioning Ferungulata or even Ungulata now creates much confusion, as many morphologists still believe either or both to be a possible natural group. Thus, in the *Discussion*, we name

and define a new superordinal group, the Fereuungulata, which consists of the living orders Cetartiodactyla, Perissodactyla, and Carnivora (plus Pholidota, not sampled here).

Interestingly, if the tree selection criterion is switched to FM least squares, with a constraint of all edges positive, the tree changes. Shark joins to the ray-finned fish edge, coelacanth is in a more usual position, and the guinea pig breaks off from the rodents and joins just below them (as if being attracted to primates), while the whale breaks loose of the cow and lies sister to the other fereuungulates (more on these points later). With invariant sites removed, the FM+ tree is the same, except that the whales now rejoin the cow. This last feature is predicted by looking at the Split Decomposition diagram for these species (Bandelt and Dress, 1992; Swofford et al., 1996), which suggests an attraction of whales to taxa deeper in the tree. This attraction does not entirely disappear with even the invariant sites-LogDet treatment (similar to the example in Waddell, 1995:170). It is an indication of how the accelerated rate of evolution for whales (perhaps twice the average for cow; see below) might be having a detectable effect, and how such effects must also be surely affecting other rapidly evolving mammalian sequences.

The ProtML tree (Molphy 2.3b; Adachi and Hasegawa, 1996b) is overall similar to Figure 1, including the edge lengths. [Note that for all the differences in trees explained herein, the tree being described can be generated by removing the underlined taxa and then readding them at the newly indicated position (for proteins the reference tree is Figure 1, for tRNA it is Figure 2).] Points of difference in the outgroups are (a) a strong preference for trout to join with cod (99% support by RELL local bootstrapping), (b) a slight preference for shark to go with ray-finned fish (54%), and (c) bichir joins with lungfish. Within the placental mammals more changes are obvious. The guinea pig now lies sister to primates, with rabbit sister to these two. The elephant has joined with the armadillo, which remains

sister to the fereuungulates. The local rearrangement RELL bootstrap support for this last clade is 89%. The RELL support of clades of special interest are shown in Table 1.

There were two maximum-parsimony trees (from a TBR search with taxa added by the closest option, PAUP4.0, Swofford, 1998), both similar to the tree in Figure 1 (and differing amongst themselves only on interbird relationships). The other differences from the tree in Figure 1 were these: parsimony, like ML, groups the trout with cod; the lungfish has moved to be a sister taxon to the lamprey; the rodents are no longer monophyletic, with the arrangement (murids, (guinea-pig, (rabbit, ((fereuungulates), (armadillo, (elephant, (primates))))))).

The pattern is unlikely to be random; it seems to reflect (a) a tendency for the elephant to be attracted to primates (and perhaps the armadillo is being dragged along with it), (b) a tendency for guinea pig to be attracted in the direction of primates, and (c) a weak attraction amongst the members of Glires.

Here, the parsimony trees weighted as described in Waddell (unpubl.), are the same as the parsimony tree. (The weighting mentioned involves taking the absolute value of component-wise logarithms of the predicted $^{-1/2}\mathbf{F}^{-1/2}$ matrix, assessed down the longest edges in the tree under a realistic evolutionary model; see Waddell and Steel, 1997, for terminology.) At other times, they are closer to the ProtML tree, at least partly because they both use the vertebrate mtDNA substitution rate matrix presented in Adachi and Hasegawa (1996a). Note the bootstrap proportions between parsimony methods and ProtML vary largely because ProtML uses a localized bootstrapping of previously estimated site likelihoods (RELL). This often gives elevated support for a group, relative to bootstrapping, when feasible alternative arrangements are not nearest neighbor interchanges.

Interesting points of comparison in Figure 1 are how the character-based methods favor certain hypotheses that the distance-based trees do not. The distance-based methods place cod outside the other ray-finned fish (favored by morphology; e.g., Stiassny et al., 1996). Rodent monophyly is supported by strong prior morphological evidence; the results here support this hypothesis. The rodent group is recovered most frequently by NJ, irrespective of the use of the LogDet.

The grouping of armadillo and elephant by the character-based methods is speculative but agrees with predictions made earlier in Waddell (unpubl.) on the basis of multiple nuclear genes. It occurs often with ProtML and parsimony, but other solutions such as elephant joining with primates, or sometimes even jumping out with hedgehog, can be close on the AA data. In fact, the landscape for tree searching is anything but easy, with similarly good solutions that may be topologically quite distinct. A curious apparent error when using observed amino acid and LogDet distances with FM (but not NJ) is putting whales outside the cluster of cow, perissodactyl, and carnivore. This bias and observed distances shows up also when using split decomposition.

Effect of Outgroup Selection

In this section we look at the effect that the outgroups are having upon hypotheses, especially those near the root of the mammalian tree. Some of the criticisms leveled against the Marsupionta hypothesis have noted the sparseness of outgroups, something well compensated for here. We define four sets of outgroups and a fifth data set minus hedgehog: (1) All taxa used as outgroups, (2) Removal of lamprey, lungfish and bichir, (3) Removal of all fish (nontetrapods), (4) Removal of all fish and new birds from Mindell et al. (1997), and (5) Set (4) minus hedgehog.

As Table 2 shows, support for Marsupionta does change as outgroups become more scarce. However, the change is in the opposite direction to that which critics would suggest, with more extensive outgroups leading to more support with both parsimony and ML methods. It is also interesting to note that Marsupionta *gains* support with the exclusion of hedgehog. Note also that Marsupionta support is noticeably higher with NJ and this is not attributable
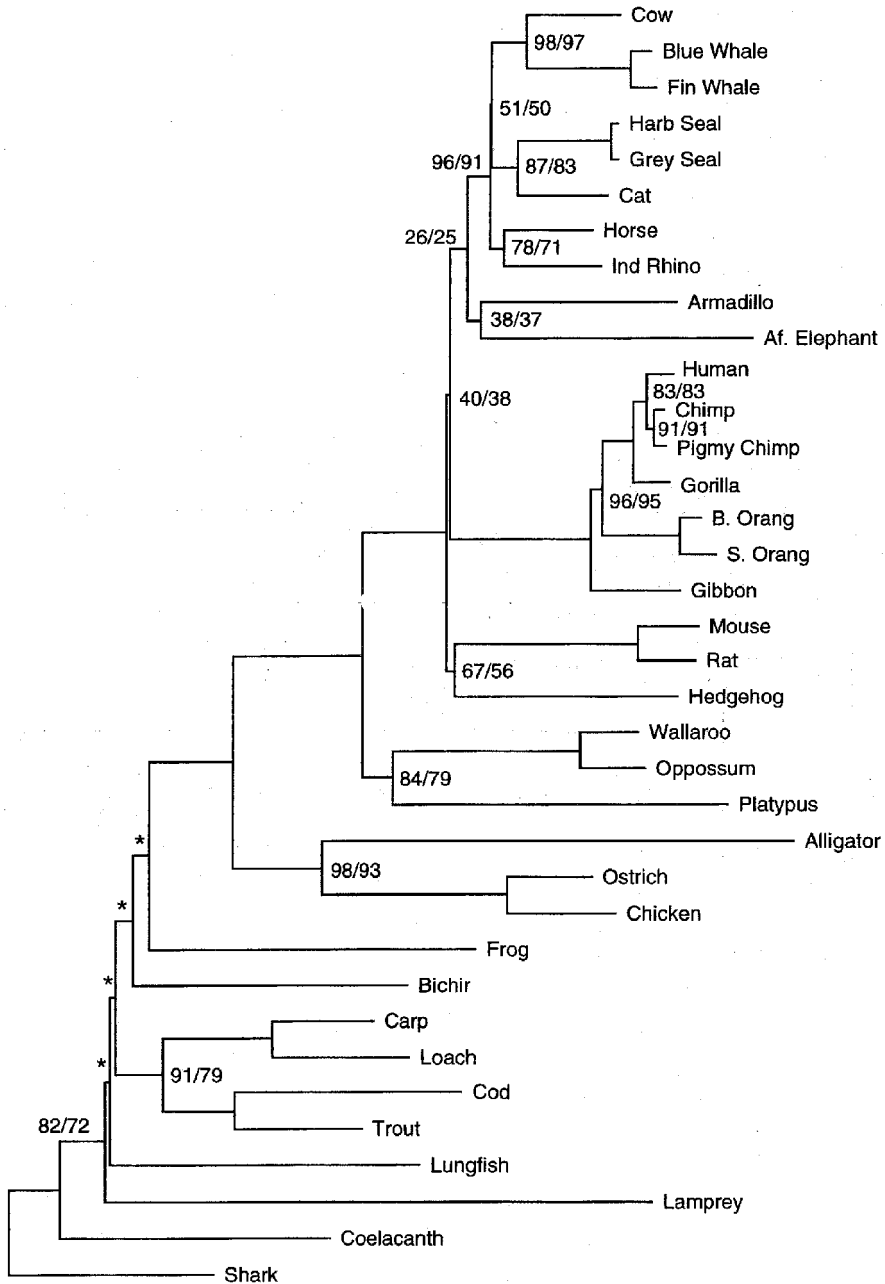
FIGURE 2.   The tRNA invariant sites-LogDet NJ tree. Bootstrap values are for 1000 replicates (no value indicates 100% support). The first number is support when all sites are treated as equally variable; the second is with the estimated proportion (HKY model estimate at 19.5%) of invariant sites removed according to their overall base frequency in these sequences (Waddell, 1995:118; Waddell and Steel, 1997). The asterisks indicate edges that do not appear in the bootstrap consensus tree, where the lamprey moves and joins with the frog, and the lungfish moves and joins with bichir. Here 1cm is ≈ 0.046 weighted substitutions per site.

to the LogDet, as using observed AA distances give similar numbers. Here, Marsupionta does not appear artifactual, but NJ may be giving a biased view of its support.

Generally, the choice of outgroups makes little difference to the support for hedgehog as first branching amongst the placentals, which is again elevated with NJ relative to the character-based methods. Rodent monophyly is certainly not challenged by these data, and it seems a sign of the instability and superficiality of some analyses (e.g., D'Ericha et al., 1996) that this is ever argued seriously. Rodent monophyly support peaks with the LogDet NJ method at about 59% and is about one-third for the character-based methods. The carnivore/perissodactyl hypothesis is generally well supported, but the exclusion of hedgehog does reduce it somewhat. Again, NJ is behaving quite unlike the character-based methods, maintaining high support and being quite insensitive to the other taxa in the analysis (a possible hallmark of the method we have noticed).

The hypothesis of armadillo and elephant together receives mixed support. ProtML gives it high support, whereas the parsimony methods tend to shift it around (more than one interchange apart), locating the elephant near the primates and occasionally near the hedgehog. NJ does not favor this hypothesis, but it clearly does not reject it either. Likewise, elephant, armadillo, and fereuungulate are seen as a possible clade, but this too receives mixed support. The ML solution of rabbit, primate, and guinea pig is seen as quite unlikely by the other methods. This shows that even ML is sensitive to something unusual in this data.

Analyses of Stripped Data

A real problem among the placentals is the lack of resolution now apparent in the "middle branching" group, that is, including rodent, primate, rabbit, elephant, armadillo, with sometimes even the hedgehog grouping with other taxa. There is now much contradiction as to what goes where; every ordinally distinct new sequence added in this part of the tree seems to bring added confusion with it. Here, we consider if this can

be resolved by removing the more rapidly evolving sites. We do this by stripping out of the data all sites that show any variability in specified groups.

A prime candidate group for stripping is primates. These taxa show a highly accelerated rate, faster perhaps than even hedgehog. Removing all sites variable within primates will shorten this edge and perhaps de-emphasize sites prone to homoplasy. This data set (SBAA-pr) contains 2718 sites and all species included except the three strange fish: lamprey, lungfish, and bichir. Using AA capture–recapture (for placentals vs. nonplacentals, set 2) yields 1131 (SE 13.3) sites estimated as invariant; removing these (SBAA-pr.-in) leaves 1587 sites. Results (Table 3) show that support then generally dropped for Marsupionta, hedgehog first, carnivores plus perissodactyls, elephant/armadillo/fereuungulates, and the rabbit/primate/cavie group and remained about the same for the other two groups. Note, methods still often disagree markedly with each other over support.

Of the methods, NJ seemed most insensitive to this method (Table 3). The NJ tree changed only when based on observed distances, with the caviomorph joined to primates in this tree (relative to Figure 1). Even so, it hardly changed the bootstrap support of the hypotheses listed in Table 2. The parsimony tree now makes the rodents monophyletic and sister to primates. Rabbit moves next to armadillo, replacing elephant, which moves as sister to all placentals except the hedgehog. On the WP tree, the rodents become sister to primates, with elephant + armadillo sister to them again. These are all sister to the fereuungulates, with the rabbit and then the hedgehog successively deeper. For parsimony, the support for Marsupionta and for Carnivora plus Perissodactyla drops considerably, while the association of elephants + armadillos and Fereuungulata remains fairly stable, as does rodent monophyly.

For likelihood, ProtML turns up some interesting new optimal trees. With primate stripped sites, the rodents turn up as a monophyletic sister taxon to primates, with

TABLE 2. AA-based bootstrap support for mammalian phylogenetic hypotheses with respect to the set (1–5) of outgroups used. ProtML used RELL with 10,000 replicates; all other bootstraps with 1000 replicates. With AA LogDet NJ and AA Dobs NJ bootstrap, there is approximately 5–10% chance the hedgehog will shift deeper so that placentals are a nonmonophyletic group.

Outgroup set →	AA Pars					WPars					Prot.ML						AA Dobs/NJ					LogDet NJ					
	1	2	3	4	5	1	2	3	4	5	1	2	2–[a]	3	4	5	1	2	3	4	5	1	2	2–[a]	3	4	5
Marsupionta	88	85	77	58	62	95	91	71	69	78	89	81	89	67	49	72	100	100	100	100	100	100	100	99	100	100	100
Hedgehog first	80	73	76	80	–	80	77	77	70	–	92	88	87	92	91	–	100	100	100	100	–	99	99	95	100	99	–
Rodent mono	32	32	30	27	30	31	30	27	26	29	–	–	–	–	–	–	50	52	53	54	50	57	59	55	58	59	44
Carn + periss	85	93	86	84	77	93	94	94	92	87	94	94	93	93	92	89	98	99	99	98	99	96	96	93	96	97	95
El/arm./ferun	23	20	15	12	22	27	25	24	24	28	81	78	82	75	79	85	12	12	12	16	13	14	14	32	14	15	16
El/arm	17	17	22	25	26	35	36	34	33	42	89	89	88	89	89	88	4	4	3	3	2	5	5	8	5	5	
Rabbit/prim/cavie	2	3	3	4	3	10	15	9	36	13	56	56	62	59	54	58	0.3	0.3	0.4	0.4	0.3	0.4	0.4	9	0.4	0.4	0.4

[a]Invariant sites removed.

rabbit sister to these again. In turn, this group is sister to the fereuungulates, with the elephant + armadillo clade sister to them again, and lastly the hedgehog outside all these. Removal of the invariant sites as well saw more changes: Elephant moved deeper as sister to primates, and not only did rodents form a monophyletic group, they also became sister to rabbit, at the position of rabbit, i.e., in Figure 1, just outside the armadillo. Thus Glires "rides again" in the mtDNA data. Interestingly, this tree is also a local NNI optimum (Swofford et al., 1996) for the data sets with all sites in, except there Glires moves outside all placental taxa except hedgehog.

A data set was prepared that stripped out sites that changed between closely related pairs of mammals (the two chimps, two orangutans, two seals, two equids, and two whales). As a result, the ProtML tree changes slightly: The murids break out of Glires and come to branch separately directly after hedgehog, while the armadillo + elephant group moves sister to the fereuungulates. Interestingly, there seems to be a second local NNI optimum on these data, which is exactly the tree from the first part of the paragraph above. Using this data set did not change the NJ trees (based on observed and AA LogDet distances) and barely changed bootstrap support.

Interestingly, on all these stripped data sets, the support for Marsupionta is down to 70–76% RELL bootstrap support on the optimal trees.

A data set was also prepared that stripped variable sites in murid rodents (rat and mouse). A third did the same for both rodents and primates. We did not see any clear support emerging for either new hypotheses or established hypotheses. However, majority NJ bootstrapping trees (e.g., with observed distances) sometimes favored trees where rodents were closer to the fereuungulates than to the primate lineage, showing that even this conclusion may not be as solid as hoped previously.

A disturbing factor is that, in theory, if the data are distributed according to a Γ distribution, and site rates are fixed relative to one another (as discussed in Waddell et al., 1997), we should easily be able to trap just those sites with the highest rates of change, which are likely to be causing a disproportionate amount of homoplasy. Such does not seem to be the case here, at least to the extent of getting better-resolved trees (Waddell, 1995:216). We have, however, caught glimpses of hypotheses thought by many to have been killed off by the mtDNA data—not just rodent monophyly, but also Glires (which received up to 84% local RELL bootstrap with one stripped data set!).

Why Caviomorpha with Primates?

A peculiar feature of the ML trees (and the best or close to best parsimony trees) is Caviomorpha (guinea pig) with Primates. Earlier we had studied a tendency of the ND1 gene to strongly group rodents with primates (Cao, Janke et al., 1998). To test this, we removed ND1 from the data. This changes the ML tree to one where the caviomorph branches immediately after murids, as does the rabbit (right after the caviomorph), suggesting an arrangement not far from Glires (elephant also then moves adjacent to primates, as though this attraction is now free to act). Almost exactly the same tree is now also recovered by parsimony and weighted parsimony, except that armadillo also moves to become a sister to the primate + elephant group.

Further examination of just the tree of more densely sampled ND1 sequences indicates that the caviomorph is slightly more strongly attracted to primates than to the murids. It is interesting to speculate that this feature may be a large part of the cause of caviomorphs sometimes dropping down the tree towards primates, which has misled some biologists into believing the data support rodent paraphyly (e.g., D'Erchia et al., 1996). Interestingly, the NJ AA trees do not show this feature. This may be a useful illustrative example where ML is misled more than parsimony, and both of these in turn are misled more than NJ. It seems ML and parsimony can go wrong when rare convergent patterns coincidentally achieve a high

TABLE 3. Bootstrap support from mtDNA proteins after variable characters within primates were stripped out (outgroup set = 2).

	AA Pars	WPars	ProtML	Dobs/NJ	LD/NJ	*ProtML	*LD/NJ[a]
Marsupionta	70	74	77	99	100	84	99
Hedgehog first	50	96	100	100	97	99	90
Rodent mono	53	36	75	43	60	90	63
Carnivores + perissodactyls	80	75	83	97	93	78	81
Elephant + armadillo + fereuungulates	6	24	–	2	1	–	5
Elephant + armadillo	17	42	89	0	2	–	11
Rabbit + primates + caviomorphs	2	2	–	0	0	–	1

[a] Invariant sites removed.

leverage (in the statistical sense, like outliers on a regression).

Actually, to rectify this problem, one need not remove the whole ND1 gene (and thus ignore what are probably sites as informative as any in the mtDNA) but perhaps remove just the 8 sites that Cao et al. (1998a) identify as the likely root of the problem. The danger of course is that rodents and primates may really be closely related, as we see next.

Comparison with tRNA Data

These data were analyzed with equivalent methods to those used for the proteins. Unfortunately, the data sets do not contain exactly the same taxa, because the tRNAs for rabbit and guinea pig were not available. This means that the hypothesis of rodent monophyly cannot be tested, whereas the lack of four sets of bird tRNAs means that there is no data set (outgroup set 4) that specifically excludes them. Table 4 shows the results for some specific hypotheses. Support for Marsupionta is more uneven and lower than with proteins. Support for carnivores plus perissodactyls is also considerably lower, and that for hedgehog first is overall very low (about 10% on average). That for elephant + armadillo + fereuungulates and elephant + armadillo is up slightly, again mostly with ML.

The NJ invariant-sites/LogDet tree, shown in Figure 2, has many interesting points. The grouping of armadillo and elephant appears again, and there is a rearrangement within the fereuungulates that disagrees with the Perissodactyl/Carnivora clade. A major insight is evidence that the root can shift. The root migrates to the edge in Figure 1, that defines a clade of all placental taxa except hedgehog and rodents. Note also the strong support for cod with trout, making us wonder whether the traditional view of trout closest to loach and carp is really correct.

This last point is interesting because the same type of method (LogDet, then NJ) gives nearly totally contradictory results for this group for the same molecule, depending on whether tRNA or proteins are used. Thus something must be wrong, and LogDet is certainly misled in at least one instance. The important point is, if such low rate groups can be wrong and contradictory, imagine the possibility for such factors amongst the much longer edges that lead to mammalian taxa.

The parsimony trees are identical to the NJ trees, except that the former still favor the Perissodactyl + Carnivore clade (see Table 4) but give perissodactyls paraphyletic with respect to carnivores.

The ML trees are even more interesting (Fig. 3). They (HKY and GTR model, with or without invariant sites estimated) support the Theria hypothesis, with platypus outside of all other mammals, and they also put the primates sister to rodents, which are sister again to the hedgehog.

One of the most interesting findings once again concerns the branch lengths. Again we

have signs of accelerated rates in hedgehogs, rodents, primates, elephants, and whales relative to cows, cats, horses, and rhinos. This time, however, many of them are clustering together. The main point is that the root has shifted. Further, the internal branch lengths on the trees (especially invariant sites-LogDet NJ) are now rather short, also as in Figure 1, compared with nearly all of the trees published for these data. As to which edge lengths are correct, an example in Waddell (1995:343) shows that ML models that ignore an excess of parallel and convergent changes tend to add many of these changes to the length of the internal edges (as does parsimony).

If the tRNA tree is correct, it may be very difficult to root the placental subtree. It seems possible still that the armadillo/elephant clade may be sister to all other placental taxa (the edges separating them from the root in the tRNA trees are quite short). Indeed, such a tree would agree closely with the rooted αA-crystallin tree (e.g., de Jong et al., 1993, which goes further and suggests Xenarthra alone at the root). So, we may have to be cautious and conclude here that rooting the placental subtree on any of the branches earlier than fereuungulates retains some possibility.

Checking the Position of Hedgehog and Elephant

These two taxa are amongst the most enigmatic on the tree. The position of hedgehog remains suspect, given its very long edge and its showing the largest base composition and AA frequency shifts, whereas elephant has a long edge and is a newly added order that probably represents many endemic African mammals (Springer et al., 1997).

To check the position of these two taxa, we have tracked their fit on the tree at the 25 positions on the invariant sites-LogDet NJ tree (Figure 1) and ProtML tree for all sites (Tables 5 and 6). We do not show the scores for ProtML on the LogDet tree (to conserve space). The taxa follow a trend similar to those on the ML tree. If, for example, the hedgehog is, in fact, associated near or

with the fereuungulate clade (as some nuclear sequences, including IRBP + vWF suggest; Springer et al., 1999; Waddell, Okada, & Hasegawa, 1999), then we would hope to see some clear improvement in the fit (even if not optimal overall) as we locate the sequence in this region. Thus, we are looking for possible local regions of attraction for these sequences.

Our interpretation of these results is as follows. For hedgehog, there is a general decline in fit as the sequence moves towards the group of true ungulates and carnivores (Table 5). Points against this trend occur when hedgehog encounters a very long edge, namely, Primates, Muridae, Caviomorpha, elephant, and deeper still, the common mammal edge (these were also most of the localities from which it was not statistically rejected; it doesn't attract to the fairly long platypus edge, however). The blip for the attraction for the common proto-mammalian edge is a worry for parsimony, as it may indicate that the hedgehog is experiencing a fairly strong tug in this direction (ML does not note it). Thus, there would seem to be good evidence for some specifically long edge attraction. We could not detect any signal to indicate that hedgehog lies next to or within the fereuungulate grouping, as has been suggested by some nuclear sequences (as mentioned earlier), except possibly that with ProtML a position sister to fereuungulates is not heavily penalized compared with other nearby edges.

For the elephant, there is a peak fit either with or adjacent to the armadillo (Table 6). There are clear violations to the trend of decaying support away from this point: attractions to primates and hedgehog in particular. As we see in Figure 1 and Waddell, Cao et al., 1999, these are the most rapidly evolving mammalian sequences. Thus, the decided grouping with armadillo (a slow-rate species) is surprising, and we take it as further evidence for this being the correct location of this sequence. The elephant only reluctantly joins to become part of the fereuungulates as a sister to all others (rejected at the 95% level by ProtML and parsimony).

TABLE 4. tRNA-based bootstrap support for hypotheses within Mammalia with respect to the outgroup set used. ProtML RELL based on 10,000 replicates; all other bootstraps used 1000 replicates. With parsimony, and data set 5, the hypothesis of marsupials plus placentals equals Theria receives 57% support.

Outgroup set	Pars				WPars				HKY ML				Dobs/NJ				LogDet/NJ			
	1	2	3	5	1	2	3	5	1	2	3	5	1	2	3	5	1	2	3	5
Marsupionta	69	69	52	43[a]	69	69	52	43[a]	67	<35[a]	<35[a]	<35[a]	78	83	91	89	78	84	92	90
Hedgehog first	7[a]	7[a]	8[a]	–	7[a]	7[a]	8[a]	–	<18[a]	<20[a]	<20[a]	–	11[a]	12[a]	16[a]	–	10[a]	10[a]	13[a]	–
Carnivores + perissodactyls	25	25	24	16[a]	25	25	24	16[a]	<37[a]	<34[a]	<34[a]	<34[a]	24[a]	24[a]	22[a]	18[a]	12[a]	12[a]	9[a]	8[a]
Elephant + armadillo + fereuungulata	44	38	34	23	44	38	34	23	93	83	83	83	28	26[a]	20[a]	17[a]	26	26	16[a]	12[a]
Elephant + armadillo	58	54	50	33	58	54	50	33	68	69	69	69	35	35[a]	31[a]	19[a]	37	38	33[a]	22[a]

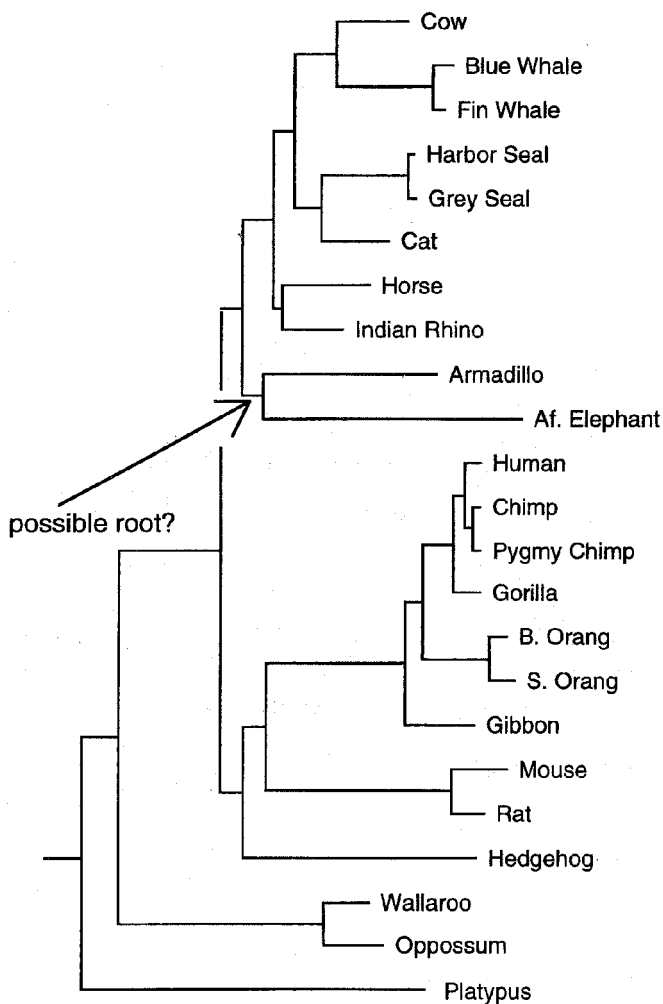[a] Group is not on the compatibility bootstrap consensus tree.

FIGURE 3.   The GTR ML tree of tRNA (with just the mammal subtree shown), with invariant sites accounted for (estimated at 21.7%). The arrow indicates that if the root comes in just a little more, the hypotheses is rather close to Epitheria (Xenarthra as sister to all other placentals). Here 1cm ≈ 0.04 substitutions per site.

The elephant is strongly rejected from joining with the other ungulates. To further test this hypotheses, we made constrained searches for the best trees with Ferungulata and Ungulata (here Perissodactyls, Cetartiodactyla, and Proboscidea). Unweighted parsimony strongly rejected ($P < 0.003$) Ungulata, but not Ferungulata. Weighted parsimony and ProtML have greater discrimination and both rejected Ferungulata and Ungulata, either separately or together, at $P < 0.05$ when tested on the best trees with and without constraints (and also when taking into account invariant sites).

So far as possible positions on the tree are concerned, we note that the Kishino–Hasegawa (KH; Kishino and Hasegawa, 1990) test with parsimony is less discriminatory than with ML; WP falls somewhere in between. The nonparametric winning sites and rank correlation (Templeton) tests for parsimony (also applicable to ML) were less sensitive than the normal approximation-based KH test, as is usually the case with

TABLE 5. Testing the position of hedgehog (the clade shown in the left column is the sister taxa to hedgehog). Numbers are length increase over the shortest tree for parsimony, and worse lnL for likelihood. The group indicated is the putative sister taxon to the hedgehog.

Hedgehog Position in ProtML tree					Hedgehog Position in LogDet tree (figure 1)		
	Pars	WPars	ProtML	ProtML–inv		Pars	WPars
PL	best	best	best	best	PL	29	2027
Mur	12	430	− 18	− 12.2	Rod	45*	2567*
PL – Mur	11	699*	− 23.1	− 15.7	Cav	49	2432
PL – Rod –Pr – Rab.	35*	2025**	− 78.1**	− 60.5*	Mur	50	2755
eleph + arm	32*	2043**	− 88**	− 71.1*	PL – Rod	44*	2844*
eleph	14	466	− 73.6*	− 65.9*	Pr	46	2793
arm	37*	2461*	–101.3*	− 81.9*	PL – Rod – Pr	61*	3817*
rab + Cav + Pr	27*	1782**	− 78.7**	− 63**	eleph	45	2331
rab	32*	1960*	− 90.1**	− 73**	PL – Rod – Pr – eleph	73*	4427*
Pr + Cav	37*	1922*	− 94**	− 76.5**	rab	81*	4633*
Cav	33*	1633	− 97.6**	− 83.2*	fer + arm	74**	4726*
Pr	33*	1724	–117.4**	–100.4**	arm	71**	4269*
Fereuungulates	41*	2219**	− 80.5*	− 61.2*	fer	73**	4789**
Cetart	76**	4203**	–150**	–113.9**	Cetart	108**	6577**
Whale	81**	4711**	–180.3**	–139.8**	Whale	116**	7051**
Cow	76**	4352**	–163.4**	–127**	Cow	109**	6666**
Pe + Ca	88**	4527**	–157.9**	–119.5**	Pe + Ca	113**	6879**
Ca	93**	4864**	–179.7**	–132.3**	Ca	123**	7182**
Pe	105**	5487**	–198.9**	–147.3**	Pe	137**	7914**
Mam	16	930	− 94.9**	− 86.2**	Mam	41	2969
Ionta	36**	2259*	–108.5**	− 91.4**	Ionta	74**	5358**
Platypus	48*	3344*	–169.8**	–147.1**	Platypus	74**	5358*
Marsup	41**	3169**	–162.4**	–139.8**	Marsup	74**	5507**
Opossum	115**	8239**	–388.4**	–334.1**	Opossum	144**	10392**
Wallaroo	119**	8478**	–389.6**	–334.1**	Wallaroo	150**	10685**

PL = Placentalia, Mur = Muridae, Rod = Rodentia, Pr = Primates, rab = rabbit, eleph = elephant, arm = armadillo, Cav = guinea pig, Cetart = Cetartiodactyla, Pe = Perissodactyla, Ca = Carnivora, Mam = Mammalia, Ionta = Marsupionta, Marsup = Marsupiala.

The minus sign (–) in column 1 means the first group minus the following taxon or taxa. Here, the best parsimony score on the ML tree is 12577, for WPars is 70933, and for ML is –75863.5. The best parsimony score on the LogDet tree is 12606, for WP is 711363. Tests are made against either the ProtML or LogDet tree.

$^*P < 0.05, ^{**}P < 0.001$.

nonparametric tests. Our tests have been more discriminatory as to where the hedgehog can go than are those of Sullivan and Swofford (1997), almost certainly because we used the amino acids rather than first and second DNA positions. This holds even with invariant sites removed, although other distributions (e.g., Γ or inverse Gaussian, Waddel et al., 1997) might see other possible positions of hedgehog, such as next to Fereuungulates, as reasonable. Thus, the results are in general agreement, in that the hedgehog could possibly slip down the backbone of the tree in the direction of the Fereuungulates.

Likelihood Ratio Test for Possible Phylogenetic Inconsistency

The nonrecombinant mtDNA molecule allows an interesting test of whether there is evidence for strong biases distorting the process of tree recovery, because both tRNA and the protein sequences must have evolved on the same tree. The test used here comes from Waddell (1995) and here uses the KH approximation of the standard error of a likelihood difference.

To test if the optimal trees for two independent data sets could be different as a result of sampling error alone, we ask: Is

TABLE 6. Testing the position of elephant (for abbreviations and explanations see Table 5)

Elephant Position in ProtML tree					Elephant Position in LogDet tree (figure 1)		
	Pars	WPars	ProtML	ProtML–inv		Pars	WPars
Hedgehog	18	626	– 71.2*	– 65.1*	Hedgehog	29	1667
PL – Hedge	21	1046*	– 57.6*	– 50.9*	PL – Hedge	45	2146
Mur	25	1345*	– 65.7*	– 57.6*	Rod	49*	2441**
PL – Hedge – Mur	13	668	– 27.6	– 22.7	Cav	50*	2741**
fer + arm	12	887*	– 20.6	– 16.2	Mur	44**	3985**
arm	best	0	0	0	PL – Hedge – Rod	46**	3336**
7 PL	31	1545*	– 85.9*	– 72.7*	Pr	61	872
rab + Cav + Pr	19	1116*	– 36.5	– 29.6	PL – Hedge – Rod – Pr	45*	2027*
rab	21	1563*	– 56.3*	– 46*	PL	73*	2930*
Cav + Pr	22	1279*	– 64.8*	– 53.1*	rab	81	2002
Cav	31*	2431**	–102.3**	– 83.6*	fer + arm	74	1757
Pr	8	783	– 76.6*	– 68.5*	arm	71	1197
Fereuungulates	13	1257*	– 31.3*	– 24.3*	fer	73	2422
Cetart	44**	3153**	–123**	–103.6**	Cetart	108*	4370**
Whale	40*	3239**	–133.6**	–114.4**	Whale	116**	4733**
Cow	54**	4153**	–160.9**	–135.4**	Cow	109**	5577**
Pe + Ca	60**	4193**	–151**	–125**	Pe + Ca	113**	4672**
Pe	46**	3338**	–128.3**	–107.1**	Pe	123**	5598**
Ca	63**	4323**	–161.2**	–133.5**	Ca	137**	5793**
Mam	56*	3505**	–222.5**	–198.3**	Mam	41**	4862**
Ionta	78**	4586**	–231.2**	–202.2**	Ionta	74**	5657**
Platypus	108**	7015**	–347.5**	–296.5**	Platypus	74**	8222**
Marsup	111**	7225**	–343.8**	–293.3**	Marsup	74**	8443**
Opossum	189**	12430**	–625.8**	–542.6**	Opossum	144**	13938**
Wallaroo	188**	12790**	–631**	–544.9**	Wallaroo	150**	14076**

there any tree (call it the median tree) such that the sum of squared standardized likelihood deviations from the data set-optimal trees is not significant? (Here, under the null model, which has the median tree specified in advance, we assume the test statistic is χ^2-distributed d.f. 2, so $\alpha A = 0.05$ is 5.99.) If the median tree is unknown in advance, it will often be taken to be the tree with the highest summed likelihood from the two data sets (although such searching will bias the test towards rejecting too rarely). Note, if the KH test for tree 1 to tree 2 on data set 1 is not significant, then tree 2 can fill the role of the median tree and one need no longer search for a better approximation to the median tree (because the deviation from tree 1 to 2 on data set 1 is nonsignificant, and the deviation from tree 2 to tree 2 on data set 2 is zero, then the overall χ^2 statistic cannot be significant).

In this instance, testing the pruned best ProtML tree against the best ML tree for the tRNA data by using the tRNA data yields a nonsignificant result. For example, under the HKY (Hasegawa et al., 1985; Swofford et al., 1996) model with κ estimated on each tree gives $P = 0.098$ (or 0.213, when the coelacanth is forced to its most likely biological position, so that the only differences are within mammals). Next, optimizing invariant sites as well gives $P = 0.082$ (and $P = 0.163$). However, if we expand to the GTR model (Barry and Hartigan, 1987, with the symmetric component of the rate matrix optimized on each tree; Swofford et al., 1996), then $P = 0.059$ (and $P = 0.148$), whereas with this model plus invariant sites, $P = 0.048^*$ (and $P = 0.103$). Lastly, if we optimize a GTR matrix (including the base composition parameters) plus a mixed invariant sites-gamma distribution (Waddell and Penny, 1996) $P = 0.032^*$ (and $P = 0.068$).

The last instance is getting closer to rejection (the sum of squared deviation being $2.17^2 + 0^2 \approx 4.7$ without coelacanth being moved), and if the apparent trend continues, we might reject the null hypothesis if we could model the process of evolution more accurately.

DISCUSSION

These analyses add strong new evidence that the traditional superordinal taxon Ungulata is not monophyletic (e.g., Miyamoto and Goodman, 1986; Irwin and Wilson, 1993; Graur et al., 1997; Springer et al., 1997; Stanhope et al., 1998; and contrary to recent morphological studies, e.g., Gaudin et al., 1996; McKenna and Bell, 1997; Shoshani and McKenna, 1998). From analyses herein, the superordinal taxon Ferungulata (Simpson, 1945), being the grouping of the living taxa Paenungulata, Artiodactyla, Perissodactyla, and Carnivora, is almost certainly not monophyletic also. There is, however, good evidence from mtDNA (e.g., Xu et al., 1996; this paper) and selected analyses of nuclear genes (Waddell, unpubl.) that a substantially modified grouping of orders is likely to be correct. This group appears to contain pangolin, carnivores, perissodactyls, cetartiodactyls, and bats, as discussed in Waddell, Okada et al., 1999.

As discussed in Waddell, Okada et al., 1999b, we are now faced with two new, and we believe strong, hypotheses of the superordinal relationships of placentals. We name the superordinal group mentioned immediately above, Scrotifera. The name comes from the word scrotum, a pouch in which the testes permanently reside in the adult male. All members of the group have a postpenile scrotum, often prominently displayed, except for some aquatic forms and pangolin (which has the testes just below the skin). It appears to be an ancestral character for this group, yet other orders generally lack this as an ancestral feature, with the probable exception of Primates. The definition of Scrotifera is the crown group defined by the common ancestor of the extant orders Pholidota, Carnivora, Cetartiodactyla (Cetacea plus Artiodactyla), Perissodactyla,

and Chiroptera (that is, this specific ancestor and all its descendants).

The second group we name is Fereuungulata, the name being derived from the Latin "fer" alluding to Ferae (the carnivores), and "euuungulata" meaning the true (e.g., hoofed) Ungulates. The definition of this taxon is the crown group defined by the common ancestor of the living orders Pholidota, Carnivora, Cetartiodactyla, and Perissodactyla. This group is expected to include the order Pholidota (pangolins), possibly as closest relatives to carnivores, while bats are probably a sister taxon to fereuungulates (Pumo et al., 1998; Waddell et al., 1999b).

An exciting finding has been support for elephant with armadillo, based on the mtDNA sequences. While support has fluctuated with different methods (from a high of 89% with ProtML local RELL, to much less with NJ), this is true of nearly all the hypotheses based on these data. Rather, it is the congruence of this result that is interesting. Between the optimal tRNA ML and ProtML trees, the only mammalian groups in common are fereuungulates, Cetartiodactyla, Placentalia, Marsupialia, and the armadillo/elephant. Thus, for the other groups, congruence has singled out all of those mammalian groups with the best support from diverse data.

Since this paper was submitted, combined 12s–16s mt rRNA data appeared in Stanhope et al. (1998), which also places the Afrotheria (of which elephant is a part) with a xenarthran. Thus, the mtDNA data seem fairly unanimous on this. Some nuclear genes (vWF and IRBP) also give this hypothesis some support (Springer et al., 1997). Further support for this grouping comes from biogeography and the coincidence of the age of this group with the opening of the South Atlantic (Waddell et al., 1999).

Surprisingly, the AA LogDet/NJ combination gave almost exactly the same trees as the observed distances with NJ—despite evidence for unequal amino acid composition, principally outside of the Fereuungulata. It is possible that even the CSR LogDet is not making sufficient correction for the nonhomogeneous model. This may be be-

cause sites evolving at different rates are nonstationary to differing degrees, but the CSR LogDet as implemented here can "see" only the base composition differences averaged across all nonzero rate classes (Waddell, 1995:127).

Of the hypotheses we set out to test, (1) Marsupionta and (2) Carnivora + Perissodactyla remain interesting hypotheses, but their support is seen to vary considerably with the part of the mtDNA and the methods used. (3) How close Xenarthra is to the root is uncertain, in that the root could come rather close to it if groups in (5) and (6) are correct. (4) Hedgehog appears to be an early split with most bootstrap analyses, but with specific pairwise tests, it can move inside rodents (and if rodents are really part of an extended Glires-like group, then this means not so uniquely deep). (5) Rodentia is moderately supported, Glires still holds promise, and even Lagomorpha + Primates crops up on rare occasions. (6) tRNA suggests the group Miyamoto and Goodman (1986) noticed and may add hedgehog to this group. (7) Elephants (representing Afrotheria) are suggested as a specific relative of armadillo, as discussed above.

On a more cautionary note, the whole pattern of the earliest divergences amongst the deeper taxa should be considered open to review. For example, in one instance, using more-conservative sites and ProtML, we obtain Glires, whereas in another, using tRNA and a GTR ML with site-rate heterogeneity, we obtain a tree with primates being sister to rodents and these being sister again to hedgehog. With such rearrangements, the root could possibly move to be sister to the elephant/armadillo group. Further, in Waddell et al. (1999b) we suggest that resolution of the earliest branching orders could be compromised by a closely spaced series of splits in the middle to early Cretaceous. It is important to reanalyze these emerging trends by adding in the now considerable evidence from well sampled nuclear protein sequences.

REFERENCES

ADACHI, J., AND M. HASEGAWA. 1996a. Model of amino acid substitution in proteins encoded by mitochondrial DNA. J. Mol. Evol. 42: 459–468.

ADACHI, J., AND M. HASEGAWA. 1996b. MOLPHY: version 2.3: Programs for molecular phylogenetics based on maximum likelihood. Comput. Sci. Monogr. 28. Institute of Statistical Mathematics, Tokyo.

ARNASON, U., A. GULLBERG, AND A. JANKE. 1997. Phylogenetic analyses of mitochondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and ferungulates. Mol. Biol. Evol. 14: 762–768.

ARNASON, U., A. GULLBERG, AND X. XU. 1996. A complete mitochondrial DNA molecule of the white-handed gibbon, *Hylobates lar*, and comparison among individual mitochondrial genes of all hominoid genera. Hereditas 124:185–189.

BANDELT, H.-J., AND A. W. M. DRESS. 1992. Split decomposition: A new and useful approach to phylogenetic analysis of distance data. Mol. Phylogenet. Evol. 1:242–252.

BARRY, D., AND J. A. HARTIGAN. 1987. Asynchronous distance between homologous DNA sequences. Biometrics 43: 261–276.

BRYANT, D., AND P. J. WADDELL. 1998. Rapid evaluation of least squares and minimum-evolution criteria or phylogenetic trees. Mol. Biol. Evol. 15:1346–1359.

CAO, Y., A. JANKE, P. J. WADDELL, M. WESTERMAN, O. TAKENAKA, S. MURATA, N. OKADA, S. PÄÄBO, AND M. HASEGAWA. 1998. Conflict among individual mitochondrial proteins in resolving the phylogeny of eutherian orders. J. Mol. Evol. 47: 307–322.

CAO, Y., P. J. WADDELL, N. OKADA, AND M. HASEGAWA. 1998. The complete mitochondrial DNA sequence of the shark *Mustelus manazo*: Evaluating rooting contradictions with living bony vertebrates. Mol. Biol. Evol. 15:1637–1646.

DE JONG, W. W., J. A. M. LEUNISSEN, AND G. J. WISTOW. 1993. Eye lens crystallins and the phylogeny of placental orders: Evidence for a macroscelid–paenungulate clade? Pages 5–12, *in* Mammal phylogeny: Placentals (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.

D'ERCHIA, A.M., C. GISSI, G. PESOLE, C. SACCONE, AND U. ARNASON. 1996. The guinea pig is not a rodent. Nature 381:597–600.

FELSENSTEIN, J. 1993. PHYLIP: Phylogeny inference package, version 3.5c. Department of Genetics, Univ. of Washington, Seattle.

GAUDIN, T. J., J. R. WIBLE, J. A. HOPSON, AND W. D. TURNBULL. 1996. Reexamination of the morphological evidence for the cohort Epitheria (Mammalia, Eutheria). J. Mammal. Evol. 3:31–79.

GRAUR, D., L. DURET, AND M. GOUY. 1996. Phylogenetic position of the order Lagomorpha. Nature 379:333–335.

GRAUR, D., M. GOUY, AND L. DURET. 1997. Evolutionary affinities of the order Perissodactyla and the phylogenetic status of the superordinal taxa Ungulata and Altungulata. Mol. Phylogenet. Evol. 7: 195–200.

GREGORY, W. K. 1910. The orders of mammals. Bull. Am. Mus. Nat. Hist. 27:1–524.

GREGORY, W. K. 1947. The monotremes and the palimpsest theory. Bull. Am. Mus. Nat. Hist. 88:1–88.

HALANYCH, K. M. 1998. Lagomorphs misplaced by more characters and fewer taxa. Syst. Biol. 47:138–146.

HARLID, A., A. JANKE, AND U. ARNASON. 1997. The mtDNA sequence of the ostrich and the divergence between paleognathous and neognathous birds. Mol. Biol. Evol. 14:754–776.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21:160–174.

IRWIN, D. M., AND A. C. WILSON. 1993. Limitations of molecular methods for establishing the phylogeny of mammals, with special reference to the position of elephants. Pages 257–267 in Mammal phylogeny: Placentals (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.

JANKE, A., AND U. ARNASON. 1997. The complete mitochondrial genome of Alligator mississippiensis and the separation between recent Archosauria (birds and crocodiles). Mol. Biol. Evol. 14:1266–1272.

JANKE, A., N. J. GEMMELL, G. FELDMAIER-FUCHS, A. VON HAESELER, AND S. PÄÄBO. 1996. The complete mitochondrial genome of a monotreme, the platypus (Ornithorhynchus anatinus). J. Mol. Evol. 42:153–159.

JANKE, A., X. XU, AND U. ARNASON. 1997. The complete mitochondrial genome of the wallaroo (Macropus robustus) and the phylogenetic relationship among Monotremata, Marsupialia and Eutheria. Proc. Natl. Acad. Sci. USA 94:1276–1281.

JOHANSEN, S., AND I. BAKKE. 1996. The complete mitochondrial DNA sequence of Atlantic cod (Gadus morhua): relevance to taxonomic studies among codfishes. Mol. Mar. Biol. Biotechnol. 5:203–214.

KISHINO, H., AND M. HASEGAWA. 1990. Evaluation of the maximum likelihood estimate of evolutionary tree topologies from DNA sequence data, and the branching order of Hominoidea. J. Mol. Evol. 29:170–179.

KRETTEK, A., A. GULLBERG, AND U. ARNASON. 1995. Sequence analysis of the complete mitochondrial DNA molecule of the hedgehog, Erinaceus europaeus, and the phylogenetic position of the Lipotyphla. J. Mol. Evol. 41:952–957.

KUMAZAWA, Y., AND M. NISHIDA. 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. J. Mol. Evol. 37:380–398.

LAKE, J. A. 1994. Reconstructing evolutionary trees from DNA and protein sequences: Paralinear distances. Proc. Natl. Acad. Sci. USA. 91:1455–1459.

LEE, W. J., AND T. D. KOCHER. 1995. Complete sequence of a sea lamprey (Petromyzon marinus) mitochondrial genome: early establishment of the vertebrate genome organization. Genetics 139:873–887.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, AND D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11:605–612.

LOPEZ, J. V., M. CULVER, S. CEVARIO, AND S. J. O'BRIEN. 1996. Complete nucleotide sequences of the domestic cat (Felis catus) mitochondrial genome and a transposed mtDNA repeat (Numt) in the nuclear genome. Genomics 33:229–246.

McKENNA, M. C. 1975. Toward a phylogenetic classification of the Mammalia. Pages 21–46 in Phylogeny of the Primates: A multidisciplinary approach (W. P. Luckett and F. Szalay, eds.). Plenum, New York.

McKENNA, M. C., AND S. K. BELL. 1997. Classification of mammals above the species level. Columbia University Press, New York.

MINDELL, D. P., M. D. SORENSON, C. J. HUDDLESTONE, H. C. MIRANDA, A. KNIGHT, S. J. SAWCHUK, AND T. YURI. 1997. Phylogenetic relationships within and among select avian orders based on mitochondrial DNA. Pages 214–243 in Avian molecular evolution and systematics (D. P. Mindell, ed.). Academic Press, San Diego.

MIYAMOTO, M. M., AND M. GOODMAN. 1986. Biomolecular systematics of eutherian mammals: Phylogenetic patterns and classification. Syst. Zool. 35:230–240.

NOVACEK, M. J. 1993. Reflections on higher mammalian phylogenetics. J. Mammal. Evol. 1:3–30.

PRESS, W. H., S. A. TEUKOLSKY, W. T. VETTERLING, AND B. P. FLANNERY. 1995. Numerical recipes in C: The art of scientific computing, 2nd edition. Cambridge Univ. Press, Cambridge, England.

PUMO, D. E., P. S. FINAMORE, W. R. FRANEK, C. J. PHILLIPS, S. TARZAMI, AND D. BALZARANO. 1998. Complete mitochondrial genome of a Neotropical fruit bat, Artibeus jamaicensis and a new hypothesis of relationships of bats to other eutherian mammals. J. Mol. Evol. 15:709–717.

RASMUSSEN, A. S., JANKE, A., AND ARNASON, U. 1998. The mitochondrial DNA molecule of the hagfish (Myxine glutinosa) and vertebrate phylogeny. J. Mol. Evol. 46:382–388.

ROWE, T. 1993. Phylogenetic systematics and the early history of mammals. Pages 129–145 in Mammal phylogeny: Mesozoic differentiation, multituberculates, monotremes, early eutherians, and marsupials (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, N.Y.

SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 183:584–598.

SEBER, G. A. F. 1982. The estimation of animal abundance. Charles Griffin, London.

SHOSHANI, J., AND M. C. McKENNA. 1998. Higher taxonomic relationships among extant mammals based

on morphology, with selected comparisons of results from molecular data. Mol. Phylogenet. Evol. 9:572–584.

SIDOW, A., T. NGUYEN, AND T. P. SPEED. 1992. Estimating the fraction of invariable codons with a capture–recapture method. J. Mol. Evol. 35:253–260.

SIMPSON, G. G. 1945. The principles of classification and the classification of mammals. Bull. Am. Mus. Nat. Hist. 85:1–350.

SPRINGER, M. S., G. C. CLEVEN, O. MADSEN, W. W. DE JONG, V. G. WADDELL, H. M. AMRINE, AND M. J. STANHOPE. 1997. Endemic African mammals shake the phylogenetic tree. Nature 388:61–64.

STANHOPE, M. J., V. G. WADDELL, O. MADSEN, W. DE JONG, C. B. HEDGES, G. C. CLEVEN, D. KAO, AND M. S. SPRINGER. 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. Proc. Natl. Acad. Sci. USA 95:9967–9972.

STIASSNY, M. L. J., L. R. PARENTI, AND G. D. JOHNSON. 1996. Interrelationships of fishes. Academic Press, San Diego.

SULLIVAN, J., AND D. L. SWOFFORD. 1997. Are guinea pigs rodents? The importance of adequate models in molecular phylogenetics. J. Mammal. Evol. 4:77–86.

SWOFFORD, D. L. 1998. PAUP*: Phylogenetic analyses using parsimony and other methods, version 4.0 (test ver. 61–64). Sinauer, Sunderland, Massachusetts.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, AND D. M. HILLIS. 1996. Phylogenetic inference. Pages 407–514 in Molecular systematics, 2nd edition (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.

SZALAY, F. S. 1993. Pedal evolution of mammals in the Mesozoic: Tests for taxic relationships. Pages 108–128 in Mammal phylogeny: Mesozoic differentiation, multituberculates, monotremes, early eutherians, and marsupials (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.

WADDELL, P. J. 1995. Statistical methods of phylogenetic analysis: Including Hadamard conjugations, LogDet transforms, and maximum likelihood. Ph.D. Thesis, Massey University, Palmerston North, New Zealand.

WADDELL, P. J., Y. CAO, M. HASEGAWA, AND D. MINDELL. 1999. Assessing the Cretaceous superordinal divergence times within birds and placental mammals by using whole mitochondrial protein sequences and an extended statistical framework. Syst. Biol. 48:119–137 (this issue).

WADDELL, P. J., P. O. LEWIS, AND D. L. SWOFFORD. In press. Distance-based methods of inferring evolutionary trees. In PAUP*: Phylogenetic analysis using parsimony and other methods, version 4.0, D. L. Swofford. Sinauer, Sunderland, Massachusetts.

WADDELL, P. J., H. OKADA, M. HASEGAWA. 1999. Toward resolving the interordinal relationships of mammals. Syst. Biol. 48:1–5 (this issue).

WADDELL, P. J., AND D. PENNY. 1996. Evolutionary trees of apes and humans from DNA sequences. Pages 53–73 in Handbook of human symbolic evolution (A.J. Lock and C. R. Peters, eds.). Oxford Univ. Press, Oxford, England.

WADDELL, P. J., D. PENNY, AND T. MOORE. 1997. Hadamard conjugations and modeling sequence evolution sequence evolution with variable rates across sites. Mol. Phylogenet. Evol. 8:33–50.

WADDELL, P. J., AND M. A. STEEL. 1997. General time reversible distances with unequal rates across sites: Mixing Γ and inverse Gaussian distributions with invariant sites. Mol. Phylogenet. Evol. 8:398–414.

XU, X., A. JANKE, AND U. ARNASON. 1996. The complete mitochondrial DNA sequence of the greater Indian rhinoceros, *Rhinoceros unicornis*, and the phylogenetic relationship among Carnivora, Perissodactyla, and Artiodactyla (+ Cetacea). Mol. Biol. Evol. 13:1167–1173.

ZARDOYA, R., AND A. MEYER. 1996. The complete nucleotide sequence of the mitochondrial genome of the lungfish (*Protopterus dolloi*) supports its phylogenetic position as a close relative of land vertebrates. Genetics 142:1249–1263.

ZARDOYA, R., AND A. MEYER. 1997. The complete DNA sequence of the mitochondrial genome of a "living fossil", the coelacanth (*Latimeria chalumnae*). Genetics 146:995–1010.