

What is a Suiforme (Artiodactyla)?

Contribution of Cranioskeletal and Mitochondrial DNA Data

Claudine Montgelard,^{*,†} Stéphane Ducrocq,[†] and Emmanuel Douzery[†]

^{*}Laboratoire de Paléontologie des Vertébrés (EPHE), and [†]Laboratoire de Paléontologie, Paléobiologie et Phylogénie, CC064, Institut des Sciences de l'Evolution, UMR 5554 (CNRS), Université Montpellier II, Place E. Bataillon, 34095 Montpellier Cedex 05, France

Received July 2, 1997; revised December 16, 1997

Suiformes (Artiodactyla) traditionally includes three families: Suidae, Tayassuidae, and Hippopotamidae but the monophyly of this suborder has recently been questioned from molecular data. A maximum parsimony analysis of molecular, morphological, and combined data was performed on the same set of taxa including representatives of the three Artiodactyla suborders (Suiformes, Ruminantia, and Tylopoda) and Perissodactyla as outgroup. Mitochondrial (cytochrome *b* and 12S rRNA) sequence comparisons support the monophyly of Suina (Suidae and Tayassuidae) and Ancodonta (Hippopotamidae) but not the monophyly of Suiformes. Inversely, our preliminary morphological analysis supports the monophyly of Suiformes whereas relationships among the three families are not resolved. The combined data set does not resolve the relationships between Suina, Ancodonta, and Ruminantia. These results are discussed in relation to morphological characters and paleontological data. Some improvements are suggested to clarify the morphological definition of Suiformes and relationships among them. © 1998 Academic Press

are nonruminant and they exhibit dental and skeletal characters that are very primitive among Artiodactyls (e.g., Gentry and Hooker, 1988; Pickford, 1993; Ducrocq, 1994; Ducrocq *et al.*, in press), making their phylogenetic relationships difficult to establish from a morphological point of view. Close relationships between Suidae and Tayassuidae are supported by paleontological data (Viret, 1961) and molecular data (Gatesy *et al.*, 1996; Randi *et al.*, 1996; Montgelard *et al.*, 1997; Gatesy, 1997). The position of Suina with respect to Hippopotamidae is more debated. From morphological and paleontological data, hippopotamids are alternatively regarded as relatives to tayassuids (Pickford, 1993) or to an extinct artiodactyl family, the Anthracotheriidae (Gentry and Hooker, 1988). However, recent molecular studies suggest that hippopotamids could be more closely related to cetaceans (whales and dolphins) than to suids and tayassuids (Irwin and Arnason, 1994; Gatesy *et al.*, 1996; Randi *et al.*, 1996; Montgelard *et al.*, 1997; Gatesy, 1997; Shimamura *et al.*, 1997). These results indicate that Suiformes may constitute a paraphyletic suborder defined on homoplastic dental and osteological characters.

The scope of this study was to test, simultaneously from molecular and morphological characters, the validity of the Suiformes clade as well as relationships between the three extant families Suidae, Tayassuidae, and Hippopotamidae. Molecular and morphological data were analyzed on the same set of taxa by the maximum parsimony approach. Our survey includes living representatives of the three families of Suiformes and also of the two other suborders of Artiodactyla (Ruminantia and Tylopoda). Cetaceans were not included in this analysis because they present too much divergence in features for morphological comparisons. Molecular studies are based on the two mitochondrial cytochrome *b* and 12S rRNA genes (Montgelard *et al.*, 1997), and morphological analyses, although preliminary, include dental and skeletal characters.

INTRODUCTION

According to traditional systematics, the suborder Suiformes Jaekel, 1911 is included in the order Artiodactyla together with Ruminantia (cows, deers, chevrotains, and others) and Tylopoda (camelids). Suiformes includes extant members of the families Suidae (pigs and hogs with five genera), Tayassuidae (peccaries with three genera) and Hippopotamidae (hippos with two genera), in addition to several other fossil families (see, for example Viret, 1961). Among living Suiformes, Simpson (1945) recognized two infraorders, Suina (suids and tayassuids) and Ancodonta (hippopotamids).

Although these three families are thought to be phylogenetically related, their representatives testify to a rather old evolutionary history. Extant Suiformes

MATERIAL AND METHODS

Molecular analysis of cytochrome b and 12S rRNA sequences. Eleven taxa were represented: (1) Suiformes: *Sus scrofa* (Suidae), *Tayassu tajacu* (Tayassuidae), *Hippopotamus amphibius*, and *Hexaprotodon liberiensis* (Hippopotamidae); (2) Tylopoda: *Camelus bactrianus* and *Lama guanicoe*; (3) Ruminantia: *Tragulus napu* (Tragulidae), *Bos taurus*, and *Cervus nippon* or *Cervus unicolor* (Pecora); and (4) two Perissodactyla (*Equus grevyi* and *Diceros bicornis* or *Ceratotherium simum*) used as outgroups. Sequence references can be found in Montgelard *et al.* (1997).

For both genes, all analyses were based on low-saturating events. For the 12S rRNA gene, transversions only in loops and both transition and transversion events in stems were conserved (Springer and Douzery, 1996) whereas indels were omitted. For cytochrome *b* low-saturating events are Conservative Nucleotide Substitutions (as defined in Arnason and Gullberg, 1996).

Morphological analysis. Seven living lineages were characterized: (1) the three families of Suiformes (Suidae, Tayassuidae, and Hippopotamidae); the suborders Tylopoda (Camelidae) and Ruminantia (Bovidae and Tragulidae); and (3) the Perissodactyla lineage Hippomorpha was used as outgroup. Forty cranioskeletal characters were defined; most of them are issued from the analysis of Gentry and Hooker (1988). The list of characters and the morphological matrix are available on the MPE web site or from S.D. upon request. Analyses were conducted using nonordered character states.

In both molecular and morphological analyses, the phylogenetic content was tested with the g_1 statistic (Hillis and Huelsenbeck, 1992), estimated by generating 10^5 trees with the random-trees option in PAUP Version 3.1.1 (Swofford, 1993). Phylogenetic tree search was conducted with PAUP using the branch-and-bound search with TBR branch-swapping option. Robustness of the nodes was assessed by the bootstrap method (Felsenstein, 1985) after 1000 replications (bootstrap option with heuristic search in PAUP) and the decay index (DI, Bremer, 1988) using topological constraints enforced in PAUP. DI corresponds to the number of extrasteps either to remove a grouping (DI positive) or to constrain the monophyly of the clade considered (DI negative).

The level of incongruence between data sets (between the two genes and between molecular and morphological data) was tested with the program ARNIE included in the Random Cladistics package (Siddall, 1996). ARNIE used the incongruence length difference (ILD) test with the parsimony approach (Farris *et al.*, 1995). One-thousand randomizations were performed with unordered characters, multiple equally parsimonious trees (mh), and branch-swapping (bb) options. Invari-

ant sites were removed from analyses (Cunningham, 1997).

RESULTS

Molecular Analysis

The number of analyzed sites represent 1137 positions (287 variable and 241 informative) for the cytochrome *b* and 866 sites (219 variable and 150 informative) for the 12S rRNA. The cytochrome *b* and 12S rRNA sequences appeared fully congruent ($P > 0.05$ with 506 variable positions).

As already noted from a more complete data set (Montgelard *et al.*, 1997), DNA sequences from either separate genes or genes in combination produced significant skewed tree-length distribution ($g_1 = -1.16$ for the cytochrome *b*, $g_1 = -1.21$ for the 12S rRNA, and $g_1 = -1.27$ for the two genes in combination; $P < 0.01$ in the three cases). Lower consistency and retention indices are obtained for the cytochrome *b* (length of the most parsimonious tree (L) = 550, CI = 0.47, and RI = 0.49) relative to the 12S rRNA (L = 309, CI = 0.53 and RI = 0.57); the two genes in combination (Fig. 1A) producing intermediate values (L = 866, CI = 0.49, and RI = 0.51).

All analyses based on separate or combined genes indicate a strong support (BP = 100%, DI = +21 to +58) for the monophyly of Hippopotamidae. A moderate to strong support (63 to 97% BP, DI = +3 to +13) is given to the monophyly of Suina (Suidae and Tayassuidae). Alternative hypotheses (Suidae + Hippopotamidae or Tayassuidae + Hippopotamidae) are poorly supported (BP range 0–2%, DI = –22 to –9).

Suiformes always appeared paraphyletic because Suina are either grouped with Ruminantia (BP = 61%) in cytochrome *b* analysis or external (BP = 60%) to the remaining Artiodactyla with the 12S rRNA, whereas the combined data set do not clearly resolve relationships between the four lineages (Suina, Hippopotamidae, Ruminantia, and Tylopoda). The monophyly of Suiformes received negligible support in all analyses (BP < 10%, DI = –4 to –2).

Morphological Analysis

The morphological matrix (40 variable and 30 informative characters) produced skewed distribution although the significance level is close to 5% ($g_1 = -0.59$, $0.05 > P > 0.01$). Two most parsimonious trees (L = 72, CI = 0.68, and RI = 0.54) were recovered (Fig. 1B).

Suiformes appears monophyletic with strong support (BP = 90%, DI = +4) whereas relationships among other Artiodactyla are poorly resolved (50% BP for Ruminantia). Among Suiformes, Suidae either clustered with Tayassuidae or with Hippopotamidae but with a very low support. The alternative Suidae + Tayassuidae hypothesis appeared in 24% of bootstrap replicates. When nonhomoplastic characters (with

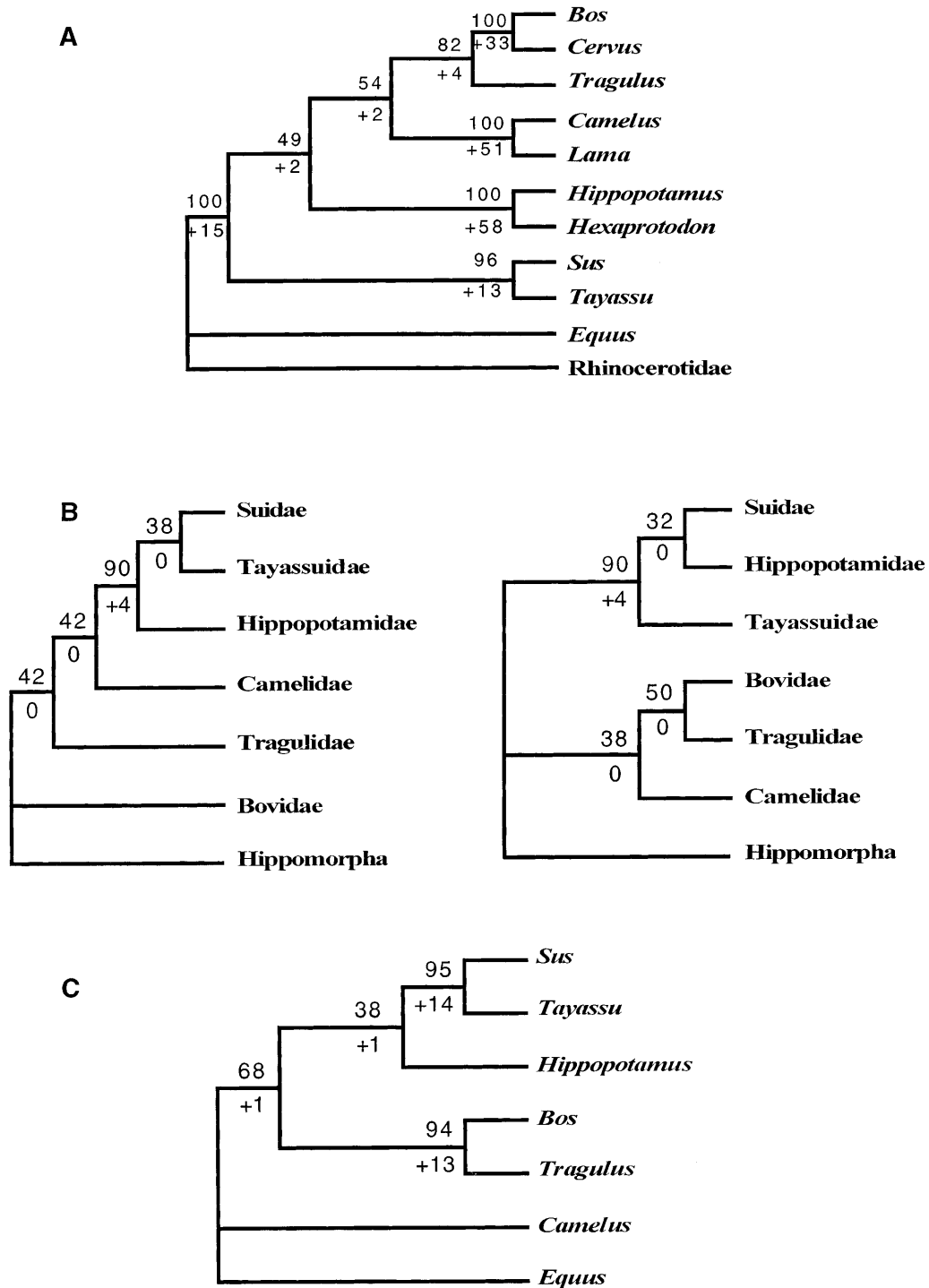


FIG. 1. Most-parsimonious trees recovered with: (A) combined cytochrome b and 12s rRNA sequences (506 variable and 291 informative characters; $L = 866$, $CI = 0.49$, $RI = 0.51$); (B) morphological analysis (40 variable and 30 informative characters; $L = 72$, $CI = 0.68$, $RI = 0.54$); (C) molecules and morphology in combination (546 variable and 321 informative characters; $L = 891$, $CI = 0.65$, $RI = 0.32$). Bootstrap proportions from 1000 replicates are indicated above the branches and decay index values are reported below.

$CI = 1$) are considered, Suiformes are characterized by a single synapomorphic change of the character amastoidy. Other unambiguous changes correspond to reversions with respect to character states defined in the hippomorpha outgroup. Among them, we found the

character bunodonty which was the traditional dental character used by morphologists to define Suiformes (Gentry and Hooker, 1988). Suidae appears to be characterized by numerous autapomorphies while several reversions appeared in the branch leading to Hippopotamidae.

Combined Analysis

Because combining incongruent data sets can reduce phylogenetic accuracy (Cunningham, 1997), our molecular and morphological partitions were tested for incongruence before combination. The taxonomic sampling was reduced to seven taxa (*Sus*, *Tayassu*, *Hippopotamus*, *Bos*, *Tragulus*, *Camelus*, and *Equus* as outgroup) which are common to morphological and molecular studies. Our data were partitioned in three data sets including the cytochrome *b* (287 variable sites), the 12S rRNA (219 variable positions), and the 40 morphological characters. The test reveals no significant incongruence between data partitions ($P > 0.05$).

The phylogenetic content of the whole data set was significant ($g_1 = -0.49$; $P = 0.01$). The most parsimonious tree ($L = 891$, $CI = 0.65$, $RI = 0.32$) recovered from the combined analysis is presented in Fig. 1C. Inversely to the molecular analysis, Suiformes appeared monophyletic but with a low support (BP = 38, $DI = +1$). Suina (*Sus* and *Tayassu*) and Ruminantia (*Bos* and *Tragulus*) appeared strongly supported whereas the associations Tayassuidae–Hippopotamidae and Suidae–Hippopotamidae are not (BP = 0% and 1%, $DI = -20$ and -18 , respectively).

DISCUSSION

The two mitochondrial genes provide strong support for the monophyly of the infraorders Suina and Ancodonta, as already noted in numerous studies on mitochondrial (Irwin *et al.*, 1991; Irwin and Arnason, 1994; Douzery and Catzeflis, 1995; Randi *et al.*, 1996; Montgelard *et al.*, 1997) or nuclear genes (Gatesy *et al.*, 1996; Gatesy, 1997). By contrast, the suborder Suiformes (Suina + Ancodonta) appears either paraphyletic or at least its monophyly is very weakly supported. With the mitochondrial cytochrome *b* gene (Irwin and Arnason, 1994; Randi *et al.*, 1996; Hasegawa and Adachi, 1996; Montgelard *et al.*, 1997), Suiformes are clearly paraphyletic due to the association between Hippopotamidae and Cetacea. This unexpected grouping is also supported by the nuclear β and κ casein (Gatesy *et al.*, 1996) and γ -fibrinogen genes (Gatesy, 1997). With other available molecular data, such as the mitochondrial 12S rRNA (Douzery and Catzeflis, 1995; Montgelard *et al.*, 1997), amino acid sequences of seven nuclear polypeptides (Miyamoto and Goodman, 1986), the pancreatic ribonuclease (Beintema *et al.*, 1988), and the α and β hemoglobins (Hasegawa and Adachi, 1996), neither the association Hippopotamidae + Cetacea nor the monophyly of Suiformes is supported. Moreover, a recent work on SINES retroposons (Shimamura *et al.*, 1997) make also Suiformes paraphyletic because it supports the monophyly of whales, hippos, and ruminants to the exclusion of pigs and camels.

According to our preliminary morphological study, Suiformes appears monophyletic but defined by a single nonhomoplastic synapomorphy: the amastoidy. This

character is however known to occur in other living (Tubulidentata, some Cetacea and Carnivora; Grassé, 1955) and fossil (Anthracotheriidae, Entelodontidae; Pearson, 1927) groups and the question of its phylogenetic validity must be clarified.

Relationships among Suiformes appear controversial. According to Pickford (1989, 1993), Hippopotamidae could be issued from an Old World Tayassuid stock and hence should be more related to Tayassuidae than to Suidae. According to Gentry and Hooker (1988), Hippopotamidae could have originated from the extinct family Anthracotheriidae, with Suidae and Tayassuidae being more closely related to each others. In our morphological analyses, although the association Hippopotamidae–Tayassuidae represents the less supported grouping (24% BP), the two other associations Suidae–Tayassuidae or Suidae–Hippopotamidae are equally supported (38% BP). Relationships among other Artiodactyla appeared also unresolved: Ruminantia are not clearly defined (50% BP). However, the level of resolution is strongly character dependent. The addition of two soft-tissue characters (thick skin and compartmented stomach; data not shown) allowed a good definition of Ruminantia and favor the cluster Suidae–Hippopotamidae. In the combined analysis, Suiformes appeared monophyletic with a low support whereas Suina remains strongly supported, which is consistent with an anthracothere ancestry for hippos.

However, it must be mentioned that our morphological study is rather preliminary and several modifications could be introduced to improve the morphological resolution. (1) It will be necessary to increase the number of morphological characters and to strengthen the analysis concerning homology of characters (for example, amastoidy or bunodonty).

(2) A further analysis will need inclusion of paleontological data. The fossil record of Suidae and Tayassuidae goes back to at least the Late Eocene or even earlier (i.e., 45 myr; Ducrocq, 1994; Ducrocq *et al.*, in press), whereas the oldest true hippos are known in deposits not older than the Late Miocene (about 8 myr; Pickford, 1993). Because of this gap between the recording of fossil Hippopotamidae and the long history of others living Suiformes (about 45 myr), it is clear that further studies need to include fossil Suiformes. The Anthracotheriidae lineage appeared as a clue taxon to resolve the question of the ancestry of Hippopotamidae and also the question of the monophyly of Suiformes. Cetaceans represent another group to include because of the increasing body of molecular results in favor of the Cetartiodactyla (Montgelard *et al.*, 1997) clade. Despite recent progresses achieved in the paleontology of cetaceans (e.g., Thewissen and Hussain, 1993), the question of a possible association between Hippopotamidae and Cetacea will be more difficult to settle from a morphological point of view, due to the very derived and specialized conditions of cetaceans. Nevertheless, it could be possible to include primitive cetacean fossils

such as *Pakicetus* (Gingerich and Russell, 1981) and others Archaeocetes displaying less-derived characters (Thewissen, 1994). Mesonychidae should not be neglected because of the phylogenetic implications of the Hippo-Cetacea clade on the Cete (mesonychians and Cetacea; Thewissen, 1994) monophyly (see discussion in Gatesy *et al.*, 1996). Furthermore, the addition of taxa will reduce the possible bias of phylogenetic reconstructions based on too few taxa (Philippe and Douzery, 1994).

Finally, the choice of the outgroup remains an important question because Suiformes are anatomically the most primitive among Artiodactyla. We used Hippomorpha but it presents numerous derived characters whereas Suiformes have retained a lot of plesiomorphies. So, Hippomorpha might not be an appropriate outgroup because it implies too many reversions in the Suiformes lineages. We performed analyses in which the outgroup was represented by an hypothetical ancestor possessing ancestral states for all characters (data not shown). In this case, the clade Suidae–Tayassuidae is the most sustained association among Suiformes whereas the monophyly of Suiformes is not supported. In a data set including extinct forms, the fossil perisodactyla *Hyracotherium* could hence represent an appropriate outgroup allowing polarization of characters.

ACKNOWLEDGMENTS

We thank the Service Commun de Biosystématique de Montpellier (attribution 12/92 to C.M.) and the Groupement de Recherche et d'Études des Génomes (GREG, attribution 82/94 to E.D.) for their financial contribution. This is publication 98-019 of the Institut des Sciences de l'Évolution de Montpellier.

REFERENCES

- Arnason, U., and Gullberg, A. (1996). Cytochrome *b* nucleotide sequences and the identification of five primary lineages of extant Cetaceans. *Mol. Biol. Evol.* **13**: 407–417.
- Beintema, J. J., Schüller, C., Irie, M., and Carsana, A. (1988). Molecular evolution of the ribonuclease superfamily. *Prog. Biochem. Mol. Biol. Evol.* **51**: 165–192.
- Bremer, K. (1988). The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* **42**: 795–803.
- Cunningham, C. W. (1997). Can three incongruence tests predict when data should be combined? *Mol. Biol. Evol.* **14**: 733–740.
- Douzery, E., and Catzeflis, F. M. (1995). Molecular evolution of the mitochondrial 12S rRNA in Ungulata (Mammalia). *J. Mol. Evol.* **41**: 622–636.
- Ducrocq, S. (1994). An Eocene peccary from Thailand and the biogeographical origins of the artiodactyl family Tayassuidae. *Palaeontology* **37**: 765–779.
- Ducrocq, S., Chaimanee, Y., Suthetham, V., and Jaeger, J.-J. (in press). The oldest known pig from the Late Eocene of Thailand. *Paleontology*.
- Farris, J. S., Källersjö, M., Kluge, A. G., and Bult, C. (1995). Testing significance of incongruence. *Cladistics* **10**: 315–319.
- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**: 783–791.
- Gatesy, J., Hayashi, C., Cronin, M. A., and Arctander, P. (1996). Evidence from the milk casein genes that Cetaceans are close relatives of hippopotamid artiodactyls. *Mol. Biol. Evol.* **13**: 954–963.
- Gatesy, J. (1997). More DNA support for a Cetacea/Hippopotamidae clade: The blood-clotting protein γ -fibrinogen. *Mol. Biol. Evol.* **14**: 537–543.
- Gentry, A. W., and Hooker, J. J. (1988). The phylogeny of Artiodactyla. In "The Phylogeny and Classification of the Tetrapods" (M. J. Benton, Ed.), pp. 235–272. Clarendon Press, Oxford.
- Gingerich, P. D., and Russell, D. E. (1981). *Pakicetus inakus*, a new archaeocete (Mammalia, Cetacea) from the early-middle Eocene Kuldana Formation of Kohat (Pakistan). *Contr. Mus. Paleont., Univ. Michigan* **25**: 235–246.
- Grassé, P. P. (1955). "Traité de Zoologie. Tome XVI (1er fascicule) Mammifères Tégument Squelette." Masson, Paris.
- Hasegawa, M., and Adachi, J. (1996). Phylogenetic position of cetaceans relative to artiodactyls: Reanalysis of mitochondrial and nuclear sequences. *Mol. Biol. Evol.* **13**: 710–717.
- Hillis, D. M., and Huelsenbeck, J. P. (1992). Signal, noise, and reliability in molecular phylogenetic analyses. *J. Hered.* **83**: 189–195.
- Irwin, D. M., Kocher, T. D., and Wilson, A. C. (1991). Evolution of the cytochrome *b* gene of mammals. *J. Mol. Evol.* **32**: 128–144.
- Irwin, D. M., and Arnason, U. (1994). Cytochrome *b* gene of marine mammals: Phylogeny and evolution. *J. Mammal. Evol.* **2**: 37–55.
- Miyamoto, M. M., and Goodman, M. (1986). Biomolecular systematics of eutherian mammals: Phylogenetic patterns and classification. *Syst. Zool.* **35**: 230–240.
- Montgelard, C., Catzeflis, F. M., and Douzery, E. (1997). Phylogenetic relationships of artiodactyls and cetaceans as deduced from the comparison of cytochrome *b* and 12S rRNA mitochondrial sequences. *Mol. Biol. Evol.* **14**: 550–559.
- Pearson, H. S. (1927). On the skulls of early Tertiary Suidae, together with an account of the otic region in some other primitive Artiodactyla. *Phil. Trans. Royal Soc. Lond. B* **215**: 389–460.
- Pickford, M. (1989). Update on hippo origins. *C. R. Acad. Sci. Paris II* **309**: 163–168.
- Pickford, M. (1993). Old World suoid systematics, phylogeny, biogeography and biostratigraphy. *Paleontol. Evol.* **26–27**: 237–269.
- Philippe, H., and Douzery, E. (1994). The pitfalls of molecular phylogeny based on four species as illustrated by the Cetacea/Artiodactyla relationships. *J. Mammal. Evol.* **2**: 133–152.
- Randi, E., Lucchini, V., and Diong, C. H. (1996). Evolutionary genetics of the suiformes as reconstructed using mtDNA sequencing. *J. Mammal. Evol.* **3**: 163–194.
- Shimamura, M., Yasue, H., Ohshima, K., Abe, H., Kato, H., Kishiro, T., Goto, M., Munechika, I., and Okada, N. (1997). Molecular evidence from retroposons that whales form a clade within even-toed ungulates. *Nature* **388**: 666–670.
- Siddall, M. (1996). "Random Cladistics," Version 4.0. Virginia Institute of Marine Science, College of William and Mary, Gloucester Point, VA. [anonymous ftp://zoo.toronto.edu/pub]
- Simpson, G. G. (1945). The principles of classification and a classification of mammals. *Bull. Am. Mus. Nat. Hist.* **85**: 1–350.
- Springer, M. S., and Douzery, E. (1996). Secondary structure, conservation of functional sites, and rates of evolution among mammalian mitochondrial 12S rRNA genes based on sequences from placentals, marsupials, and a monotreme. *J. Mol. Evol.* **43**: 357–373.
- Swofford, D. L. (1993). "PAUP: Phylogenetic Analysis Using Parsimony," Version 3.1.1. Illinois Natural History Survey, Champaign, IL.
- Thewissen, J. G. M. (1994). Phylogenetic aspects of cetacean origins: A morphological perspective. *J. Mammal. Evol.* **2**: 157–183.
- Thewissen, J. G. M., and Hussain, S. T. (1993). Origin of underwater hearing in whales. *Nature* **361**: 444–445.
- Viret, J. (1961). Artiodactyla. In "Traité de Paléontologie (Tome VI)" (J. Piveteau, Ed.), pp. 887–1104. Masson, Paris.