University of Alberta

Phylogenetic inference of relationships among caniform carnivores based on complete mitochondrial protein-coding sequences

by

Isabelle Delisle 

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in Systematics and Evolution

Department of Biological Sciences

Edmonton, Alberta

Fall 2004

Your file Votre référence
ISBN: 0-612-95921-X
Our file Notre référence
ISBN: 0-612-95921-X

Acknowledgements

This thesis project could not have been done without the help and support of many. First, I am very grateful to my supervisor Curtis Strobeck for giving me the opportunity to do this research.

I would like to thank my supervisory committee, including Curtis Strobeck, Sean Graham, and Ian Stirling, whose help and support of my work were invaluable. Also, I am indebted to Bruce Rannala, who joined the committee on late notice, and Mark Wilson, and François-Joseph Lapointe, who were members of my examining committee. All of these people have contributed to help improve my research and my thesis.

A very special thanks goes to Renee Polziehn who, from acting as my mentor upon my arrival at the University of Alberta to becoming such a supportive friend, guided me through the whole process.

Corey Davis, with his amazing lab skills and the great discussions we shared, has been instrumental in helping me through tough times in the lab and in making the project gain in quality and in scale.

My other labmates, especially Chris Kyle, Lindsey Carmichael, Tara Gallagher, and Jen Bonneville, have made the many days working in their company cheerful and fun. They've been great friends.

I also wish to thank the National Science and Engineering Research Council (NSERC) for the scholarships that have supported me along the way. Also, the Department of Biological Sciences and the Faculté St-Jean have provided me with a Teaching Assistantship and several teaching opportunities that have helped me not only financially, but have helped develop my skills and passion for education.

I am thankful to Shane Good of the Cleveland Metroparks Zoo, and Wanda Angermeyer, of the Calgary Zoo, for helping me obtain important samples.

Finally, my friends and family, particularly my parents, my partner Guillaume, my sister Cathy, my friends Véronique and Crystelle, Jean, Annie, and Amaya, have all contributed to the completion of this project in innumerable ways, particularly by keeping my stress level in check, thank you so much, merci de tout coeur, je vous aime.

Table of contents

List of Tables

List of Figures

Chapter 1

General introduction

Among the 273 species comprising the order Carnivora can be found dog, wolf, grizzly bear, giant panda, harp seal, wolverine, raccoon, cat, lion, and tiger. These animals all share, or have shared at some time in history, a particular relationship with humans. These relationships may be of commercial, competitive, emotional, or symbolic nature. This certainly has had an influence on the fact that carnivores have been extensively studied (see for example Gittleman, 1989, 1996). Despite this fact, phylogenetic relationships among Carnivora remain unclear. The present chapter reviews the systematics and evolution of carnivores, highlights unresolved issues, and discusses why and how these questions should be revisited today.

The systematic position of the order Carnivora

Definition of Carnivora

The mammalian order Carnivora is primarily defined by a single dental trait. In adult carnivores, the shearing function is restricted to the specialized pair P^4/M_1 (carnassial teeth). Other morphological characters of the carnassial teeth, auditory bulla, circulatory

patterns and claws have also been used to group living carnivores as a monophyletic clade (Flynn et al., 1988).

Evolution of carnivores

True carnivores — species belonging to the order Carnivora — emerged at the beginning of the Tertiary period from a stock of small insectivores that gradually evolved to kill larger prey (Eisenberg, 1981), an event observed after the disappearance of dinosaurian predators. True carnivores of the superfamily Miacoidea, ancestral to present-day carnivores, were then in competition with other carnivorous groups: the order Creodonta, and early ungulates in the Condylarthra. As the vegetation diversified and grasses appeared during the late Eocene/early Oligocene, large herbivores, rodents and lagomorphs evolved, driving a new radiation of carnivores (Ewer, 1973; Werdelin, 1996). True carnivores are thought to have been more successful at this radiation because their dental plan allowed more possibilities for adaptation than that of other carnivorous lineages. Major lines of carnivores, bear-like, dog-like, and cat-like, emerged at that time (Eisenberg, 1981). Most extant species of carnivores, however, evolved more recently from these major lineages, in a radiation that occurred at the Miocene-Pliocene boundary, seven to five million years ago (Martin, 1989).

Closely related groups

The orders Creodonta and Carnivora have a common origin in the root group Palaeonyctidae of the early Cretaceous (Thenius, 1989). Simpson (1945) places both

orders in the superorder Ferae. Fossil and extant members of the two orders share with the order Pholidota (pangolins) a single, possibly synapomorphous, trait: a thin, oblique lamina separating the cerebral and cerebellar fossa (Novacek et al., 1988). The three orders have therefore sometimes been included in Ferae (McKenna and Bell, 1997), a relationship also evidenced by molecular analyses (Shoshani, 1986; Shoshani and McKenna, 1998; Arnason et al., 2002; Amrine-Madsen et al., 2003). In contrast, many other shared characteristics suggest a closer affinity of Pholidota with Edentata (sloths), the two orders forming a clade that diverges at the base of the eutherian tree (Novacek and Wyss, 1986; Novacek et al., 1988).

Higher level relationships of Carnivora and Ferae have also been debated. In 1945, Simpson placed the superorder Ferae in the cohort Ferungulata, a large group that also included the extant order Tubulidentata (aardvark), the superorder Paenungulata comprising Proboscidea (elephants), Hyracoidea (hyraxes) and Sirenia (manatees), the superorder Mesaxonia including Perissodactyla (horses and rhinoceroses) and Artiodactyla (pigs, deer, *etc.*), and fossil allies. Simpson stated that his ferungulate clade was based on fossil evidence, but his conclusions were not supported by later morphological studies (Novacek et al., 1988). Morphology-based phylogenies have placed carnivores in a variety of superordinal clades: as the sister group to Primates (Shoshani, 1986), in a polytomy with Archonta [Scandentia (tree shrews), Dermoptera (flying lemurs), Chiroptera (bats), and Primates], Insectivora (shrews, moles, *etc.*), and Ungulata (McKenna, 1975), or in a large polytomy involving all eutherian orders except

Edentata and Pholidota (Novacek and Wyss, 1986). Such discrepancies indicate that, on the basis of morphological characters alone, relationships of Carnivora cannot be fully resolved (Flynn et al., 1988).

Molecular evidence, however, is providing a better resolved and more consistent solution. Mitochondrial and nuclear DNA data confirm the close affinity of Pholidota and Carnivora, allied to Perissodactyla and Artiodactyla (including Cetacea: whales and dolphins), forming an emended Ferungulata. Ferungulata is further joined by Chiroptera and some groups from the order Insectivora, which has been found to be polyphyletic (Madsen et al., 2001; Murphy et al., 2001a; Murphy et al., 2001b; Arnason et al., 2002; Amrine-Madsen et al., 2003). Internal relationships within this new superordinal clade, named Laurasiatheria, are as yet uncertain.

Phylogenetic relationships among Carnivora

Current taxonomy

In contrast with the uncertainty of their placement within the mammalian tree, the monophyly of carnivores is undisputed. Major clades within Carnivora were described by Flower (1869) on the basis of characteristics of their auditory bulla. Arctoidea lack a longitudinal septum within the bulla; Cynoidea have a low septum in the anterior part of the cavity; and Aeluroidea have a complete septum dividing the bulla into two. Arctoid

families are Ursidae (bears), Mustelidae (weasels), Procyonidae (raccoons), Ailuridae (red panda), Otariidae (sea lions and fur seals), Odobenidae (walruses), and Phocidae (true seals). The latter three aquatic families form the Pinnipedia. The Canidae (dogs) are the only family constituting Cynoidea. Felidae (cats), Viverridae (civets), Hespestidae (mongooses), and Hyaenidae (hyenas) comprise the Aeluroidea (also called Feloidea).

Arctoidea and Cynoidea are thought to be more closely related to each other than they are to Aeluroidea. Carnivora are therefore divided into two suborders: Caniformia grouping the first two, and Feliformia (McKenna and Bell, 1997).

Less recent classifications, however, divided Carnivora into two different suborders, separating terrestrial (Fissipedia) from aquatic (Pinnipedia). Most authors then recognized the arctoid affinity of Pinnipedia (Flower, 1869; Simpson, 1945; Ewer, 1973; Wozencraft, 1989), but thought the distinction between them and fissipeds so important that it justified the non-cladistic classification

Though different name designations and hierarchical levels may be used depending on the authors (see discussion in Simpson, 1945; and Flynn et al., 1988), the main divisions of Carnivora, described in the previous paragraphs and presented in Figure 1-1, are generally agreed upon.

Areas of agreement and points of contention

The higher level phylogeny presented in Figure 1-1 represents a backbone of the phylogeny of Carnivora. Relationships among families comprised in these clades are however contentious. Within the monophyletic infraorder Arctoidea, inter-familial relationships are largely unresolved.

First, the positions of the three pinniped families within Arctoidea are the subject of an ongoing controversy. The question of their ancestral affinities is at the centre of the debate, with Ursidae and Mustelidae being two possible sister groups to Pinnipedia (Flynn et al., 1988). Early twentieth century morphological analyses (Simpson, 1945), and all molecular data (Vrana et al., 1994; Arnason et al., 1995; Ledje and Arnason, 1996a, 1996b; Flynn and Nedbal, 1998; Flynn et al., 2000), support grouping the three pinniped families as a monophyletic group. But another alternative, convincingly put forward by McLaren (1960) and Tedford (1976), is that of a diphyletic origin of pinnipeds, with phocids closely related to mustelids, and otariids and odobenids closer to ursids (for a review, see Bininda Emonds and Russell, 1996). This view remained predominant until the work of Wyss (1987; 1988), who revived the monophyly theory. Authors and approaches however disagree on the placement of the Odobenidae, found as sister to either the Otariidae in the case of older morphological analyses and molecular work, or to the Phocidae in recent revisions of the morphological evidence.

Another current problematic issue is the position of the red panda, *Ailurus fulgens*. This small-sized member of the order Carnivora, named for its resemblance to a cat, has alternately been placed closer to the giant panda (Ewer, 1973), raccoons (Simpson, 1945; O'Brien et al., 1985; Dragoo and Honeycutt, 1997; McKenna and Bell, 1997; Nowak, 1999), bears (Sarich, 1973; Wilson and Reeder, 1993), skunks (Flynn et al., 2000), or dogs (Ledje and Arnason, 1996a). It has also been suggested that it shares a more distant ancestor with a bear and pinniped clade (Vrana et al., 1994), or with musteloids (weasels, skunks, and raccoons, see Schmidt-Kittler, 1981; Flynn and Nedbal, 1998; Bininda-Emonds et al., 1999).

The superfamily Musteloidea, grouping Mustelidae and Procyonidae as sister families, is a recent resolution of a part of the arctoid polytomy (Schmidt-Kittler, 1981, who also included Ailuridae; Miyamoto and Goodman, 1986; Flynn et al., 1988) that is confirmed by DNA sequence analyses (Vrana et al., 1994; Ledje and Arnason, 1996a, 1996b; Dragoo and Honeycutt, 1997).

Skunks are also at the center of a proposed taxonomic rearrangement. Skunks have traditionally been placed as the subfamily Mephitinae among mustelids but are, with increasing support, placed in their own family, the Mephitidae. The main reason for this re-classification is that many recent molecular studies have shown that the family Procyonidae is closer to the rest of the Mustelidae (i.e. *sensu stricto*) than skunks are (Arnason and Widegren, 1986; Wayne et al., 1989; Vrana et al., 1994; Ledje and

Arnason, 1996a, 1996b; Dragoo and Honeycutt, 1997). The elevation of skunks to the family level does not change the overall composition of the Musteloidea but redefines internal relationships within the clade.

Similar unresolved relationships exist within the Feliformia. There is a conflicting distribution of putatively derived morphological characters that support alternative topologies where Hespestidae, Viverridae, or Felidae, are the sister family to the three other feliform families (Flynn et al., 1988). Moreover, family designations are being challenged. Molecular analyses confirm the polyphyly of Viverridae, showing that the genus *Nandinia*, which has a unique cartilaginous bulla (Flynn et al., 1988), is the sister genus to all other members of the suborder Feliformia, and should therefore belong to its own monotypic family (Flynn and Nedbal, 1998). More recent analyses of wider samples of taxa continue to refute traditional familial and subfamilial designations (Gaubert and Veron, 2003; Veron et al., 2004).

More work is therefore needed to resolve the phylogenetic relationships at all levels among Carnivora.

Importance of good phylogenies

Systematic purposes

Since the middle of the twentieth century, the science of systematics has been based, primarily, on the Hennigian principles of cladistics. In this framework, classifications and taxonomies depend on phylogenies, as they must reflect the phylogenetic history of the species. In recent years, endeavours such as the Tree of Life (Maddison, 2001) have renewed the interest in accurate phylogenies.

Framework for studies in other disciplines

Confidence in phylogenies allows hypothesis testing in many different disciplines of biology, such as biogeography, evolution and co-evolution, development, and many others (Moritz and Hillis, 1996). A recent study by Lindenfors et al. (2003) is an example of such an application for a carnivore phylogeny: they used the supertree of Bininda-Emonds et al. (1999) to investigate the evolution of delayed implantation. In such studies, the conclusions reached are highly dependent on the accuracy of the phylogeny used.

Conservation and research decisions

Rarity and uniqueness are often used as criteria when determining conservation and management priorities. The phylogenetic position of a group constitutes one way to evaluate uniqueness. For example, if it is shown that the red panda, *Ailurus fulgens*, is

the only extant representative of a monotypic family (Ailuridae) and therefore not a member of Procyonidae, then an argument could be made to rank its conservation priority higher. Particular phylogenetic positions may also be the impetus for focusing research efforts on a species or a group of species. For example, as traditional members of the family Mustelidae, skunks have often been assumed to share attributes and characteristics found in studies involving other (true) mustelids. Generally perceived as pests, skunks have been the subject of very little direct research. This has been the case, for instance, in population genetics, where mustelids such as the COSEWIC-listed wolverine and the commercially valuable mink have been preferred research subjects. Now, as more phylogenetic analyses suggest that skunks are highly divergent from true mustelids, more direct investigation is required.

Molecular systematics

Molecules as markers

In the late 1960's, the pioneering work of Eck and Dayhoff (1966) and Sarich and Wilson (1967) introduced the use of molecular data to phylogenetic inference and divergence dating. The birth of molecular systematics simultaneously gave rise to controversies: a "molecules vs. morphology" debate emerged, as well as many criticisms of the methods and assumptions used in the new discipline (see Page and Holmes, 1998). Early molecular markers, such as immunological distance and DNA-DNA hybridization, were

criticized as phenetic, and early analytic methods using clustering according to overall similarity were perceived as in opposition to the cladistic movement based on shared derived characters. Discrete data (such as protein and DNA sequences) and methods of analysis that optimize a criterion of fit (such as maximum parsimony, Farris et al., 1970; Fitch, 1971; and maximum likelihood, Felsenstein, 1981) reconciled some of the philosophical aspects of the molecular and traditional approaches.

While molecules were recognized as one line of phylogenetic evidence among others, molecular systematic developed as a distinct field. Findings often very different from traditional views kept fuelling the debate of molecules vs. morphology. In recent years, however, there has been a renewal of interest for the development of more integrative approaches aiming at finding an explanation of all the data available. Using total evidence in a single analysis (Kluge, 1989), or obtaining a consensus between separate analyses of different data (Miyamoto and Fitch, 1995), are two different practices opposed in a perduring philosophical argument (Hillis, 1987; Bull et al., 1993; deQueiroz, 1993; Chippindale and Wiens, 1994; Levasseur and Lapointe, 2001). Beyond the debate, new methods aiming at taking all of the data into consideration have been developed, namely supermatrices (Gatesy et al., 2002), and supertrees (Sanderson et al., 1998; Bininda-Emonds et al., 1999).

Recent advances

Technical progress has made the use of molecular markers very appealing to the systematist. First, the polymerase chain reaction technique (Kleppe et al., 1971; Mullis and Faloona, 1987; Ochman et al., 1988; Saiki et al., 1988), jointly with automated cycle sequencing (the dideoxynucleotides chain termination method of Sanger et al., 1977; using fluorescently labelled terminators, Jett et al., 1989; Davis et al., 1991; Harding and Keller, 1992), now makes the collection of a large amount of homologous sequence data from a large number of taxa possible. These methods allow the simultaneous application of two approaches to resolving difficult problems in phylogeny: adding more characters (Mindell and Thacker, 1996; Swofford et al., 1996; Soltis et al., 1998), and increasing taxon sampling (Hillis, 1996; 1998; Rannala et al., 1998).

In addition to improved methods of data collection, methods of analysis have become more complex in their use of more realistic models of DNA evolution. These models can be implemented within a maximum likelihood framework, albeit at the cost of very intensive computation time. Faster computers are gradually making the implementation of these models more feasible, as are new tree inference methods using a Bayesian framework (Rannala and Yang, 1996; Yang and Rannala, 1997; Larget and Simon, 1999). The Bayesian approach is performed in a maximum likelihood framework but does not aim at finding the best tree but rather uses a Markov chain Monte Carlo process to sample the posterior distribution of parameters (including topology) describing a given data set.

Finally, new statistical tests target a longstanding problem of phylogenetic inference: comparing the result (the best tree) to different hypotheses (alternative topologies), and evaluating the significance of the result. Likelihood-based tests of topologies devised by Shimodaira and Hasegawa (1999; SH test) and Swofford et al. (1996; SOWH test) evaluate the significance of the value δ, the difference in log-likelihood between the best tree and an alternative tree (or possibly more than one, in the SH test). A main difference between these tests is how they determine the null distribution of the test statistic δ. The SH test is nonparametric and uses pseudo-replicates sampled with replacement from the data. The SOWH test is parametric and uses estimated parameters of the DNA evolution model that best fits the data to simulate new data sets under the null hypothesis. These tests are reviewed in Goldman et al. (2000)

Likelihood-based tests, especially the SOWH, can be very computationally intense, involving a primary search for the best tree and then multiple rounds of evaluation of simulated data sets. The Bayesian approach has a double advantage as a faster tree searching method, but also in the fact that probability estimates are compiled in the same process. The posterior probability of a tree is the number of times it is visited by the Markov chain over the total number of trees sampled. It has been suggested that the Bayesian posterior probability of a tree is a direct measure of the probability of it being the true tree given the observed data, and, as such, can be used as a statistical measure (Rannala and Yang, 1996; Larget and Simon, 1999; Lewis, 2001).

Shortcomings

Although offering several advantages, the use of molecular data for phylogenetic inference also has potential caveats. The main problem related to the use of genes as phylogenetic markers is that while a gene contains multiple characters used to reconstruct a phylogeny, it (typically) constitutes a single hereditary unit that may or may not reflect the evolutionary history of the group under study (Pamilo and Nei, 1988). A potential incongruence between a gene tree and a species tree may stem from different sources of error such as gene duplication or loss, deep coalescence, and lateral gene transfer (Pamilo and Nei, 1988; Moore, 1995; Brower et al., 1996). Mitochondrial (or chloroplastic) DNA is often used as a phylogenetic marker as it comprises several genes and may minimize some of the noted problems (Moore, 1995; but see also Hoelzer, 1997; and Moore, 1997); one must however be aware that it still represents a single linkage group. As many different data sets become available for the same taxa, individual gene trees will represent individual lines of evidence that can be compared to, and analyzed alongside, others to provide a better picture of the species' genealogy.

Objectives of this study

The general goal of this study was to further resolve relationships among the caniform families of the Carnivora. First, methods of molecular data collection were optimized in a new protocol to rapidly and efficiently obtain sequence of the mitochondrial genome

from numerous species of the order Carnivora (Chapter 2). The method developed was used to obtain complete mitochondrial protein-coding sequences from 32 species (Chapter 3 and Davis et al., 2004). Combined to six previously published sequences, the new data set was used in a phylogenetic analysis of the relationships within Caniformia (Chapter 3). In Chapter 4, different measures of support and statistical tests of topologies were used to further investigate the interesting result of a sister relationship between the red panda and the skunk, obtained in the general analysis. Finally, this work allows a reflection on the potential and limits of mitochondrial DNA sequences in resolving the phylogeny of the order Carnivora.

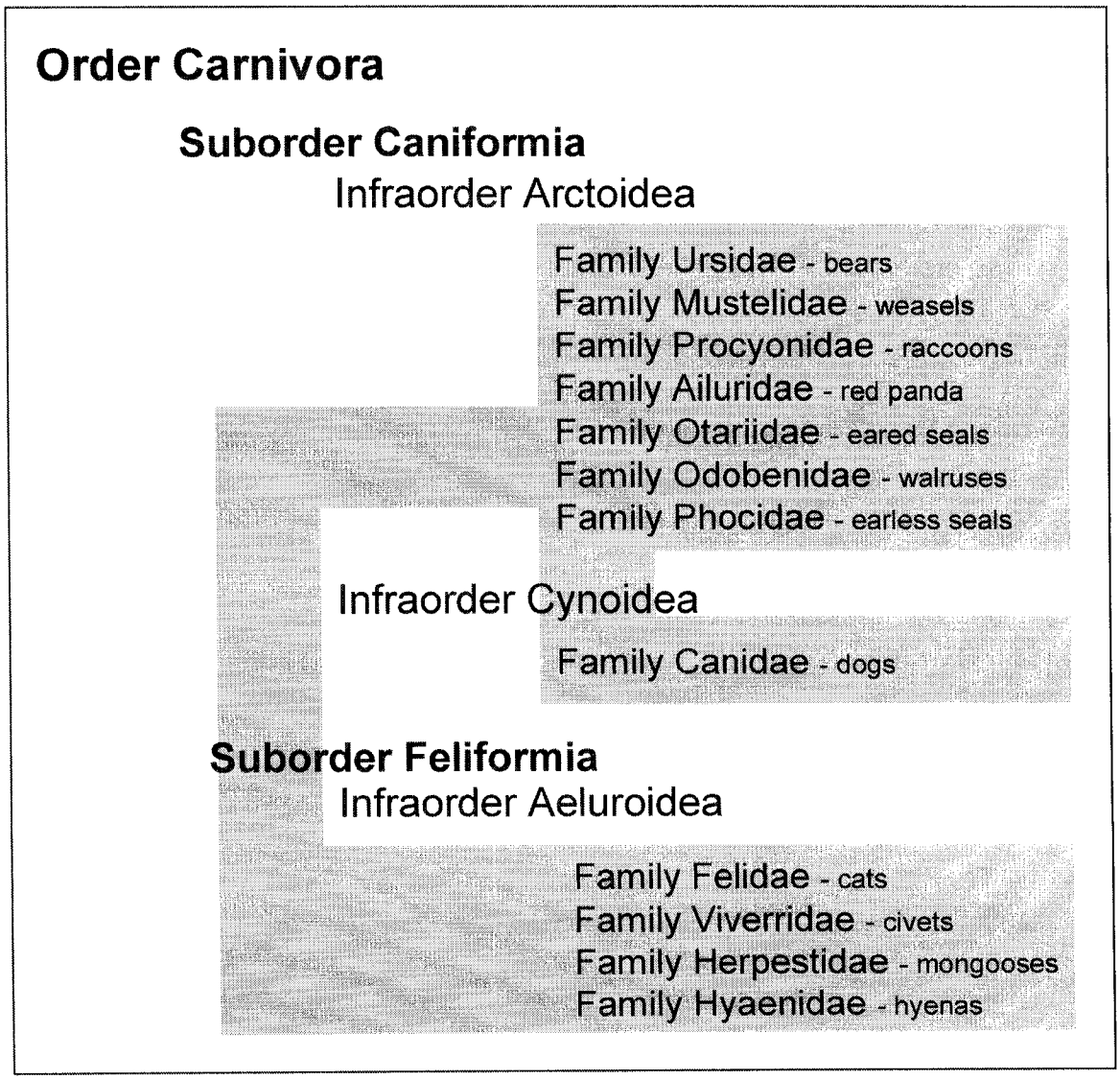


Figure 1-1 Higher level taxonomy and phylogenetic relationships of the order Carnivora.

Bibliography

Amrine-Madsen, H., K.-P. Koepfli, R. K. Wayne, and M. S. Springer. 2003. A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. Molecular Phylogenetics and Evolution 28:225-240.

Arnason, U., J. A. Adegoke, K. Bodin, et al. 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. Proceedings of the National Academy of Sciences USA 99:8151-8156.

Arnason, U., K. Bodin, A. Gullberg, C. Ledje, and S. Mouchaty. 1995. A molecular view of pinniped relationships with particular emphasis on the true seals. Journal of Molecular Evolution 40:78-85.

Arnason, U., and B. Widegren. 1986. Pinniped phylogeny enlightened by molecular hybridizations using highly repetitive DNA. Molecular Biology and Evolution 3:356-365.

Bininda Emonds, O. R. P., and A. P. Russell. 1996. A morphological perspective on the phylogenetic relationships of the extant phocid seals (Mammalia: Carnivora: Phocidae). Bonner zoologische Monographien 41:1-256.

Bininda-Emonds, O. R. P., J. L. Gittleman, and A. Purvis. 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). Biological Reviews Of The Cambridge Philosophical Society 74:143-175.

Brower, A. V. Z., R. DeSalle, and A. Vogler. 1996. Gene trees, species trees, and systematics: a cladistic perpective. Annual Review of Ecology and Systematics 27:423-450.

Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford, and P. J. Waddell. 1993. Partitioning and combining data in phylogenetic analysis. Systematic Biology 42:384-397.

Chippindale, P. T., and J. J. Wiens. 1994. Weighting, partitioning, and combining characters in phylogenetic studies. Systematic Biology 43:278-287.

Davis, C. S., I. Delisle, D. B. Siniff, I. Stirling, and C. Strobeck. 2004. A phylogeny of the extant Phocidae inferred from complete mitochondrial DNA coding regions. Molecular Phylogenetics and Evolution (In press).

Davis, L. M., F. R. Fairfield, C. A. Harger, et al. 1991. Rapid DNA sequencing based upon single molecule detection. Genetic Analysis-Biomolecular Engineering 8:1-7.

deQueiroz, A. 1993. For consensus (sometimes). Systematic Biology 42:368-372.

Dragoo, J. W., and R. L. Honeycutt. 1997. Systematics of mustelid-like carnivores. Journal of Mammalogy 78:426-443.

Eck, R. V., and M. O. Dayhoff. 1966. Atlas of protein sequence and structure. National Biomedical Research Foundation.

Eisenberg, J. F. 1981. The Mammalian Radiations. University of Chicago Press, Chicago.

Ewer, R. F. 1973. The Carnivores. Weidenfeld and Nicolson, London.

Farris, J. S., A. G. Kluge, and M. J. Eckardt. 1970. A numerical approach to phylogenetic systematics. Systematic Zoology 19:172-189.

Felsenstein, J. 1981. Evolutionary trees from DNA-sequences - a maximum-likelihood approach. Journal of Molecular Evolution 17:368-376.

Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Systematic Zoology 20:406-416.

Flower, W. H. 1869. On the value of the characters of the base of the cranium in the classification of the order Carnivora and the systematic position of *Bassaris* and other disputed forms. Proceedings of the Zoological Society of London 1869:4-37.

Flynn, J. J., and M. A. Nedbal. 1998. Phylogeny of the Carnivora (Mammalia): Congruence vs incompatibility among multiple data sets. Molecular Phylogenetics and Evolution 9:414-426.

Flynn, J. J., M. A. Nedbal, J. W. Dragoo, and R. L. Honeycutt. 2000. Whence the red panda? Molecular Phylogenetics and Evolution 17:190-199.

Flynn, J. J., N. A. Neff, and R. H. Tedford. 1988. Phylogeny of the Carnivora. Pages 73-115 *in* The phylogeny and classification of the tetrapods. Vol. 2: Mammals (M. J. Benton, ed.). Clarendon Press, Oxford.

Gatesy, J., C. Matthee, R. DeSalle, and C. Hayashi. 2002. Resolution of a supertree/supermatrix paradox. Systematic Biology 51:652-664.

Gaubert, P., and G. Veron. 2003. Exhaustive sample set among Viverridae reveals the sister-group of felids: the linsangs as a case of extreme morphological convergence within Feliformia. Proceedings of the Royal Society of London Series B-Biological Sciences 270:2523-2530.

Gittleman, J. L. (ed) 1989. Carnivore behavior, ecology and evolution. Volume 1. Cornell University Press, Ithaca, NY.

Gittleman, J. L. (ed) 1996. Carnivore behavior, ecology and evolution. Volume 2. Cornell University Press, Ithaca, NY.

Goldman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49:652-670.

Harding, J. D., and R. A. Keller. 1992. Single-molecule detection as an approach to rapid DNA sequencing. Trends in Biotechnology 10:55-57.

Hillis, D. M. 1987. Molecular versus morphological approaches to systematics. Annual Review of Ecology and Systematics 18:23-42.

Hillis, D. M. 1996. Inferring complex phylogenies. Nature 383:130-131.

Hillis, D. M. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Systematic Biology 47:3-8.

Hoelzer, G. A. 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-genes trees revisited. Evolution 51:622-626.

Jett, J. H., R. A. Keller, J. C. Martin, N. K. Seitzinger, and E. B. Shera. 1989. Single molecule detection in flowing sample streams as an approach to DNA sequencing. Abstracts of Papers of the American Chemical Society 198:41.

Kleppe, K., E. Ohtsuka, R. Kleppe, I. Molineux, and H. G. Khorana. 1971. Studies on polynucleotides XCVI. Repair replication of short synthetic DNAs as catalyzed by DNA polymerases. Journal of Molecular Biology 56:341-361.

Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). Systematic Biology 38:7-25.

Larget, B., and D. Simon. 1999. Markov chains Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Molecular Biology and Evolution 16:750-759.

Ledje, C., and U. Arnason. 1996a. Phylogenetic analyses of complete cytochrome b genes of the order Carnivora with particular emphasis on the Caniformia. Journal of Molecular Evolution 42:135-144.

Ledje, C., and U. Arnason. 1996b. Phylogenetic relationships within caniform carnivores based on analyses of the mitochondrial 12S rRNA gene. Journal of Molecular Evolution 43:641-649.

Levasseur, C., and F.-J. Lapointe. 2001. War and peace in phylogenetics: a rejoinder on total evidence and consensus. Pages 881-891 *in* Systematic Biology Taylor & Francis Ltd.

Lewis, P. O. 2001. Phylogenetic systematics turns over a new leaf. Trends in Ecology & Evolution 16:30-37.

Lindenfors, P., L. Dalen, and A. Angerbjorn. 2003. The monophyletic origin of delayed implantation in carnivores and its implications. Evolution 57:1952-1956.

Maddison, D. R. 2001. The Tree of life web project. http://tolweb.org/tree/phylogeny.html.

Madsen, O., M. Scally, C. J. Douady, et al. 2001. Parallel adaptive radiations in two major clades of placental mammals. Nature 409:610-614.

Martin, L. D. 1989. Fossil history of the terrestrial Carnivora. Pages 536-568 *in* Carnivore behavior, ecology and evolution. Vol. 1 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

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. S. Szalay, eds.). Plenum Press, New York.

McKenna, M. C., and S. K. Bell. 1997. Classification of mammals : above the species level. Columbia University Press, New York.

McLaren, I. A. 1960. Are the Pinnipedia biphyletic? Systematic Zoology 9:18-28.

Mindell, D. P., and C. E. Thacker. 1996. Rates of molecular evolution: phylogenetic issues and applications. Annual Review of Ecology and Systematics 27:279-303.

Miyamoto, M. M., and W. M. Fitch. 1995. Testing phylogenies and phylogenetic methods with congruence. Systematic Biology 44:64-76.

Miyamoto, M. M., and M. Goodman. 1986. Biomolecular systematics of eutherian mammals: phylogenetic patterns and classification. Systematic Zoology 35:230-240.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-genes trees. Evolution 49:718-726.

Moore, W. S. 1997. Mitochondrial-gene trees versus nuclear-gene trees, a reply to Hoelzer. Evolution 51:627-629.

Moritz, C., and D. M. Hillis. 1996. Molecular systematics: Context and controversies. Pages 1-13 *in* Molecular Systematics (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland MASS.

Mullis, K. B., and F. A. Faloona. 1987. Specific synthesis of DNA invitro via a polymerase-catalyzed chain-reaction. Methods in Enzymology 155:335-350.

Murphy, W. J., E. Eizirik, W. E. Johnson, Y. P. Zhang, O. A. Ryder, and S. J. O' Brien. 2001a. Molecular phylogenetics and the origins of placental mammals. Nature 409:614-618.

Murphy, W. J., E. Eizirik, S. J. O'Brien, et al. 2001b. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294:2348-2351.

Novacek, M. J., and A. R. Wyss. 1986. Higher-level relationships of the recent eutherian orders: morphological evidence. Cladistics 2:257-287.

Novacek, M. J., A. R. Wyss, and M. C. McKenna. 1988. The major groups of eutherian mammals. Pages 31-71 *in* The phylogeny and classification of the tetrapods. Vol. 2: Mammals (M. J. Benton, ed.). Clarendon Press, Oxford.

Nowak, R. M. 1999. Walker's mammals of the world, 6th edition. Johns Hopkins University Press, Baltimore.

O'Brien, S. J., W. G. Nash, D. E. Wildt, M. E. Bush, and R. E. Benveniste. 1985. A molecular solution to the riddle of the giant panda's phylogeny. Nature 317:140-144.

Ochman, H., A. S. Gerber, and D. L. Hartl. 1988. Genetic applications of an inverse polymerase chain-reaction. Genetics 120:621-623.

Page, R. D. M., and E. C. Holmes. 1998. Molecular evolution, a phylogenetic approach. Blackwell Science.

Pamilo, P., and M. Nei. 1988. Relationships between gene trees and species trees. Molecular Biology and Evolution 5:568-583.

Rannala, B., J. P. Huelsenbeck, Z. Yang, and R. Nielsen. 1998. Taxon sampling and the accuracy of large phylogenies. Systematic Biology 47:702-710.

Rannala, B., and Z. H. Yang. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. Journal of Molecular Evolution 43:304-311.

Saiki, R. K., D. H. Gelfand, S. Stoffel, et al. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. Science 239:487-491.

Sanderson, M. J., A. Purvis, and C. Henze. 1998. Phylogenetic supertrees: assembling the trees of life. Trends in Ecology & Evolution 13:105-109.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74:5463-5467.

Sarich, V. M. 1973. The giant panda is a bear. Nature 245:218-220.

Sarich, V. M., and A. C. Wilson. 1967. Immunological time scale for hominid evolution. Science 158:1200-1203.

Schmidt-Kittler, N. 1981. Zur Stammesgeschichte der marderverwandten Raubtiergruppen (Musteloidea, Carnivora). Eclogae Geologicae Helvetiae 74:753-801.

Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16:1114-1116.

Shoshani, J. 1986. Mammalian phylogeny: a comparison of morphological and molecular results. Molecular Biology and Evolution 3:222-242.

Shoshani, J., and M. McKenna. 1998. Higher taxonomic relationships among extant mammals based on morphology, with selected comparisons of results from molecular data. Molecular Phylogenetics and Evolution 9:572-584.

Simpson, G. G. 1945. The principles of classification and a classification of mammals. The American Museum of Natural History. Bulletins, 85, New York.

Soltis, D. E., P. S. Soltis, M. E. Mort, M. W. Chase, V. Savolainen, S. B. Hoot, and C. M. Morton. 1998. Inferring complex phylogenies using parsimony: An empirical approach using three large DNA data sets for angiosperms. Systematic Biology 47:32-42.

Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference. Pages 407-514 *in* Molecular Systematics (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland MASS.

Tedford, R. H. 1976. Relationship of pinnipeds to other carnivores (Mammalia). Systematic Zoology 25:363-374.

Thenius, E. 1989. Carnivores: Phylogeny. Pages 370-383 *in* Grzimek's Encyclopedia of Mammals. Vol. 3 (B. Grzimek, ed.). McGraw-Hill, New-York.

Veron, G., M. Colyn, A. E. Dunham, P. Taylor, and P. Gaubert. 2004. Molecular systematics and origin of sociality in mongooses (Herpestidae, Carnivora). Molecular Phylogenetics and Evolution 30:582-598.

Vrana, P. B., M. C. Milinkovitch, J. R. Powell, and W. C. Wheeler. 1994. Higher level relationships of the arctoid Carnivora based on sequence data and "total evidence". Molecular Phylogenetics and Evolution 3:47-58.

Wayne, R. K., R. E. Benveniste, D. N. Janczewski, and S. J. O'Brien. 1989. Molecular and biochemical evolution of the Carnivora. Pages 465-494 *in* Carnivore behavior, ecology and evolution. Vol. 1 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Werdelin, L. 1996. Carnivoran ecomorphology: A phylogenetic perspective. Pages 582-624 *in* Carnivore behavior, ecology, and evolution. Vol. 2 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Wilson, D. E., and D. M. Reeder. 1993. Mammal species of the world : a taxonomic and geographic reference, 2nd edition. Smithsonian Institution Press, Washington, D.C.

Wozencraft, W. C. 1989. Appendix - Classification of the recent Carnivora. Pages 569-593 *in* Carnivore behavior, ecology and evolution. Vol. 1 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Wyss, A. R. 1987. The walrus auditory region and the monophyly of pinnipeds. American Museum Novitates 2871:1-13.

Wyss, A. R. 1988. Evidence from flipper structure for a single origin of pinnipeds. Nature 334:427-428.

Yang, Z., and B. Rannala. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. Molecular Biology and Evolution 14:717-724.

Chapter 2

Conserved primers for rapid sequencing of the complete mitochondrial genome of carnivores, applied to three species of bears[1]

Introduction

Phylogenetic relationships of the order Carnivora have been studied extensively. However, depending on the type of data examined, the extent of species sampling, and method of analysis used, carnivores have been placed in nearly every major position throughout the eutherian phylogenetic tree (for a review, see Novacek, 1992). The integrity of the order itself has, however, remained intact. Undisputedly monophyletic, the order Carnivora nonetheless constitutes a very adaptable and heterogeneous group (Wayne et al., 1989) whose evolution has been marked by several events of parallel or convergent evolution, and rapid radiation (Ewer, 1973; Martin, 1989). Consequently, phylogenetic relationships among and within many carnivore families are contentious. For example, phylogenetic relationships are unresolved between the families Procyonidae (raccoons), Ailuridae (red panda), Mustelidae (weasels and skunks), Ursidae (bears), Phocidae (true or earless seals), Otariidae (eared seals) and Odobenidae (walrus). The latter three families are grouped together as Pinnipedia, or aquatic carnivores. Within

[1] A version of this chapter is published in Molecular Biology and Evolution (2002), Volume 19(3), pages 357-361.

Pinnipedia, the closer affinity of Odobenidae to either Phocidae or Otariidae has long been debated. Also, the monophyly of the family Mustelidae has been challenged by molecular studies, where true Mustelidae and Mephitidae (skunks) have been shown to belong to different branches within the superfamily Musteloidea, a group that also includes Procyonidae, and perhaps Ailuridae (Arnason and Widegren, 1986; Ledje and Arnason, 1996; Dragoo and Honeycutt, 1997; Flynn et al., 2000). While all of the above illustrate the uncertainty of inter-familial relationships among carnivores, there is also uncertainty below the family level. The case of the phylogeny of the family Ursidae is a typical example where the integrity of the family is recognized but internal relationships between species are unresolved. The giant panda (Ailuropodinae) and the spectacled bear (Tremarctinae) lineages are known to have split towards the base of the ursid clade. The remaining six species of bears, grouped in the subfamily Ursinae, are thought to have diverged at the beginning of the Pliocene, through such a rapid radiation event (Thenius, 1990) that the order of species divergence is difficult to determine. Hence, the speciation event that led to extant ursine bears (American black bear, *Ursus americanus*; brown bear, *U. arctos*; polar bear, *U. maritimus*; Asiatic black bear, *U. thibetanus*; sun bear, *U. malayanus*; and sloth bear, *U. ursinus*) is usually represented as a polytomy.

The order of species divergence occurring during rapid radiations remains unresolved largely because available data contain very few informative changes essential to infer the correct phylogenetic tree. Given a consistent method of phylogenetic inference, greater resolving power can be achieved by using larger character sets to provide a sufficient

number of informative characters (Mindell and Thacker, 1996; Swofford et al., 1996; Soltis et al., 1998). In molecular studies, this implies obtaining longer DNA sequences. Moreover, it has been shown that the combination of multiple genes in a single large data set has the potential to raise a weak phylogenetic signal above the noise level (Bull et al., 1993). Combining these two advantages, complete mitochondrial (mtDNA) sequences (a total of approximately 17 kb, and 13 protein-coding genes) offer great potential for recovering phylogenies.

Materials and methods

Sequencing strategy

With the specific goal of resolving the ursine polytomy, and keeping in mind the variety of questions related to carnivore phylogeny that have yet to be resolved, a new approach to complete mitochondrial genome sequencing is described here for obtaining complete mitochondrial sequences from closely related carnivore species. A series of primers was designed based on conserved regions identified from an alignment of published complete mitochondrial genomes from carnivores [harbour seal, *Phoca vitulina*, NC_001325, (Arnason and Johnsson, 1992); grey seal, *Halichoerus grypus*, genome NC_001602, (Arnason et al., 1993); domestic cat, *Felis catus*, NC_001700, (Lopez et al., 1996); and dog, *Canis familiaris*, NC_002008, (Kim et al., 1998)]. Eleven primer pairs were designed for the amplification of fragments covering the entire mitochondrial genome.

Fragments are 1.4 to 3.2 kb in size and always span more than one gene, in an attempt to limit the risk of amplifying nuclear copies while increasing the chance of detecting them. The primer design also ensures that there is sufficient overlap of the fragments in order to obtain the sequence of the primer sites and their flanking nucleotides. Also, three to twelve internal primers were designed to complete the sequence of each fragment in both directions. The program Oligo™ (Rychlik, 1997) was used to evaluate primer characteristics and combinability in amplifying pairs. A list of all the primers used to sequence the complete mitochondrial genome from a bear species is presented in Appendix A.

Sequencing conditions

Conditions for amplification were 1X PCR buffer (50 mM KCl, 10 mM Tris-HCl pH 8.8, 0.1% Triton and 0.16 mg/mL bovine serum albumin), 2.5 mM $MgCl_2$, 0.16 mM dNTPs, 0.8 μM each primer, approximately 1-5 U of *Taq* DNA polymerase in a total reaction volume of 100 μL containing approximately 250 ng of DNA. PCR cycles were 3 minutes at 94 °C followed by 35-45 cycles of 1 min 94 °C, 1 min at 50 °C and 1.5 min at 72 °C. Amplifying and internal primers were used for direct sequencing of PCR products.

Application to three bear species

This approach was used to sequence the mitochondrial genome of the three North American ursine species: *U. americanus*, *U. arctos*, and *U. maritimus*. The total genomic DNA of a representative of each species was extracted from tissue samples using a

DNeasy Tissue Extraction Kit (QIAgen) and was used directly in PCR, as described above. Resulting fragments were electrophoresed on 0.8-1% agarose gels, bands cut out and products purified using the QIAquick Gel Extraction Kit (QIAgen). Direct sequencing was then performed using Applied Biosystems dRhodamine or BigDye (v. 3.0 or 3.1) sequencing kits, following the manufacturer recommended protocols. Unincorporated dideoxyribonucleotide removal was performed using the DyeEx Kit (QIAgen). Sequencing products were resolved on an ABI 377 automated DNA sequencer. Sequences were analyzed using Sequence Analysis v. 3.4.1, and assembled using Sequence Navigator v 1.0.1 (both programs from Applied Biosystems).

Phylogenetic analysis I – Cytochrome b gene of bears

A preliminary phylogenetic analysis was performed to verify the position of the new sequences and to evaluate the potential of the proposed sequencing strategy for future work. To facilitate comparisons with previous studies, the weighted parsimony method used by Talbot and Shields (1996), as well as the maximum likelihood approach using the HKY (Hasegawa et al., 1985) + gamma (Yang, 1993) model used by Waits et al. (1999) were applied to a data set comprising the cytochrome b sequences from Talbot and Shields (1996), an additional Asiatic black bear sequence (Matsuhashi et al., 1999), and the three new bear sequences. Domestic dog, grey seal, and harbour seal sequences were used as outgroups.

Phylogenetic analysis II – Complete genome

The complete mitochondrial sequences were used in an analysis along with previously published sequences from other carnivores (as listed above). Various analysis strategies were used in order to evaluate the impact of the method used on tree inference. Maximum parsimony (MP; Farris et al., 1970; Fitch, 1971), maximum likelihood (ML; Felsenstein, 1981) and neighbour-joining (NJ; Saitou and Nei, 1987) methods were used. Models used were HKY (Hasegawa et al., 1985) with and without a gamma distribution for ML searches, and uncorrected *p* and LogDet distances for NJ. Phylogenetic analyses were performed using various data subsets from the complete genome nucleotide sequence: the whole coding region, the 12 H-strand protein-coding genes with the two ribosomal RNA genes, the 12 H-strand protein-coding genes only, and the amino acid translation of the 12 genes (used in MP and NJ analyses). Artiodactyl and perissodactyl sequences [*Bos taurus*, V00654, (Anderson et al., 1982); and *Equus caballus*, X79547, (Xu and Arnason, 1994)] were used to root the trees. All analyses were performed using PAUP* (Swofford, 2003).

Results and Discussion

The sequences obtained for the three North American ursine species *U. americanus*, *U. arctos*, and *U. maritimus* show that the organization of the bear mitochondrial genome conforms to that of other mammalian species. The length of the mtDNA of the black,

brown, and polar bear individuals used in this study is 16 841 bp, 17 020 bp, and 17 017 bp. These numbers are approximate however, as a multiple 10 bp tandem repeat region located in the control region of the mitochondrial genome is too long to be sequenced directly. This region of the mitochondrial molecule, reported to show high levels of heteroplasmy by Hoelzel et al. (1994), is approximately 350 bp long in the black bear individual sequenced, and 600 bp in brown and polar bears. The sequences obtained from separate fragments matched each other in regions of overlap throughout the circular molecule. Moreover, gene sequences showed no occurrence of premature stop codons and aligned well with previously published homologous sequences. These two results imply that the sequences obtained are not from pseudogene copies. These complete mitochondrial sequences can be accessed through Genbank (Accession numbers NC_003426, NC_003427, NC_003428).

To confirm the ursine identity and the phylogenetic position of the individuals sequenced, the cytochrome b gene from the newly sequenced complete mitochondrial genomes was placed in a data matrix along with previously published cytochrome b sequences from the eight extant species of bears. The three newly sequenced individuals grouped with their conspecifics. However, the parsimony and likelihood approaches resulted in two different trees. This is not surprising since previous studies have shown that such short sequences are insufficient to resolve the ursine polytomy with sufficient certainty (Talbot and Shields, 1996; Waits et al., 1999). At this point in my study, my purpose is not to thoroughly evaluate bear phylogeny but mainly to validate the approach, first by

confirming that it can provide reliable sequence data, and then by showing that the addition of data to the ursid and carnivore data set has the potential to result in a better resolved tree. A few findings are relevant to this objective. The cytochrome b sequence from the black bear used in this study is representative of the western haplotype previously identified in several independent studies (Cronin et al., 1991; Paetkau and Strobeck, 1996; Byun et al., 1997; Wooding and Ward, 1997). A "western" North American black bear, and an Asiatic black bear from Japan (Matsuhashi et al., 1999) were therefore added to the original study of Talbot and Shields (1996), who used an "eastern" black bear and an Asiatic black bear of an unknown origin. The distance analysis showed that the two divergent North American black bears differed by an uncorrected *p* distance of 0.038 and the two Asiatic black bears by 0.047. These distances were the largest among representatives of the same species included here, and larger than that separating the two seal species used as outgroup. These additions to the data set improved the bootstrap support of the clade comprising the two black bear species from 75% to 98%. This analysis suggests that the addition of taxa, here divergent conspecifics can substantially improve confidence in the tree topology. A few representatives from each ursid species, corresponding to subspecific clades, should be helpful in resolving all of the nodes of the ursine polytomy, thereby emphasizing the need for an efficient method to obtain sequence data from very closely related taxa.

Maximum parsimony, maximum likelihood, and neighbour-joining methods for estimating trees from the complete sequence data sets all resulted in the same tree

topology. This corroborates the conclusions of Russo et al. (1996) showing that when many mitochondrial genes are used, all methods appear to converge towards a single tree. The tree topology obtained corresponds to what is currently accepted as the "putative true tree" (see Honeycutt and Adkins, 1993; Cao et al., 1994; Kuma and Miyata, 1994; Janke et al., 1997; Cao et al., 1998). This analysis suggests that extensive sequence from the mitochondrial genome provides reliable data and sufficient resolution for building mammalian phylogenies.

The sequencing strategy developed here offers many advantages when the ultimate goal is to obtain homologous sequences from numerous individuals of the same or closely related species. A complete mitochondrial sequence can be obtained in just 73 sequencing reactions, a number at least ten times less than procedures that involve cloning of random fragments, where a fivefold coverage of the genome is necessary to approach a complete sequence (Fleischmann et al., 1995). Moreover, the sequencing procedure is completely PCR-based, therefore bypassing purification of mitochondria, DNA restriction, and cloning steps common to other methods. The reliance on PCR also has its potential downfalls. First, there is the possibility of amplification of pseudogenes. Users of this technique must be vigilant when analyzing the data and look out for telltale signs such as highly divergent sequences, misaligning fragments, and frequent apparent frameshift and nonsense mutations. PCR-induced errors could also occur. This should lead to a mixture of fragments differing at the site of error, and should therefore be detectable at the sequencing step. A few replicates of the amplification procedure are

generally necessary to obtain enough of the template DNA for the sequencing of each fragment: overlapping sequences from different replicates can serve as controls in detecting errors.

Finally, the eleven primer pairs used for the amplification of mitochondrial fragments are located in highly conserved regions so as to transfer easily to other carnivores. They have indeed been used for amplification of DNA from numerous carnivore species (see Chapter 3). While the majority of the amplifying and internal primers are expected to be widely useful, some of the internal primers presented here will show unavoidable mismatches that may lead to poorer PCR performance. It is therefore likely that a few new primers will be required to complete the sequence of any given carnivore mitochondrial genome. On the other hand, as more species are sequenced, a list of alternative primers will become available to facilitate further sequencing. This opens the door to the use of complete mtDNA sequences in various phylogenetic analyses at the order, family, subfamily and species levels among carnivores, and should certainly lead to a better understanding of carnivore phylogeny.

Bibliography

Anderson, S., M. H. de Bruijn, A. R. Coulson, I. C. Eperon, F. Sanger, and I. G. Young. 1982. Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. Journal of Molecular Biology 156:683-717.

Arnason, U., A. Gullberg, E. Johnsson, and C. Ledje. 1993. The nucleotide sequence of the mitochondrial molecule of the grey seal, *Halichoerus grypus*, and a comparison with mitochondrial sequences of other true seals. Journal of Molecular Evolution 43:41-45.

Arnason, U., and E. Johnsson. 1992. The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*. Journal of Molecular Evolution 34:493-505.

Arnason, U., and B. Widegren. 1986. Pinniped phylogeny enlightened by molecular hybridizations using highly repetitive DNA. Molecular Biology and Evolution 3:356-365.

Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford, and P. J. Waddell. 1993. Partitioning and combining data in phylogenetic analysis. Systematic Biology 42:384-397.

Byun, S. A., B. F. Koop, and T. E. Reimchen. 1997. North American black bear mtDNA phylogeography: implications for morphology and the Haida Gwaii glacial refugium controversy. Evolution 51:1647-1653.

Cao, Y., J. Adachi, A. Janke, S. Paabo, and M. Hasegawa. 1994. Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. Journal of Molecular Evolution 39:519-527.

Cao, Y., A. Janke, P. J. Waddell, et al. 1998. Conflict among individual mitochondrial proteins in resolving the phylogeny of eutherian orders. Journal of Molecular Evolution 47:307-322.

Cronin, M. A., S. C. Amstrup, G. W. Garner, and E. R. Vyse. 1991. Interspecific and intraspecific mitochondrial variation in North American bears (*Ursus*). Canadian Journal of Zoology 69:2985-2992.

Dragoo, J. W., and R. L. Honeycutt. 1997. Systematics of mustelid-like carnivores. Journal of Mammalogy 78:426-443.

Ewer, R. F. 1973. The Carnivores. Weidenfeld and Nicolson, London.

Farris, J. S., A. G. Kluge, and M. J. Eckardt. 1970. A numerical approach to phylogenetic systematics. Systematic Zoology 19:172-189.

Felsenstein, J. 1981. Evolutionary trees from DNA-sequences - a maximum-likelihood approach. Journal of Molecular Evolution 17:368-376.

Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Systematic Zoology 20:406-416.

Fleischmann, R. D., M. D. Adams, O. White, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496-498, 507-512.

Flynn, J. J., M. A. Nedbal, J. W. Dragoo, and R. L. Honeycutt. 2000. Whence the red panda? Molecular Phylogenetics and Evolution 17:190-199.

Hasegawa, M., H. Kishino, and T. A. Yano. 1985. Dating of the Human Ape Splitting by a Molecular Clock of Mitochondrial-DNA. Journal of Molecular Evolution 22:160-174.

Hoelzel, A. R., J. V. Lopez, G. A. Dover, and S. J. O'Brien. 1994. Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. Journal of Molecular Evolution 39:191-199.

Honeycutt, R. L., and R. M. Adkins. 1993. Higher level systematics of eutherian mammals: an assessment of molecular characters and phylogenetic hypotheses. Annual Review of Ecology and Systematics 24:279-305.

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. Proceedings of the National Academy of Sciences USA 94:1276-1281.

Kim, K. S., S. E. Lee, H. W. Jeong, and J. H. Ha. 1998. The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. Molecular Phylogenetics and Evolution 10:210-220.

Kuma, K., and T. Miyata. 1994. Mammalian phylogeny inferred from multiple protein data. Japanese Journal of Genetics 69:555-566.

Ledje, C., and U. Arnason. 1996. Phylogenetic relationships within caniform carnivores based on analyses of the mitochondrial 12S rRNA gene. Journal of Molecular Evolution 43:641-649.

Lopez, J. V., S. Cevario, and S. J. O'Brien. 1996. Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. Genomics 33:229-246.

Martin, L. D. 1989. Fossil history of the terrestrial Carnivora. Pages 536-568 *in* Carnivore behavior, ecology and evolution. Vol. 1 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Matsuhashi, T., R. Masuda, T. Mano, and M. C. Yoshida. 1999. Microevolution of the mitochondrial DNA control region in the Japanese brown bear (*Ursus arctos*) population. Molecular Biology and Evolution 16:676-684.

Mindell, D. P., and C. E. Thacker. 1996. Rates of molecular evolution: phylogenetic issues and applications. Annual Review of Ecology and Systematics 27:279-303.

Novacek, M. J. 1992. Mammalian phylogeny: shaking the tree. Nature 356:121-125.

Paetkau, D., and C. Strobeck. 1996. Mitochondrial DNA and the phylogeography of Newfoundland black bears. Canadian Journal of Zoology 74:192-196.

Russo, C. A. M., N. Takezaki, and M. Nei. 1996. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Molecular Biology and Evolution 13:525-536.

Rychlik, W. 1997. Oligo^{TM}, version 5.1. National Biosciences, Inc.

Saitou, N., and M. Nei. 1987. The neighbor-joining method - a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406-425.

Soltis, D. E., P. S. Soltis, M. E. Mort, M. W. Chase, V. Savolainen, S. B. Hoot, and C. M. Morton. 1998. Inferring complex phylogenies using parsimony: An empirical approach using three large DNA data sets for angiosperms. Systematic Biology 47:32-42.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0b10. Sinauer Associates.

Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference. Pages 407-514 *in* Molecular Systematics (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland MASS.

Talbot, S. L., and G. F. Shields. 1996. A phylogeny of bears (Ursidae) inferred from complete sequences of three mitochondrial genes. Molecular Phylogenetics and Evolution 5:567-575.

Thenius, E. 1990. Carnivores: Phylogeny. Pages 370-383 *in* Encyclopedia of Mammals. Vol. 3 (Grzimek, ed.). McGraw-Hill, New-York.

Waits, L. P., J. Sullivan, S. J. O'Brien, and R. H. Ward. 1999. Rapid radiation events in the family Ursidae indicated by likelihood phylogenetic estimation from multiple fragments of mtDNA. Molecular Phylogenetics and Evolution 13:82-92.

Wayne, R. K., R. E. Benveniste, D. N. Janczewski, and S. J. O'Brien. 1989. Molecular and biochemical evolution of the Carnivora. Pages 465-488 *in* Carnivore behavior, ecology and evolution (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Wooding, S., and R. Ward. 1997. Phylogeography and Pleistocene evolution in the North American black bear. Molecular Biology and Evolution 14:1096-1105.

Xu, X. F., and U. Arnason. 1994. The complete mitochondrial-DNA sequence of the horse, *Equus caballus* - extensive heteroplasmy of the control region. Gene 148:357-362.

Yang, Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Molecular Biology and Evolution 10:1396-1401.

Chapter 3

A phylogeny of the Caniformia based on complete protein-coding mitochondrial sequences

Introduction

The phylogeny of the mammalian order Carnivora has been studied extensively. The order Carnivora is traditionally divided at the suborder level into Caniformia, dog-like carnivores, and Feliformia, cat-like carnivores. Both suborders are further subdivided into several family groups. Various work based on morphological characters (Wyss and Flynn, 1993), karyotypes (Arnason, 1977), DNA-DNA hybridization techniques (Arnason and Widegren, 1986), DNA sequence data (Arnason et al., 1995; Ledje and Arnason, 1996a, 1996b; Flynn et al., 2000), total evidence (Vrana et al., 1994; Dragoo and Honeycutt, 1997; Flynn and Nedbal, 1998), and supertree strategies (Bininda-Emonds et al., 1999), generally agree on the monophyly of these two suborders and of most carnivore families, but have failed to resolve the tree with significant support either above or below the family level (Flynn and Nedbal, 1998). The present work aims at resolving contentious relationships among caniform families.

Caniformia are commonly subdivided into two infraorders (also designated superfamilies by some authors, see discussion in Simpson, 1945; and Flynn et al., 1988): Cynoidea (or Canoidea) comprising a single family, Canidae (dogs and foxes), sister to Arctoidea, a more diverse infraorder composed of several families of bear-like carnivores. Traditional arctoid families are: Ursidae (bears), Mustelidae (weasels), Procyonidae (raccoons), Phocidae (true or earless seals), Otariidae (eared seals) and Odobenidae (walrus). The latter three families are grouped together as Pinnipedia, or aquatic carnivores.

Inter-familial relationships within a monophyletic infraorder Arctoidea are largely unresolved. There are several points of contention, described in detail in the Introduction: the monophyly of Pinnipedia, the relative position of the three pinniped families, the ancestral affinities of Pinnipedia, Ursidae and Mustelidae, the position of the red panda, *Ailurus fulgens*, and the traditional inclusion of skunks within Mustelidae, are all being questioned.

The great diversity and adaptability of carnivore species translates into multiple instances of convergent evolution (Martin, 1989) that have complicated morphological and paleontological systematic studies. Also, the evolutionary history of carnivores appears to have been punctuated by multiple events of rapid radiation (Ewer, 1973; Martin, 1989). These explosive and relatively fast bursts of evolution are represented on evolutionary trees by very short branches. Short branches contain, by their nature, few informative characters needed for reconstructing phylogeny. In these cases, the amount

of data collected is often simply insufficient to resolve relationships generated during explosive radiation events. On the other hand, efforts at collecting large data sets often do so at the expense of taxon representation. The resolution of the issues stated above therefore necessitates a large data set and representative taxa sampling. In the present study, a previously developed method to efficiently acquire large amounts of sequence data from the mitochondrial DNA of many carnivores (D elisle and Strobeck, 2002; Chapter 2) was used. Combined with previously published sequences, a data set of twelve mitochondrial genes (a total length of 10 842 nucleotides) for 38 representative carnivore species was assembled. Analysis using maximum parsimony, maximum likelihood and Bayesian approaches provided a unique and well supported solution to most contentious relationships within Caniformia.

Materials and methods

Data collection

Thirty-eight species were included in this study and are listed in Table 3-1. The suborder Caniformia was represented by 35 species while three species of the suborder Feliformia were used as outgroup taxa. Sequences were obtained for 12 mitochondrial genes [Cytochrome c oxidase (COX) subunits 1, 2, and 3, cytochrome b (Cyt b), NADH dehydrogenase (ND) subunits 1, 2, 3, 4, 4L, and 5, and ATP synthase F0 (ATP) subunits 6 and 8]. A thirteenth gene, ND6, is encoded on the opposite strand of the mitochondrial

DNA molecule and is thought to be subject to different evolutionary constraints. Consequently, it is generally excluded from phylogenetic studies using large mitochondrial data sets (see for example Arnason et al., 2002; and Lin et al., 2002). Based on this, and in an effort to make this study comparable to others using similar data, ND6 sequences were not included in the present study. New sequences were obtained using the method of Delisle and Strobeck (2002; Chapter 2). Additional primers were designed to accommodate regions particularly difficult to sequence in some species. All primers are listed in Appendix A. When available, sequences were retrieved from Genbank. Details of sequence provenance are given in Table 3-2. The 12 mitochondrial genes were assembled in a partitioned matrix. Mitochondrial genes are often *a priori* thought to constitute a single data partition, being located on the mitochondrial chromosome, a DNA molecule generally assumed to be exempt of recombination and hence behaving as a single evolutionary unit. I however verified this assumption using the partition homogeneity test (ILD test; Farris et al., 1994; 1995) implemented in PAUP* (Swofford, 2003) to compare gene partition in a multiple partition test and in all possible pair-wise combinations. A different way to partition the data, according to codon position, was also tested. Also, each potential partition was tested on its own against the rest of the data. ILD tests were performed on 100-1000 random addition sequence replicates. An *a posteriori* approach to evaluate partition combinability was also used: the maximum likelihood tree obtained from the analysis of the data set as a whole was tested, using the Shimodaira-Hasegawa test (SH test; Shimodaira and

Hasegawa, 1999, see below) against the best tree obtained for each gene analyzed separately.

Phylogenetic analysis

Based on *a priori* assumptions and partition homogeneity tests, the 12 mitochondrial genes were analyzed as a concatenated data set of a total nucleotide length of 10 842. Sequences were analyzed using maximum parsimony (MP; Farris et al., 1970; Fitch, 1971), maximum likelihood (ML; Felsenstein, 1981) and Bayesian (Rannala and Yang, 1996; Yang and Rannala, 1997; Larget and Simon, 1999) inference methods.

MP analysis was performed using PAUP* (Swofford, 2003) first on the data coded with the four nucleotide states A, T, G, and C, and then on the same data recoded into two states, purine or pyrimidine (hereafter called RY-coding). Most parsimonious trees were obtained from heuristic searches (using TBR tree swapping method) of ten random addition replicates. Bootstrap values were obtained by performing 1000 pseudo-replicates of the same analyses.

The program Modeltest (Posada and Crandall, 1998) was used on the MP topology to determine the model of substitution that best fits the data, for use in ML and Bayesian searches.

An iterative approach to searching the ML tree was used as the computational time was made prohibitive by the size of the data set. This approach consists in using a starting tree obtained from a fast method (I used the MP tree and the uncorrected *p* distance neighbour-joining tree in two independent searches) to estimate all the parameters of the best model specified by Modeltest. The estimated parameters are then used in a heuristic search for the ML tree. The tree obtained at this step is then used to re-estimate the parameters. The new parameters are used in a second heuristic search. The cycle of parameter estimation and tree searching is performed until the estimates have stabilized. PAUP* (Swofford, 2003) was used at both steps of the iterative cycle.

Bayesian analysis performed using MrBayes (Ronquist and Huelsenbeck, 2003) was run twice to verify consistency, each time with four chains of which three were heated, for 2 000 000 generations. Trees were sampled every 100 generations and the first 10% were discarded as "burn-in". The remaining trees were imported into PAUP (Swofford, 2003) and posterior probabilities at each node were calculated using the majority rule (50%) consensus method. Credible intervals of likelihood and model parameters (visualized in the program Tracer, Rambaut and Drummond, 2003), credible set of trees, and posterior probabilities of specific topologies, were all computed after "burn-in". Both runs were congruent and were combined to calculate overall posterior probabilities. In addition, another Bayesian analysis (5 000 000 generations) was performed using a more complex model where parameters are estimated for each gene partition.

All trees were rooted using three species of the suborder Feliformia (*Felis catus*, *Lynx canadensis*, and *Puma concolor*; Table 3-1). Moreover, to verify that no bias related to the choice of outgroup was observed, MP and ML analyses were also performed using various combinations of further outgroups from the orders considered the closest extant relatives to Carnivora: Perissodactyla (*Equus caballus*, NC_001640; Xu and Arnason, 1994), Artiodactyla (*Bos taurus*, NC_001567; Anderson et al., 1982), Pholidota (*Manis tetradactyla*, NC_004027; Arnason et al., 2002), and Chiroptera (*Pteropus scapulatus*, NC_002619; Lin and Penny, 2001).

Competing tree topology hypotheses were compared using the nonparametric Shimodaira-Hasegawa test (Shimodaira and Hasegawa, 1999 - SH test using 10 000 RELL approximation replicates, as implemented in PAUP*). SH tests results were compared to partial optimizations (*posPpud* of Goldman et al., 2000) of the SOWH test (Swofford et al., 1996) performed on a taxa subset of ten species, one per family. Species selected as representatives were: domestic dog, domestic cat, red panda, striped skunk, raccoon, wolverine, polar bear, walrus, South American sea lion and crabeater seal (see Tables 3-1 and 3-2 for Latin names and sources). Prior to performing SOWH tests, the pruned taxa set was subjected to model testing and maximum likelihood heuristic search and nonparametric bootstrapping using 1000 RELL pseudo-replicates. The ML topology, here favouring the Pinnipedia-Ursidae sister relationship with a bootstrap proportion of 59%, was tested against the other two alternatives. Results from the SH and SOWH tests were compared to Bayesian posterior probabilities (BP) of each tree, a measure that has

been suggested to represent the probability of it being the true tree given the observed data and a model (Larget and Simon, 1999; Lewis, 2001). The BP of a tree is the number of times it is visited by the Markov chain over the total number of trees sampled.

Results

The ILD tests showed no significant ($\alpha = 0.05$) conflict between any of the partitions. The best fit model of substitution for the concatenated data set selected by Modeltest was the general time reversible model taking into account the proportion of invariable sites and rate variation among sites (GTR + I + Γ). Point estimates of best fit GTR + I + Γ model parameters obtained from the ML analysis were included within the 95% credible intervals of the same parameters inferred from Bayesian analyses and are detailed in the legend to Figure 3-1. The *a posteriori* test of partition homogeneity using SH tests showed that the ML tree was a possible explanation of the sequence data for each gene, with the exception of ND3, where the ML tree was rejected with a *P*-value of 0.0314. This value is however deemed not significant if a Bonferroni correction is applied to account for the number of similar tests performed ($\alpha_c = 0.05/12 = 0.004$). Also, a ML analysis using the same data set minus the ND3 gene sequence gives the same tree topology as the complete data set. Given these results, all further analyses were performed on the complete data set.

The ML tree obtained from two separate iterative searches is presented in Figure 3-1. The two different models used in Bayesian analysis, using the same parameters estimates for the whole data set, or different parameters for each gene partition, gave similar results. Bayesian posterior probabilities calculated (after "burn-in") from the analysis using uniform parameters are shown on the ML tree. These posterior probabilities are generally high and in agreement with the ML topology. Four nodes received weaker support and are pictured with dotted lines. Maximum parsimony results obtained from the analysis of the four-state data differed from those obtained from the RY-coded data (Figure 3-2). The four-state data generated a tree with one significant (parsimony bootstrap support higher than 70%) disagreement with the ML / Bayesian tree presented in Figure 3-1. In parsimony analysis, the red panda – skunk clade is sister to the ursid clade, as opposed to sister to musteloids. RY-coding however reconciled parsimony analysis with ML and Bayesian analyses. The use of additional outgroups did not change the internal relationships among Caniformia (data not shown).

The two different methods employed to measure support, Bayesian posterior probabilities and parsimony bootstrap using four-state and RY-coded data (values in parentheses below in that order), provided identical values of 100% for the monophyly of Arctoidea. The monophyly of major arctoid clades was strongly supported for Pinnipedia (100/100/100, further discussed in Davis et al., 2004), and Ursidae (100/100/100), but divided in the case of Musteloidea (100/23/82). There was, however, no strong evidence about the relative position of these three arctoid clades, effectively resulting in a

polytomy (see Table 3-3). Within these clades, relationships between families were well supported. Within Pinnipedia, the families Otariidae and Odobenidae were shown to form a clade (100/100/100), sister to Phocidae. High support was also found for the sister relationship between true mustelids and procyonids (100/70/94), and between ailurids and mephitids (100/73/77). There were four other less supported nodes, in areas of the tree that are situated outside the major focus of this chapter, which concerns inter-familial relationships within Arctoidea. Two of these nodes are within Phocidae: the node placing the Ross seal (*Ommatophoca rossi*) as the sister species to the Weddell and leopard seals clade (*Leptonychotes weddelli* and *Lobodon carcinophagus*), and the node joining the grey seal (*Halichoerus grypus*) and the ringed seal (*Pusa hispida*) as sister species received overall low support (discussed in Davis et al., 2004). Another weakly supported node was within felids, the outgroup for this study. Finally, the relationship of the sun bear (*Ursus malayanus*) as the sister species to the polar-brown bear clade (*Ursus maritimus* and *Ursus arctos*), although highly supported by Bayesian analysis, received moderate support in parsimony analysis.

This analysis therefore fully resolves inter-familial relationships within Arctoidea, with the exception of a single node. Table 3-3 presents a further investigation of support obtained for each of the three possible strictly dichotomous relationships at this node. First, support values for each topology were compared. Bayesian analysis marginally favoured (50% posterior probability) placing the bears as the sister clade to pinnipeds. Parsimony bootstrap support was divided, but RY parsimony bootstrap also more

strongly supported (70%) the placement of Ursidae and Pinnipedia as sister clades. Statistical tests of topologies were then performed. The Shimodaira-Hasegawa test showed no significant difference between these three possible topologies. SOWH tests performed using a subset of ten species (identified in Table 3-1) found that the topology where the Ursidae is sister to the Pinnipedia is a significantly better tree.

Discussion

Inter-familial relationships among Caniformia

Inter-familial relationships among Caniformia were almost completely resolved (Figures 3-1 and 3-2). The mitochondrial trees obtained here support prevailing views of arctoid (100/100/100) and pinniped (100/100/100) monophyly (see Flynn, 1996, and references therein). It is also in agreement with most molecular studies in placing the families Otariidae and Odobenidae on a common branch (100/100/100), as the sister group to Phocidae (Vrana et al., 1994; Arnason et al., 1995; Ledje and Arnason, 1996a, 1996b; Flynn and Nedbal, 1998; Flynn et al., 2000). A similar arrangement is also recovered in the supertree analysis of Bininda-Emonds et al. (1999).

A close affinity between Mustelidae and Procyonidae, grouped in the superfamily Musteloidea, is generally accepted (see Flynn et al., 1988). Skunks are however increasingly often excluded from the true mustelids and placed in their own family, the

Mephitidae, branching off at the base of a true mustelid-procyonid clade (Arnason and Widegren, 1986; Ledje and Arnason, 1996a, 1996b; Dragoo and Honeycutt, 1997; Flynn et al., 2000). The present analysis corroborates these findings (true mustelids-procyonids clade support values: 100/70/94) and suggests that analyses that constrained the Mustelidae to include skunks, i.e. Mustelidae *sensu lato* (Bryant et al., 1993; Bininda-Emonds et al., 1999), may be inherently flawed. The growing uncertainty of the true mustelid affinity of the skunks has led some authors to omit them from their phylogenetic studies of the Mustelidae (Koepfli and Wayne, 1998; 2003). Both strategies, of enforcing monophyly or of omitting "problematic" taxa, may lead to the loss of important phylogenetic information and can potentially obscure unexpected phylogenetic relationships. For example, the sister relationship between skunks and the red panda suggested by Flynn et al. (2000), has been largely overlooked, probably because of the novelty of the suggestion and the low bootstrap support it received from their four gene analysis of 17 (some chimeric) taxa from the order Carnivora. The present analysis of twelve mitochondrial genes conferred more confidence in this recent hypothesis (with posterior probability/parsimony bootstrap values of 100/73/77). While a sister relationship of the red panda to the Musteloidea has been suggested before (Flynn and Nedbal, 1998; Bininda-Emonds et al., 1999), the demonstration of its possible closer affinity to mephitids was precluded for the reasons stated above.

Relationships below the family level

Relationships below the family level were also strongly supported. For instance, a highly supported phylogeny of the extant Phocidae, showing the monophyly of the subfamilies Phocinae (Northern phocids; *Phoca* and allies; 100/100/99) and Monachinae (Southern phocids, elephant seals and monk seals; *Monachus* and allies; 100/100/100) was obtained. The genus *Monachus*, comprising the two monk seal species, was monophyletic (100/99/81), branching off at the base of the Monachinae (100/96/96), hence resolving two previously contentious issues (further discussed in Davis et al., 2004). Within the family Otariidae, sea lions (subfamily Otariinae) were shown to be paraphyletic. This result is in agreement with the cytochrome b/control region analysis of Wynen et al. (2001). This adds to the growing evidence from various sources indicating that the subfamilial designations within Otariidae, based on the single character of presence/absence of underfur, are inappropriate (Repenning et al., 1971; Lento et al., 1995; Berta and Sumich, 1999; Bininda-Emonds et al., 1999; Brunner, 2000). Relationships between the five true mustelids included in this work were in agreement with more focused studies that included more taxa and less data (Koepfli and Wayne, 1998, 2003; Sato et al., 2003). Strongly supported familial phylogenies, and corroboration of other studies indicate that large mitochondrial data sets contain relevant phylogenetic information at this level of divergence as well.

Advantages of a large data set

Early molecular phylogenetic analyses of the Carnivora, although sometimes based on very similar data sets, have shown contradicting results, depending on the choice of method of analysis, taxon sampling and *a priori* assumptions concerning the monophyly of certain clades or character weighting schemes (for example, compare Zhang and Ryder, 1993; Vrana et al., 1994; Ledje and Arnason, 1996a, 1996b; Dragoo and Honeycutt, 1997). More recent work has shown that combining sequence data from several genes, and the use of more appropriate models of analysis, have great potential to resolve these points of contention (see Flynn and Nedbal, 1998; Flynn et al., 2000; Koepfli and Wayne, 2003). The present analysis, including 38 carnivore species and data sequence from 12 mitochondrial genes in a concatenated data set of 10 842 nucleotides, showed that previous conflicting hypotheses were in fact mainly the result of a lack of sufficient informative data. The three methods of analysis used here agreed on a unique tree of carnivore families that contains a single unresolved node. The different support statistics (non-parametric bootstrapping and Bayesian posterior probabilities), and statistical comparisons of topologies (SH and SOWH tests) used were also in agreement on strong and weaker parts of the species tree. Agreement between different methods is a good indication that the analysis is showing a phylogenetic signal above the level of noise in the data, resulting in a phylogenetic tree independent of biases specific to any particular method (Kim, 1993). This is good indication that the analysis is converging on the true mitochondrial tree.

Several of these findings could only be obtained by an analysis that imposes no constraints on the analysis, does not assume the monophyly of particular groups, and uses of appropriate outgroups to leave ingroup relationships vary freely. Large scale data collection techniques, better evolutionary models, more efficient tree searching algorithms, and faster processors, are constantly alleviating the needs for using restricted data and taxa sets. The collection and analysis using several methods of a large data (10 842 bp) and taxon (35 caniform species) set was therefore possible, and from it emerged a strongly supported phylogeny of the Caniformia. The phylogenetic tree of the Caniformia showed relationships that are in agreement with several earlier studies and, in this sense, represents a refinement and a consolidation of earlier findings based on fewer genes and taxa. This improved confidence in the phylogeny is a strong argument in favour of gene concatenation in phylogenetic analysis.

Further work

The generally strongly supported tree obtained in this work nevertheless still contains some weaker nodes and an important trichotomy involving pinnipeds, ursids, and musteloids. The basal arctoid trichotomy reflects the very rapid evolution of carnivore families that occurred during the Eocene-Oligocene transition (Ewer, 1973; Werdelin, 1996). Furthermore, even though this work represents the most extensive molecular data set collected so far, it still only comprises less than a quarter of all caniform species. It is therefore possible that the addition of other key taxa, such as the deep branching ursid *Ailuropoda melanoleuca* (Giant panda), and more representatives of the speciose

musteloid families, may help resolve this node and strengthen weaker relationships. More improvements to this study will be possible in the future. First, ongoing sequencing of mitochondrial genes from carnivore species will expand the present taxon set. Also, the mitochondrial data representing a maternal lineage, and being subject potential problems such as lineage sorting, analyses from other lines of evidence such as nuclear genes and the paternally inherited Y-chromosome will be essential in corroborating these findings. Also, as has often been the case, a review of the morphological data in light of recent hypotheses suggesting novel relationships may provide new insights. Ultimately, comparative analyses of the various data available, as well as total evidence approaches, supertrees and supermatrix methods, should result in a clearer picture of carnivore phylogeny, and a robust framework for a better understanding of the evolution of the Carnivora.

Table 3-1 Species list. Genbank accession numbers are in Table 3-2. Species marked with an asterisk were used as the family representatives for the SOWH test taxa set.

Carnivora
Caniformia
Mustelidae
Martes americana (American marten)
Lontra canadensis (North American river otter)
Taxidea taxus (American badger)
* *Gulo gulo* (Wolverine)
Mustela vison (Mink)
Mephitidae
* *Mephitis mephitis* (Striped skunk)
Procyonidae
* *Procyon lotor* (Raccoon)
Ailuridae
* *Ailurus fulgens* (Red panda)
Ursidae
Ursus arctos (Brown bear)
Ursus americanus (American black bear)
* *Ursus maritimus* (Polar bear)
Ursus malayanus (Sun bear)
Phocidae
Cystophora cristata (Hooded seal)
Erignatus barbatus (Bearded seal)
Hydrurga leptonyx (Leopard seal)
* *Lobodon carcinophagus* (Crabeater seal)
Mirounga angustirostris (Northern elephant seal)
Mirounga leonine (Southern elephant seal)
Monachus schauinslandi (Hawaiian monk seal)
Monachus monachus (Mediterranean monk seal)
Ommatophoca rossi (Ross seal)
Pagophilus groenlandicus (Harp seal)
Pusa hispida (Ringed seal)
Phoca largha (Spotted seal)
Leptonychotes weddelli (Weddell seal)
Halichoerus grypus (Grey seal)
Phoca vitulina (Harbour seal)
Otariidae
Arctocephalus australis (Southern fur seal)
* *Otaria byronia* (South American sea lion)
Arctocephalus fosteri (New Zealand fur seal)
Eumetopias jubatus (Steller sea lion)
Odobenidae
* *Odobenus rosmarus* (Walrus)
Canidae
* *Canis familiaris* (Domestic dog)
Canis lupus (Wolf)
Alopex lagopus (Arctic fox)

Table 3-1 Continued.

Feliformia
Felidae
* *Felis catus* (Domestic cat)
Lynx canadensis (Canada lynx)
Puma concolor (cougar)

Table 3-2 List of species and sequence Genbank accession numbers.

Sequences from this study

From Chapter 2

Ursus arctos (Brown bear)	From mitochondrial genome NC_003427 (Delisle and Strobeck, 2002)
Ursus americanus (American black bear)	From mitochondrial genome NC_003426 (Delisle and Strobeck, 2002)
Ursus maritimus (Polar bear)	From mitochondrial genome NC_003428 (Delisle and Strobeck, 2002)

From Chapter 3

	Gene					
	COX 1	COX2	COX3	Cyt b	ND1	ND2
Lynx canadensis (Canada lynx)	AY598472	AY598473	AY598474	AY598475	AY598476	AY598477
Puma concolor (Cougar)	AY598484	AY598485	AY598486	AY598487	AY598488	AY598489
Canis lupus (Wolf)	AY598496	AY598497	AY598498	AY598499	AY598500	AY598501
Alopex lagopus (Arctic fox)	AY598508	AY598509	AY598510	AY598511	AY598512	AY598513
Ailurus fulgens (Red panda)	AY598520	AY598521	AY598522	X94919[a]	AY598523	AY598524
Mephitis mephitis (Striped skunk)	AY598531	AY598532	AY598533	X94927[a]	AY598534	AY598535
Martes americana (American marten)	AY598542	AY598543	AY598544	AY121352[b]	AY598545	AY598546
Lontra canadensis (North American river otter)	AY598553	AY598554	AY598555	AF057121[c]	AY598556	AY598557
Taxidea taxus (American badger)	AY598564	AY598565	AY598566	AF057132[c]	AY598567	AY598568
Procyon lotor (Raccoon)	AY598575	AY598576	AY598577	X94930[a]	AY598578	AY598579
Ursus malayanus (Sun bear)	AY598586	AY598587	AY598588	U18899[d]	AY598589	AY598590

	Gene					
	ND3	ND4	ND4L	ND5	ATP6	ATP8
Lynx canadensis (Canada lynx)	AY598478	AY598479	AY598480	AY598481	AY598470	AY598471
Puma concolor (Cougar)	AY598490	AY598491	AY598492	AY598493	AY598482	AY598483
Canis lupus (Wolf)	AY598502	AY598503	AY598504	AY598505	AY598494	AY598495
Alopex lagopus (Arctic fox)	AY598514	AY598515	AY598516	AY598517	AY598506	AY598507
Ailurus fulgens (Red panda)	AY598525	AY598526	AY598527	AY598528	AY598518	AY598519
Mephitis mephitis (Striped skunk)	AY598536	AY598537	AY598538	AY598539	AY598529	AY598530
Martes americana (American marten)	AY598547	AY598548	AY598549	AY598550	AY598540	AY598541
Lontra canadensis (North American river otter)	AY598558	AY598559	AY598560	AY598561	AY598551	AY598552
Taxidea taxus (American badger)	AY598569	AY598570	AY598571	AY598572	AY598562	AY598563
Procyon lotor (Raccoon)	AY598580	AY598581	AY598582	AY598583	AY598573	AY598574
Ursus malayanus (Sun bear)	AY598591	AY598592	AY598593	AY598594	AY598584	AY598585

Table 3-2 Continued.

Sequences in Davis et al. (2004)

	Gene					
	COX 1	COX2	COX3	Cyt b	ND1	ND2
Odobenus rosmarus (Walrus)	AY377148	AY377171	AY377263	X82299[e]	AY377361	AY377281
Arctocephalus australis (Southern fur seal)	AY377150	AY377173	AY377265	AY377329	AY377363	AY377283
Otaria byronia (South American sea lion)	AY377149	AY377172	AY377264	AY377328	AY377362	AY377282
Cystophora cristata (Hooded seal)	AY377144	AY377167	AY377259	X82294[e]	AY377357	AY377277
Erignatus barbatus (Bearded seal)	AY377143	AY377166	AY377258	X82295[e]	AY377356	AY377276
Hydrurga leptonyx (Leopard seal)	AY377134	AY377157	AY377249	AY377323	AY377350	AY377270
Lobodon carcinophagus (Crabeater seal)	AY377130	AY377153	AY377245	AY377321	AY377348	AY377268
Mirounga angustirostris (Northern elephant seal)	AY377138	AY377161	AY377253	AY377325	AY377352	AY377272
Mirounga leonina (Southern elephant seal)	AY377140	AY377163	AY377255	AY377326	AY377353	AY377273
Monachus schauinslandi (Hawaiian monk seal)	AY377141	AY377164	AY377256	X72209[f]	AY377354	AY377274
Monachus monachus (Mediterranean monk seal)	AY377142	AY377165	AY377257	AY378327	AY377355	AY377275
Ommatophoca rossi (Ross seal)	AY377132	AY377155	AY377247	AY377322	AY377349	AY377269
Pagophilus groenlandicus (Harp seal)	AY377145	AY377168	AY377260	X82303[e]	AY377358	AY377278
Pusa hispida (Ringed seal)	AY377146	AY377169	AY377261	X82304[e]	AY377359	AY377279
Phoca largha (Spotted seal)	AY377147	AY377170	AY377262	X82305[e]	AY377360	AY377280
Leptonychotes weddelli (Weddell seal)	AY377136	AY377159	AY377251	AY377324	AY377351	AY377271
Gulo gulo (Wolverine)	AY377151	AY377174	AY377266	AB051245[g]	AY377364	AY377284
Mustela vison (Mink)	AY377152	AY377175	AY377267	AB026109[h]	AY377365	AY377285

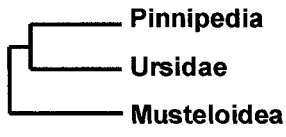
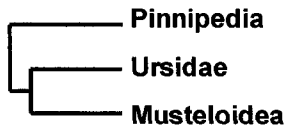
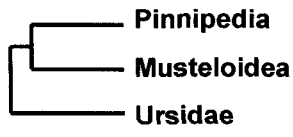
	Gene					
	ND3	ND4	ND4L	ND5	ATP6	ATP8
Odobenus rosmarus (Walrus)	AY377217	AY377343	AY377240	AY377380	AY377310	AY377194
Arctocephalus australis (Southern fur seal)	AY377219	AY377345	AY377242	AY377382	AY377312	AY377196
Otaria byronia (South American sea lion)	AY377218	AY377344	AY377241	AY377381	AY377311	AY377195
Cystophora cristata (Hooded seal)	AY377213	AY377339	AY377236	AY377376	AY377306	AY377190
Erignatus barbatus (Bearded seal)	AY377212	AY377338	AY377235	AY377375	AY377305	AY377189
Hydrurga leptonyx (Leopard seal)	AY377203	AY377332	AY377226	AY377369	AY377296	AY377180
Lobodon carcinophagus (Crabeater seal)	AY377199	AY377330	AY377222	AY377367	AY377292	AY377176
Mirounga angustirostris (Northern elephant seal)	AY377207	AY377334	AY377230	AY377371	AY377303	AY377184
Mirounga leonina (Southern elephant seal)	AY377209	AY377335	AY377232	AY377372	AY377302	AY377186
Monachus schauinslandi (Hawaiian monk seal)	AY377210	AY377336	AY377233	AY377373	AY377303	AY377187

Table 3-2 Continued.

	Gene					
	ND3	ND4	ND4L	ND5	ATP6	ATP8
Monachus monachus (Mediterranean monk seal)	AY377211	AY377337	AY377234	AY377374	AY377304	AY377188
Ommatophoca rossi (Ross seal)	AY377201	AY377331	AY377224	AY377368	AY377294	AY377178
Pagophilus groenlandicus (Harp seal)	AY377214	AY377340	AY377237	AY377377	AY377307	AY377191
Pusa hispida (Ringed seal)	AY377215	AY377341	AY377238	AY377378	AY377308	AY377192
Phoca largha (Spotted seal)	AY377216	AY377342	AY377239	AY377379	AY377309	AY377193
Leptonychotes weddelli (Weddell seal)	AY377205	AY377333	AY377228	AY377370	AY377298	AY377182
Gulo gulo (Wolverine)	AY377220	AY377346	AY377243	AY377383	AY377313	AY377197
Mustela vison (Mink)	AY377221	AY377347	AY377244	AY377384	AY377314	AY377198
Sequences from other sources						
Felis catus (Domestic cat)	From mitochondrial genome NC_001700 (Lopez et al., 1996)					
Canis familiaris (Domestic dog)	From mitochondrial genome NC_002008 (Kim et al., 1998)					
Arctocephalus fosteri (New Zealand fur seal)	From mitochondrial genome NC_004023 (Lin et al., 2002)					
Eumetopias jubatus (Steller sea lion)	From mitochondrial genome NC_004030 (Arnason et al., 2002)					
Halichoerus grypus (Grey seal)	From mitochondrial genome NC_001602 (Arnason et al., 1993)					
Phoca vitulina (Harbour seal)	From mitochondrial genome NC_001325 (Arnason and Johnsson, 1992)					

[a] indicates sequences from Ledje and Arnason (1996a)
[b] indicates sequence from Stone et al. (2002)
[c] indicates sequences from Koepfli and Wayne (1998)
[d] indicates sequence from Talbot and Shields (1996)
[e] indicates sequences from Arnason et al. (1995)
[f] indicates sequence from Arnason et. (1993)
[g] indicates sequence from Hosoda et al.(2000)
[h] indicates sequence from Kurose et al. (2000)

Table 3-3 Comparisons of competing topologies at arctoid polytomy. The three possible resolved topologies are evaluated; see the text for details on specific measures and tests. Bolded values indicate the arrangement favoured by individual support measures. For the SOWH tests, $\alpha_c = 0.05/2 = 0.025$.

	Resolved topologies		
Support measure			
Likelihood of best tree satisfying the topological constraint	128 894.45826	**128 894.27576**	128 895.88154
Bayesian posterior probability (ML tree)	**8.9%**	8.8%	1.8%
Bayesian posterior probability (node)	**50%**	42%	8%
Parsimony bootstrap	17%	**44%**	18%
RY parsimony bootstrap	**70%**	28%	2%
SH test	0.6632	**best**	0.5815
SOWH test	**best**	<0.01*	0.01*

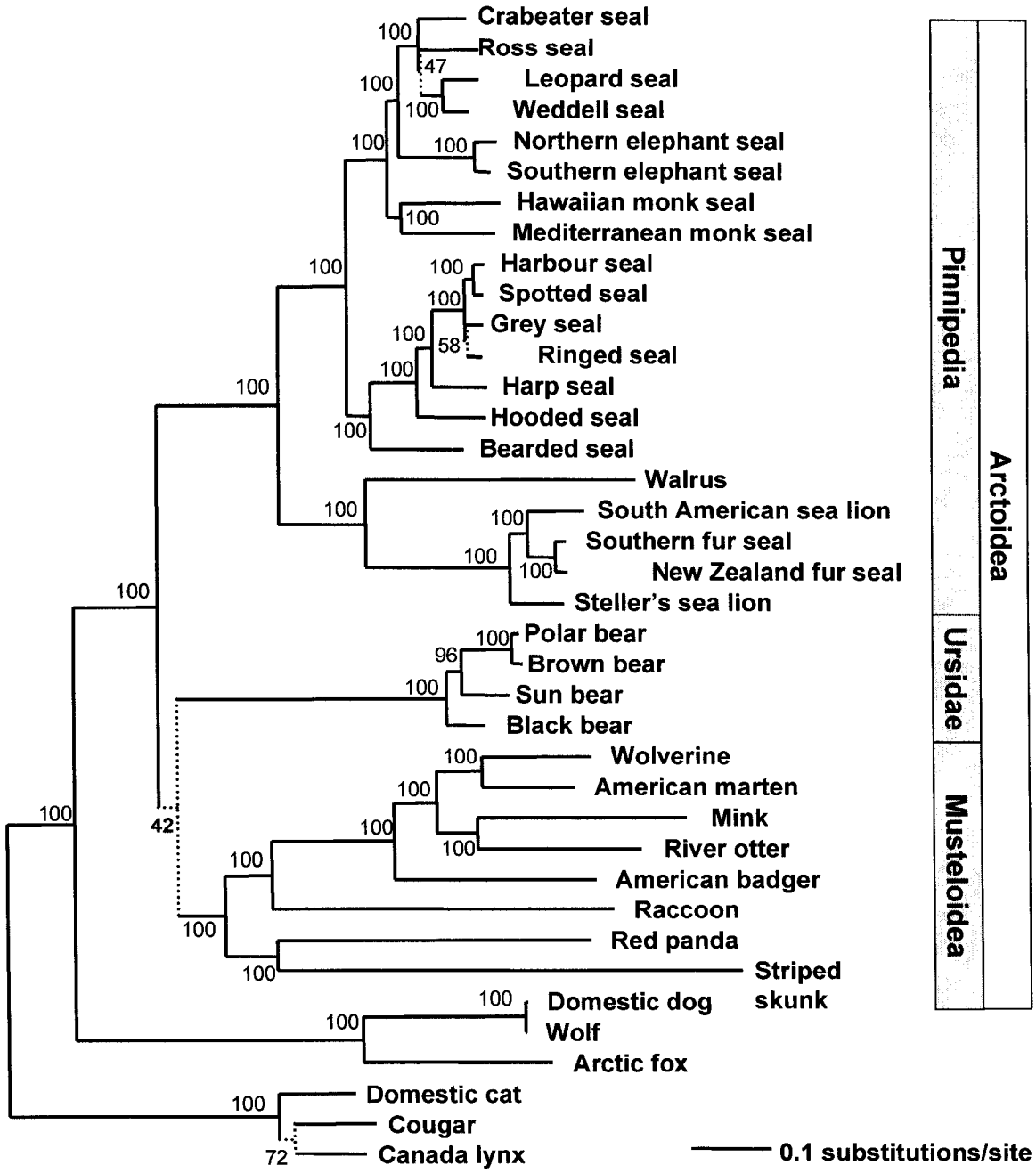


Figure 3-1 Maximum likelihood tree of 38 carnivore species obtained from two separate iterative tree searches. Support values for each node are Bayesian posterior probabilities (BP). Relationships supported by BP less than 95% are shown with dotted lines. (-lnL = 128 894.20268). GTR + I + Γ parameters estimates were: base frequencies A 0.361458, C 0.331058, G 0.071673, T 0.235811; rate matrix AC 0.44266, AG 19.29533, AT 0.78722, CG 0.41188, CT 14.11779, GT 1.00000; proportion of invariable sites I 0.460967; and gamma shape, , 0.927588.

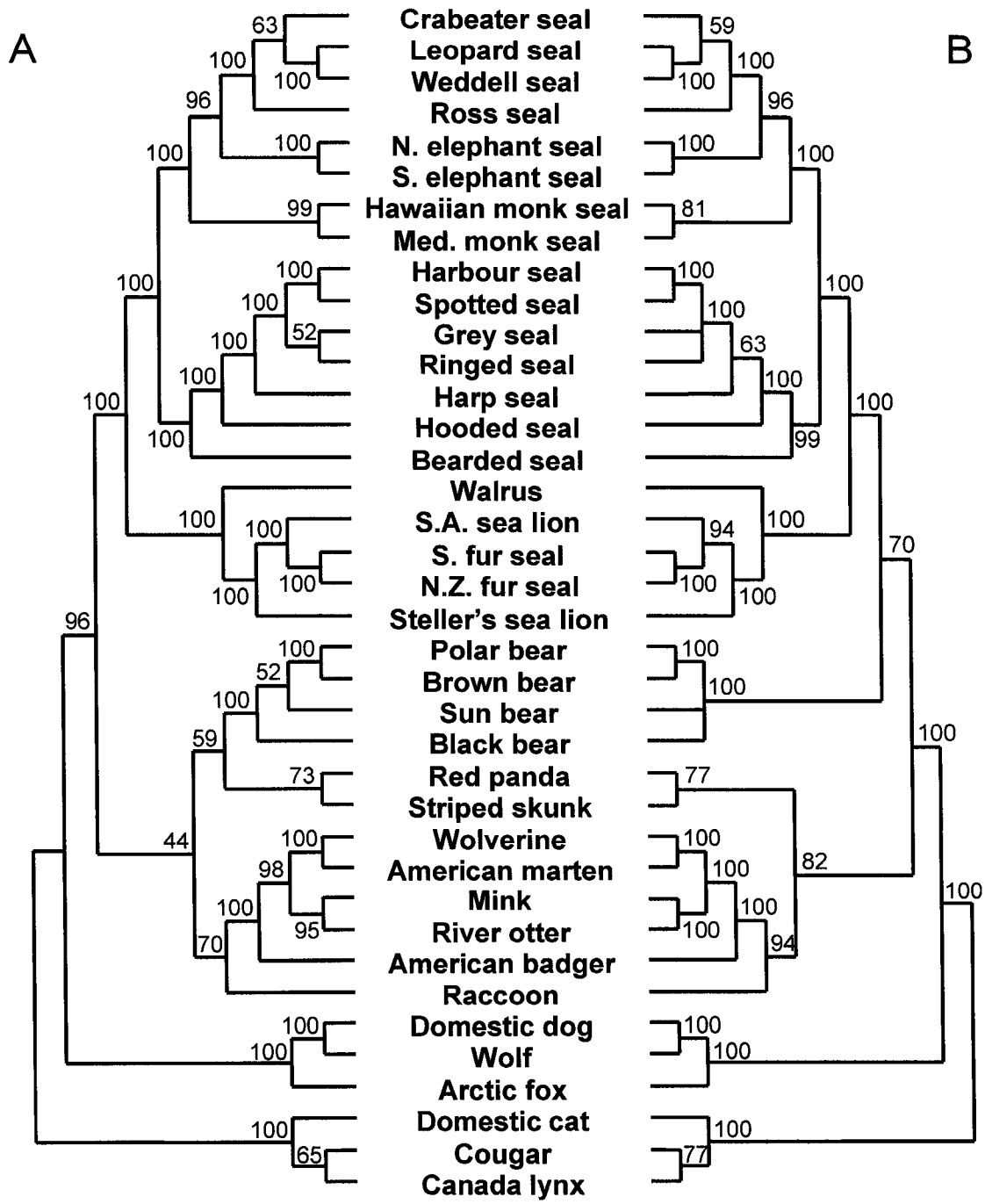


Figure 3-2 Parsimony analysis of relationships between 38 carnivore species. **A.** Single most parsimonious tree obtained from four-state-coded data. **B.** Strict consensus of 6 equally parsimonious trees obtained from the analysis of the RY-coded data. Support values are from 1000 bootstrap pseudo-replicates of the respective data.

Bibliography

Anderson, S., M. H. de Bruijn, A. R. Coulson, I. C. Eperon, F. Sanger, and I. G. Young. 1982. Complete sequence of bovine mitochondrial DNA. Conserved features of the mammalian mitochondrial genome. Journal of Molecular Biology 156:683-717.

Arnason, U. 1977. The relationship between the four principal pinniped karyotypes. Hereditas 87:227-242.

Arnason, U., J. A. Adegoke, K. Bodin, et al. 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. Proceedings of the National Academy of Sciences USA 99:8151-8156.

Arnason, U., K. Bodin, A. Gullberg, C. Ledje, and S. Mouchaty. 1995. A molecular view of pinniped relationships with particular emphasis on the true seals. Journal of Molecular Evolution 40:78-85.

Arnason, U., A. Gullberg, E. Johnsson, and C. Ledje. 1993. The nucleotide sequence of the mitochondrial molecule of the grey seal, *Halichoerus grypus*, and a comparison with mitochondrial sequences of other true seals. Journal of Molecular Evolution 43:41-45.

Arnason, U., and E. Johnsson. 1992. The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*. Journal of Molecular Evolution 34:493-505.

Arnason, U., and B. Widegren. 1986. Pinniped phylogeny enlightened by molecular hybridizations using highly repetitive DNA. Molecular Biology and Evolution 3:356-365.

Berta, A., and J. L. Sumich. 1999. Marine mammals : evolutionary biology. Academic Press, San Diego.

Bininda-Emonds, O. R. P., J. L. Gittleman, and A. Purvis. 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). Biological Reviews Of The Cambridge Philosophical Society 74:143-175.

Brunner, S. 2000. Cranial morphometrics of fur seals and sea lions (Family: Otariidae) – systematics, geographic variation and growth. Veterinary Anatomy and Pathology. University of Sydney, NSW 2006. Sydney. 254 pages

Bryant, H. N., A. P. Russell, and W. D. Fitch. 1993. Phylogenetic relationships within the extant Mustelidae (Carnivora): appraisal of the cladistic status of the Simpsonian subfamilies. Zoological Journal of the Linnean Society 108:301-334.

Davis, C. S., I. Delisle, D. B. Siniff, I. Stirling, and C. Strobeck. 2004. A phylogeny of the extant Phocidae inferred from complete mitochondrial DNA coding regions. Molecular Phylogenetics and Evolution (In press).

Delisle, I., and C. Strobeck. 2002. Conserved primers for rapid sequencing of the complete mitochondrial genome from carnivores, applied to three species of bears. Molecular Biology and Evolution 19:357-361.

Dragoo, J. W., and R. L. Honeycutt. 1997. Systematics of mustelid-like carnivores. Journal of Mammalogy 78:426-443.

Ewer, R. F. 1973. The Carnivores. Weidenfeld and Nicolson, London.

Farris, J. S., M. Kallersjo, A. G. Kluge, and C. Bult. 1994. Testing significance of incongruence. Cladistics 10:315-319.

Farris, J. S., M. Kallersjo, A. G. Kluge, and C. Bult. 1995. Constructing a significance test for incongruence. Systematic Biology 44:570-572.

Farris, J. S., A. G. Kluge, and M. J. Eckardt. 1970. A numerical approach to phylogenetic systematics. Systematic Zoology 19:172-189.

Felsenstein, J. 1981. Evolutionary trees from DNA-sequences - a maximum-likelihood approach. Journal of Molecular Evolution 17:368-376.

Fitch, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Systematic Zoology 20:406-416.

Flynn, J. J. 1996. Carnivore phylogeny and rates of evolution: morphological, taxic, and molecular. Pages 542-581 *in* Carnivore behavior, ecology, and evolution. Vol. 2 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Flynn, J. J., and M. A. Nedbal. 1998. Phylogeny of the Carnivora (Mammalia): Congruence vs incompatibility among multiple data sets. Molecular Phylogenetics and Evolution 9:414-426.

Flynn, J. J., M. A. Nedbal, J. W. Dragoo, and R. L. Honeycutt. 2000. Whence the red panda? Molecular Phylogenetics and Evolution 17:190-199.

Flynn, J. J., N. A. Neff, and R. H. Tedford. 1988. Phylogeny of the Carnivora. Pages 73-115 *in* The phylogeny and classification of the tetrapods. Vol. 2: Mammals (M. J. Benton, ed.). Clarendon Press, Oxford.

Goldman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49:652-670.

Hosoda, T., H. Suzuki, M. Harada, et al. 2000. Evolutionary trends of the mitochondrial lineage differentiation in species of genera Martes and Mustela. Genes and Genetic Systems 75:259-67.

Kim, J. 1993. Improving the accuracy of phylogenetic estimation by combining different phylogenetic methods. Systematic Biology 42:331-340.

Kim, K. S., S. E. Lee, H. W. Jeong, and J. H. Ha. 1998. The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. Molecular Phylogenetics and Evolution 10:210-220.

Koepfli, K. P., and R. K. Wayne. 1998. Phylogenetic relationships of otters (Carnivora:Mustelidae) based on mitochondrial cytochrome *b* sequences. Journal of Zoology 246:401-416.

Koepfli, K. P., and R. K. Wayne. 2003. Type ISTS markers are more informative than cytochrome b in phylogenetic reconstruction of the Mustelidae (Mammalia : Carnivora). Systematic Biology 52:571-593.

Kurose, N., A. V. Abramov, and R. Masuda. 2000. Intrageneric diversity of the cytochrome b gene and phylogeny of Eurasian species of the genus *Mustela* (Mustelidae, Carnivora). Zoological Science 17:673.

Larget, B., and D. Simon. 1999. Markov chains Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Molecular Biology and Evolution 16:750-759.

Ledje, C., and U. Arnason. 1996a. Phylogenetic analyses of complete cytochrome b genes of the order Carnivora with particular emphasis on the Caniformia. Journal of Molecular Evolution 42:135-144.

Ledje, C., and U. Arnason. 1996b. Phylogenetic relationships within caniform carnivores based on analyses of the mitochondrial 12S rRNA gene. Journal of Molecular Evolution 43:641-649.

Lento, G., R. Hickson, G. Chambers, and D. Penny. 1995. Use of spectral analysis to test hypotheses on the origin of pinnipeds. Molecular Biology and Evolution 12:28-52.

Lewis, P. O. 2001. Phylogenetic systematics turns over a new leaf. Trends in Ecology & Evolution 16:30-37.

Lin, Y.-H., P. A. McLenachan, A. R. Gore, M. J. Phillips, R. Ota, M. D. Hendy, and D. Penny. 2002. Four new mitochondrial genomes and the increased stability of evolutionary trees of mammals from improved taxon sampling. Molecular Biology and Evolution 19:2060-2070.

Lin, Y.-H., and D. Penny. 2001. Implications for bat evolution from two new complete mitochondrial genomes. Molecular Biology and Evolution 18:684-688.

Lopez, J. V., S. Cevario, and S. J. O'Brien. 1996. Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. Genomics 33:229-246.

Martin, L. D. 1989. Fossil history of the terrestrial Carnivora. Pages 536-568 *in* Carnivore behavior, ecology and evolution. Vol. 1 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Posada, D., and K. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Rambaut, A., and A. Drummond. 2003. Tracer, version v1.0.1.

Rannala, B., and Z. H. Yang. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. Journal of Molecular Evolution 43:304-311.

Repenning, C. A., R. S. Peterson, and C. L. Hubbs. 1971. Contributions to the systematics of the southern fur seals, with particular reference to the Juan Fernandez and Guadelupe species. Pages 1-52 *in* Antarctic Pinnipedia (W. H. Burt, ed.). American Geophysical Union.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Sato, J. J., T. Hosoda, M. Wolsan, K. Tsuchiya, M. Yamamoto, and H. Suzuki. 2003. Phylogenetic relationships and divergence times among mustelids (Mammalia: Carnivora) based on nucleotide sequences of the nuclear interphotoreceptor retinoid binding protein and mitochondrial cytochrome b genes. Zoological Science 20:243.

Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16:1114-1116.

Simpson, G. G. 1945. The principles of classification and a classification of mammals. The American Museum of Natural History. Bulletins, 85, New York.

Stone, K. D., R. W. Flynn, and J. A. Cook. 2002. Post-glacial colonization of northwestern North America by the forest-associated American marten (*Martes americana*, Mammalia: Carnivora: Mustelidae). Molecular Ecology 11:2049-2063.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0b10. Sinauer Associates.

Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference. Pages 407-514 *in* Molecular Systematics (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland MASS.

Talbot, S. L., and G. F. Shields. 1996. A phylogeny of bears (Ursidae) inferred from complete sequences of three mitochondrial genes. Molecular Phylogenetics and Evolution 5:567-575.

Vrana, P. B., M. C. Milinkovitch, J. R. Powell, and W. C. Wheeler. 1994. Higher level relationships of the arctoid Carnivora based on sequence data and "total evidence". Molecular Phylogenetics and Evolution 3:47-58.

Werdelin, L. 1996. Carnivoran ecomorphology: A phylogenetic perspective. Pages 582-624 *in* Carnivore behavior, ecology, and evolution. Vol. 2 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Wynen, L. P., S. D. Goldsworthy, S. J. Insley, et al. 2001. Phylogenetic relationships within the eared seals (Otariidae : Carnivora): Implications for the historical biogeography of the family. Molecular Phylogenetics and Evolution 21:270-284.

Wyss, A. R., and J. J. Flynn. 1993. A phylogenetic analysis and definition of the Carnivora. Pages 32-52 *in* Mammal Phylogeny: Placentals. Vol. 2 (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.

Xu, X. F., and U. Arnason. 1994. The complete mitochondrial-DNA sequence of the horse, *Equus caballus* - extensive heteroplasmy of the control region. Gene 148:357-362.

Yang, Z., and B. Rannala. 1997. Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte Carlo method. Molecular Biology and Evolution 14:717-724.

Zhang, Y.-P., and O. A. Ryder. 1993. Mitochondrial DNA sequence evolution in the Arctoidea. Proceedings of the National Academy of Sciences USA 90:9557-9561.

Chapter 4

Evolutionary affinity of red panda and skunks strongly supported by mitochondrial multi-gene data

Introduction

The epithet "enigmatic" commonly associated with the red panda (*Ailurus fulgens*) alludes to its unresolved evolutionary affinities. This small mammal of the order Carnivora, named for its resemblance to a cat, has alternately been placed closer to the giant panda (Ewer, 1973), raccoons (Simpson, 1945; O'Brien et al., 1985; Dragoo and Honeycutt, 1997; McKenna and Bell, 1997; Nowak, 1999), bears (Sarich, 1973; Wilson and Reeder, 1993), skunks (Flynn et al., 2000), or dogs (Ledje and Arnason, 1996a). It has also been suggested that it shares a more distant ancestor with a bear and pinniped clade (Vrana et al., 1994), or with musteloids (weasels, skunks, and procyonids, see Schmidt-Kittler, 1981; Flynn and Nedbal, 1998; Bininda-Emonds et al., 1999). Paucity of shared derived characters in morphological studies (discussed in Dragoo and Honeycutt, 1997), insufficient data and taxon representation in molecular analyses, and ingrained traditional views biasing some studies by imposing constraints or inappropriate outgroups, are the principal causes of this large number of potential evolutionary scenarios, and of the lack of convincing support in favour of any one of them. Here, a

phylogenetic analysis based on twelve mitochondrial genes from thirty-eight carnivore species, including representatives of all families of interest, provides strong evidence for a sister relationship between the red panda and the skunks.

Both the red panda and the skunks belong to the order Carnivora, suborder Caniformia, traditionally broken down into seven extant families: the terrestrial Canidae (dogs, foxes), Ursidae (bears), Procyonidae (raccoon, kinkajou), and Mustelidae (weasels), and the three aquatic families forming the higher level group Pinnipedia: Phocidae (earless seals), Otariidae (eared seals), and Odobenidae (walruses) (Flynn et al., 1988; Corbet and Hill, 1991; McKenna and Bell, 1997). Mainly as a result of its unresolved phylogenetic position, the red panda is sometimes placed on its own, in an eighth family, the Ailuridae.

While the endangered status of the red panda has brought the attention of non-specialists to its phylogenetic position and uniqueness, the less popular skunks have also been at the centre of major taxonomical rearrangements. In contrast to the classic classification of skunks as the subfamily Mephitinae within Mustelidae (following Simpson, 1945), numerous studies (Arnason and Widegren, 1986; Ledje and Arnason, 1996b; Dragoo and Honeycutt, 1997; Flynn et al., 2000) indicate that they are only distantly related to other "true" mustelids. A newly defined higher level clade Musteloidea has emerged from these studies, placing sister families Mustelidae (*sensu stricto*, i.e., without skunks) and Procyonidae together as the sister clade to skunks, thereby elevated to their own family, the Mephitidae.

The inclusion of skunks among Mustelidae, often constrained in past analyses (Bryant et al., 1993; Bininda-Emonds et al., 1999), has precluded any early suggestion of their distant relationship to true mustelids, and, even more so of an affinity between them and the red panda. This relationship has been suggested only once (Flynn et al., 2000), based on an analysis of DNA sequence data from four genes, three mitochondrial and one nuclear. Although Flynn et al's analysis may be criticized on the basis of it showing disagreement between the trees obtained from maximum parsimony and maximum likelihood methods, its use of chimeric data, and the low nodal support values obtained (parsimony bootstrap, 54%; Bremer support, 3), the historical uncertainty and large number of already existing hypotheses are more likely the cause of the newly proposed relationship being largely overlooked.

These issues, and several other points of contention at various levels among Caniformia, motivated the collection of a large molecular dataset which combines the advantages of a large amount of data (twelve mitochondrial genes for a total of 10,842 nucleotides) and considerable taxon representation (38 species including 35 caniforms from all widely accepted and putative families). The data matrix was examined using a comprehensive array of phylogenetic methods, and relationships among the carnivore suborder Caniformia were investigated as presented in Chapter 3. Here I explore in more detail the relationships among mustelids, procyonids, and, especially, the skunks and the red panda.

Materials and methods

Data collection and phylogenetic analysis

Details of data collection protocols and of phylogenetic analysis performed on the complete data set using Bayesian, maximum likelihood (ML), and maximum parsimony (MP) methods are given in Chapter 3. A list of the species included is given in Table 3-1.

Statistical tests of topologies

Two different statistical tests (for a comparison and review, see Goldman et al., 2000) were used to compare previous hypotheses of red panda and skunks affinities to the sister relationship between the red panda and the skunk obtained from the phylogenetic analysis detailed in Chapter 3. Alternative hypotheses tested are listed in Table 4-1. They are: -1- the traditional family Mustelidae including skunks (cf. Simpson, 1945), -2- the sister relationship of the red panda and bears (Sarich, 1973; Wilson and Reeder, 1993), -3- the sister relationship of the red panda and the family Canidae (Ledje and Arnason, 1996a), -4- the red panda as a member of the family Procyonidae (Simpson, 1945; O'Brien et al., 1985; Dragoo and Honeycutt, 1997; McKenna and Bell, 1997; Nowak, 1999), -5- the red panda as the sister taxon to an Ursidae-Pinnipedia clade (Vrana et al., 1994), and -6- the red panda as the sister taxon to the three musteloid families, Mustelidae, Procyonidae, and Mephitidae (Schmidt-Kittler, 1981; Flynn and Nedbal, 1998; Bininda-Emonds et al., 1999).

The nonparametric test of Shimodaira and Hasegawa (SH test; Shimodaira and Hasegawa, 1999) was used to compare all competing hypotheses against the ML topology (see Chapter 3). For each test, the best tree that fits the constraint imposed by the hypothesis was tested against a set of three possible arrangements, taking into account the unresolved relationship between Ursidae, Pinnipedia and Musteloidea. SH tests were performed in PAUP* (Swofford, 2003), using 10,000 RELL pseudo-replicates (corresponding to the *posNPcnd* test of Goldman et al., 2000).

SH tests have been shown to be conservative (tending to accept the hypothesis that trees are equally probable explanations of the data when they are not; Goldman et al., 2000; Buckley, 2002). Consequently, the less conservative but more computationally intensive SOWH parametric test was only used in the case of a non-significant result using SH. The rationale for this is the empirical observation that *P*-values from SH tests are always larger than those obtained in comparable SOWH tests (Goldman et al., 2000; Buckley, 2002). Because of the computational requirements of the fully optimized test (*posPfud* of Goldman et al., 2000), a reduced taxa set consisting of ten representative species, one for each family, was used. The species that were selected are identified in Table 3-1. Prior to performing SOWH tests, the pruned taxa set was subjected to model testing, and maximum likelihood heuristic search and nonparametric bootstrapping using 1000 RELL pseudo-replicates.

Bayesian posterior probabilities

It has been suggested that the Bayesian posterior probability (BP) of a tree is a direct measure of the probability of it being the true tree given the observed data and the model, and, as such, can be used as a statistical measure (Rannala and Yang, 1996; Larget and Simon, 1999; Lewis, 2001). The BP of a tree is the number of times it is visited by the Markov chain over the total number of trees sampled. Two different BP values were compiled for each of the alternative hypotheses. The first more restrictive one is the BP of the best tree that satisfies the stated constraint [BP (ML solution) in Table 4-1]. This makes the Bayesian resemble ML-based tests, which compare fully defined topologies. The second Bayesian posterior probability computed is the total BP of all the trees sampled that satisfy the constraint [BP (constraint) in Table 4-1], which corresponds to a nodal support measure. Therefore, the BP of a specific best tree that maximizes the likelihood given a constraint [BP (ML solution)] is always lower than or equal to the BP of a group of trees that satisfy the same constraint [BP (constraint)].

Results

Phylogenetic analysis

The sister relationship of the red panda and the skunk was strongly supported in all types of analysis. Bayesian posterior probability was 100%, and parsimony bootstrap values were 73% or 77% when four-state or RY-coded data were used (Chapter 3). A likelihood

bootstrap analysis performed using the same smaller taxa set as for the SOWH test, comprising a single representative per family including Ailuridae and Mephitidae, was in agreement with these results, giving a support value of 79%. Overall, Bayesian, maximum likelihood, and RY parsimony analyses agreed on a single, strongly supported solution for the relationships between mustelid, procyonid, mephitid, and ailurid species (Figure 4-1). Parsimony analysis of the four-state coded data differed in placing the red panda-skunk clade as the sister to bears (Chapter 3, Figure 3-2).

Statistical tests of topologies

SH tests were used to compare previously suggested relationships of the red panda and the skunks to the mitochondrial DNA solution of a sister relationship between the two, as found here. The alternative trees are shown in Table 4-1. Each individual SH test evaluated the null hypothesis that the alternative tree was as probable as the *a posteriori* set of three best trees, identical except for the deeper relationships between Pinnipedia, Ursidae, and Musteloidea. Significance values obtained in the SH tests are shown in Table 4-1. Hypotheses 3 (red panda as the sister species to the Canidae) and 4 (red panda as a member of the Procyonidae) were strongly rejected. Hypotheses 1 (traditional position of skunks among the Mustelidae), 2 (red panda as a member of the family Ursidae) and 5 (placing the red panda as the sister species to an Ursidae-Pinnipedia clade) received a support value below $\alpha = 0.05$ but above the Bonferroni corrected $\alpha_c = 0.0083$. Finally, the hypothesis 6 where the red panda family, Ailuridae, is the sister family to the other three musteloid families Mustelidae, Procyonidae, and Mephitidae, received a more

significant support value of 0.3894. SOWH tests were performed for all hypotheses that received significant of marginally significant *P*-values in the SH test. All proposed hypotheses were rejected by the SOWH tests.

Bayesian posterior probabilities

All of the constrained ML trees corresponding to previously suggested relationships of skunks and red panda, presented in Table 4-1, had a BP value of less than 0.00003. This result corresponds to the complete absence of the six constrained trees from the set of 36 002 trees sampled by the Markov chains. The probability of trees simply satisfying the constraint were the same, except for a proportion of 0.0001 trees (5 out of 36 002) that were congruent with the constraint that the Ailuridae is the first family branching off at the base of the Musteloidea clade but differed from the ML topology in other areas of the tree. All alternative hypotheses are therefore rejected by the Bayesian posterior probability statistic used as a *P*-value.

Discussion

The phylogenetic analysis of twelve mitochondrial genes performed in this study strongly agrees with the hypothesis of Flynn et al. (2000) suggesting a close affinity between the red panda and the skunks. Their conclusions were based on the sequence of four genes (cyt b, 12S and 16S rRNAs, and transthyretin intron I) from 17 taxa (some chimeric), but

there was disagreement between methods and support was weak (parsimony bootstrap, 54%; Bremer support, 3). The additional taxon representation and data sampling used here provide additional strong support and an increased stability of the tree.

In addition to providing evidence for a close relationship of Ailuridae and Mephitidae, the use of statistical tests of topologies allowed for the rejection of competing hypotheses of phylogenetic relationships involving the red panda and the skunks. An agreement with the longstanding contention that skunks are part of the Mustelidae (Number 1 in Table 4-1, cf. Simpson, 1945) was never obtained in a phylogenetic analysis using several different methods and was rejected as a possibility by most statistical tests. This conclusion that skunks are distant from true mustelids, put forward by biochemical and molecular studies (Wurster and Benirschke, 1968; Arnason and Widegren, 1986; Wayne et al., 1989; Ledje and Arnason, 1996b; Dragoo and Honeycutt, 1997; Flynn et al., 2000) exemplifies the need for a re-evaluation of the morphological characters upon which the monophyly of the Mustelidae, including the skunks, is based. It has been suggested (Dragoo and Honeycutt, 1997) that there are in fact very few morphological characters to support this assertion, and that most may in fact be either homoplasies or plesiomorphies.

The two most frequently proposed positions of the red panda, as a member of Procyonidae, or as a relative to the family Ursidae, were also both rejected (Tests 2 and 4, respectively, in Table 4-1). This is in agreement with the criticisms that these hypothesized relationships are, in fact, due mostly to a superficial resemblance of the red

panda and the raccoon (Vrana et al., 1994) or to incomplete taxon sampling (Flynn et al., 2000) as the basis of a procyonid affinity, or to homoplasious characters related to similar habitat and diet in common with the giant panda, as the basis of an ursid affinity (Vrana et al., 1994; Flynn et al., 2000). Ewer's (1973) classification of the two pandas together as part of the Procyonidae therefore stems from a combination of the aforementioned systematic errors.

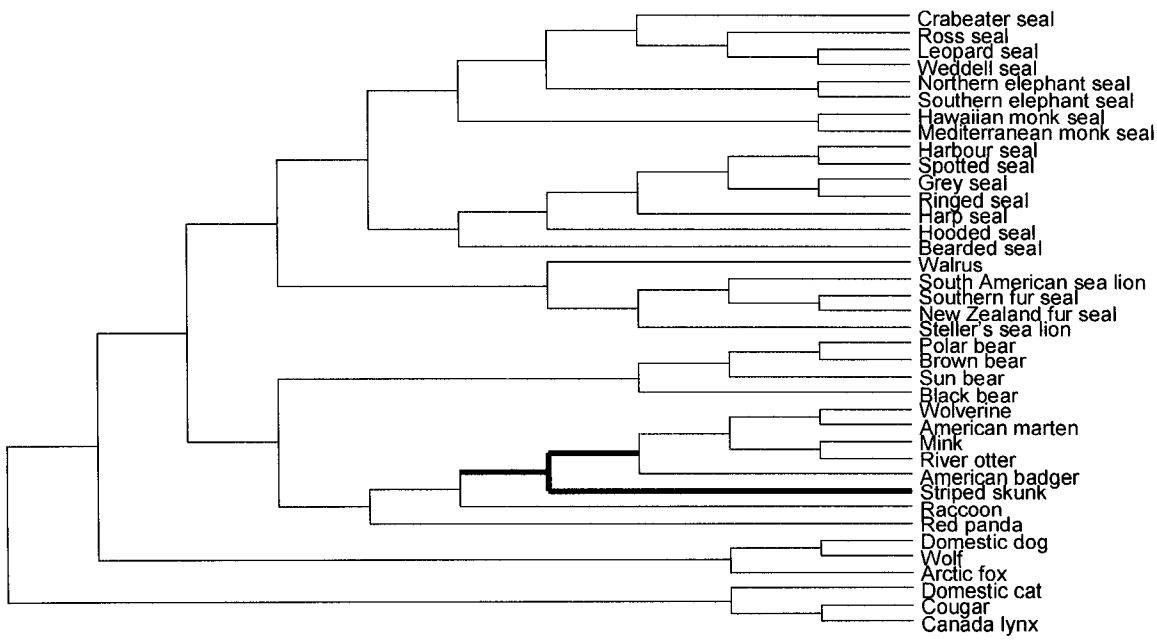
The hypothesis that the red panda is the sister species to the remaining of the Musteloidea clade (Mustelidae, Procyonidae, and Mephitidae; test 6 in Table 4-1), suggested by Schmidt-Kittler (1981), Flynn and Nedbal (1998), and Bininda-Emonds et al. (1999), generated more ambiguous results. The SH test (P=0.3894) accepted this hypothesis as an equally probable explanation of the data, while the SOWH test (P<0.01) strongly rejected that it is the true tree. Goldman et al. (2000) and Buckley et al. (2002) report empirical evidence that such different results are often seen where these two tests are performed. One possible explanation lies in the different formulation of the two tests. The SH test verifies that the competing topologies (often more than two, as is the case in this study) are equally good explanations of the data while the SOWH test verifies whether there exist significantly better topologies than the proposed one (Goldman et al., 2000). This may make the SH test more conservative than the SOWH. Also, the SH test is non-parametric, meaning that the data are directly resampled with replacement in each pseudo-replicate, while the data are simulated from the null topology and parameters in the parametric SOWH test, making it model-dependent. These different strategies may

have opposite impacts: an appropriate model may make the SOWH test more powerful, but an inaccurate one may make it prone to Type 1 errors (rejecting the null hypothesis when it should accept it; Buckley, 2002). Additionally, implementations of the SOWH test are computationally intense and here did not take advantage of the full data matrix, as I only included ten representative species. However, Bayesian posterior probabilities were in agreement with the SOWH test ($P<0.0003$). The Bayesian approach does have the advantage of using all of the available data, but is thought to suffer from the same sensitivity to model misspecifications as the SOWH test (Buckley, 2002).

Finally, this study strongly supports the close association of the red panda and the skunks as the sister families Ailuridae and Mephitidae, and their placement as sister to a Procyonidae-true Mustelidae clade, in a redefined Musteloidea. Further work will however be necessary to acquire more confidence in our rejection of the marginally supported alternative topology placing the Ailuridae as sister to the three other musteloid families. The addition of taxa to the present mitochondrial data set, as well as data from other markers may improve the confidence in the arrangement proposed here. Also, it will be of the utmost interest to see whether a re-evaluation of morphological and paleontological data will support this emerging new picture of the Musteloidea tree and the affinity between the red panda and the skunks, and bring new light to our understanding of their evolutionary history.

Table 4-1 Statistical tests of alternative positions of the red panda. Previously suggested hypotheses were treated as constraints, shown in bold branches, and used to generate an ML solution, illustrated. Tests were performed as outlined in the text. Bonferroni corrections give $\alpha_c = 0.0083$ (SH), and $\alpha_c = 0.0125$ (SOWH). * indicates significance.

Hypothesis - Constraint	SH test	SOWH test	BP (ML solution)	Total BP (constraint)
1. Skunks in Mustelidae	0.0434	0.01*	<0.00003*	<0.00003*



2. Red panda sister to Ursidae	0.0188	<0.01*	<0.00003*	<0.00003*

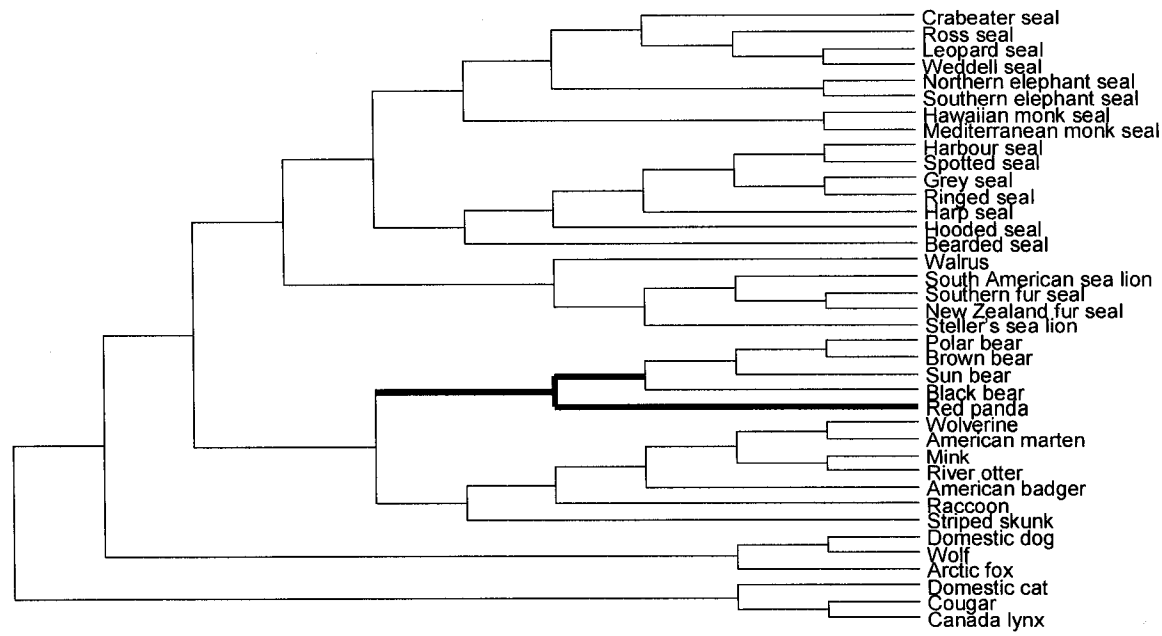
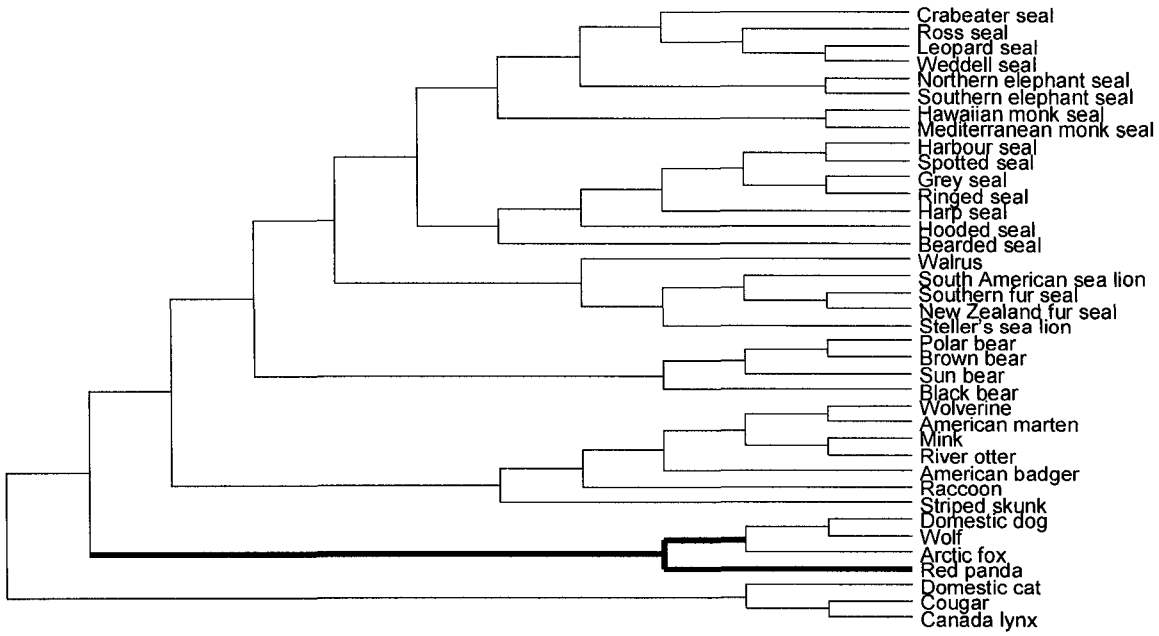


Table 4-1 Continued.

Hypothesis - Constraint	SH test	SOWH test	BP (ML solution)	Total BP (constraint)
3. Red panda sister to Canidae	0.0001*	-	<0.00003*	<0.00003*



4. Red panda in Procyonidae	0.0031*	-	<0.00003*	<0.00003*

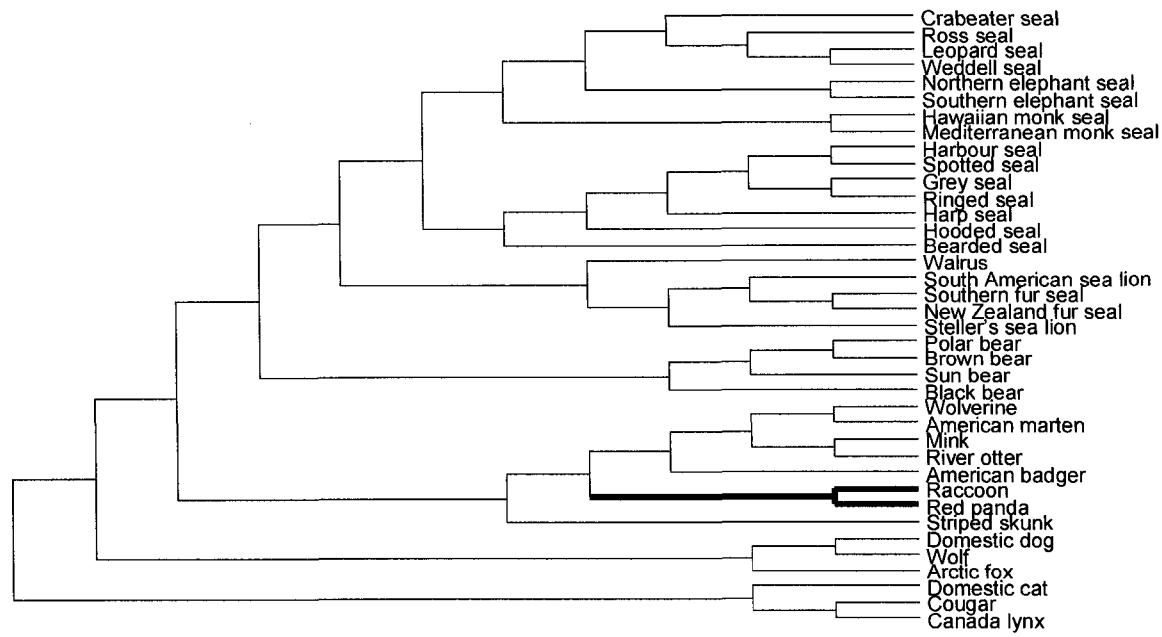
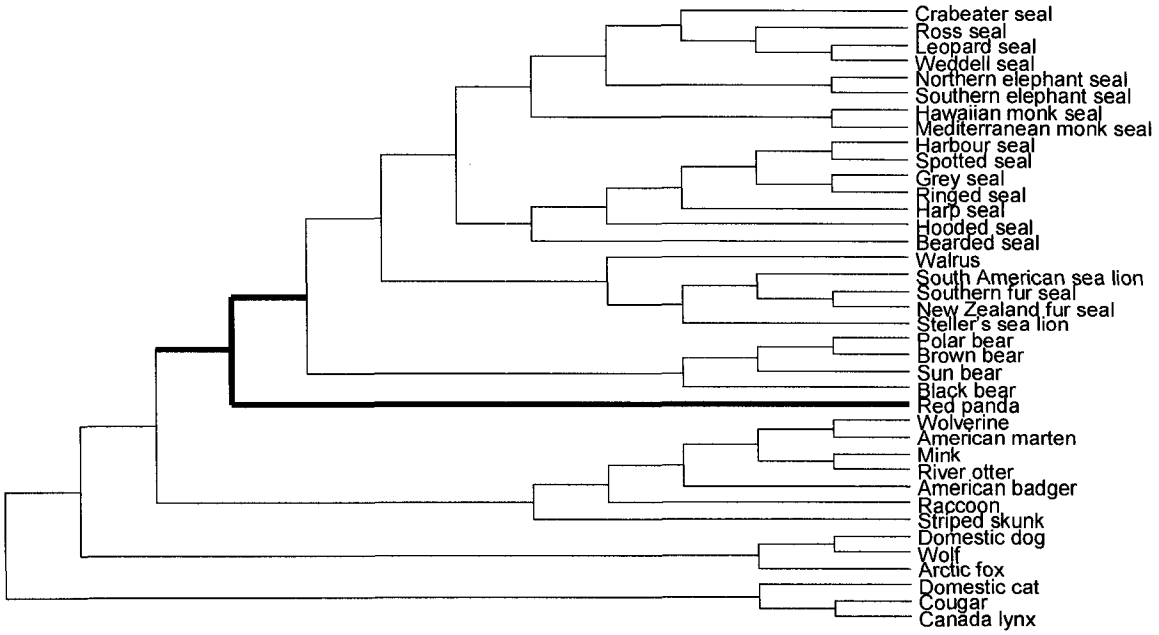
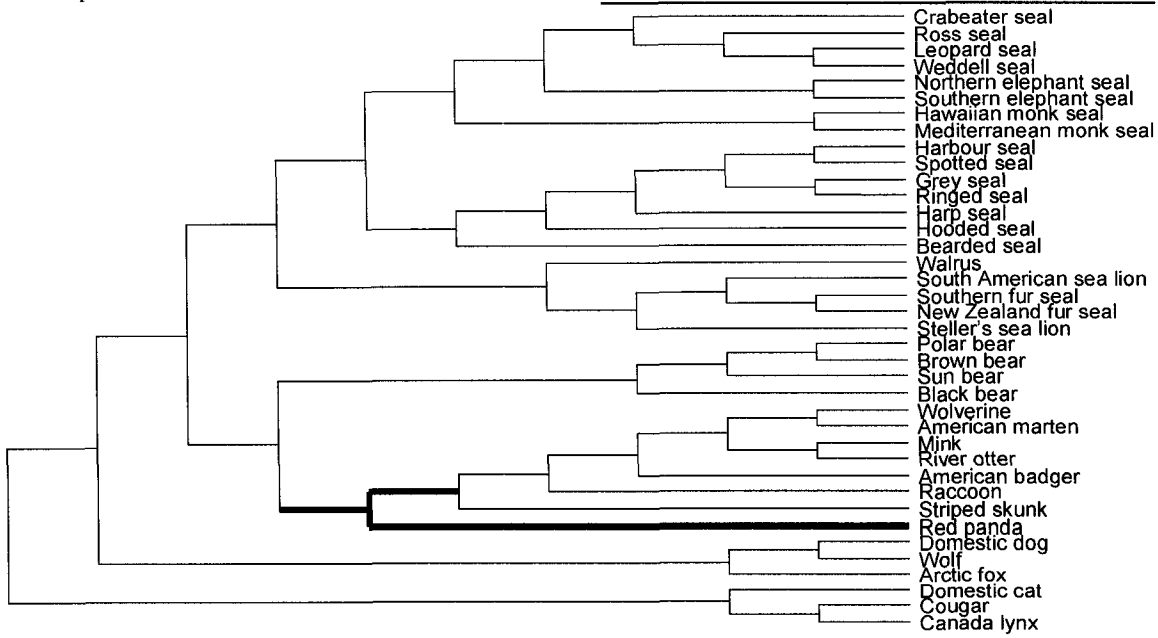


Table 4-1 Continued.

Hypothesis - Constraint	SH test	SOWH test	BP (ML solution)	Total BP (constraint)
5. Red panda sister to Ursidae-Pinnipedia clade	0.0099	<0.01*	<0.00003*	<0.00003*



6. Red panda sister to other musteloid families	0.3894	<0.01*	<0.00003*	0.0001*



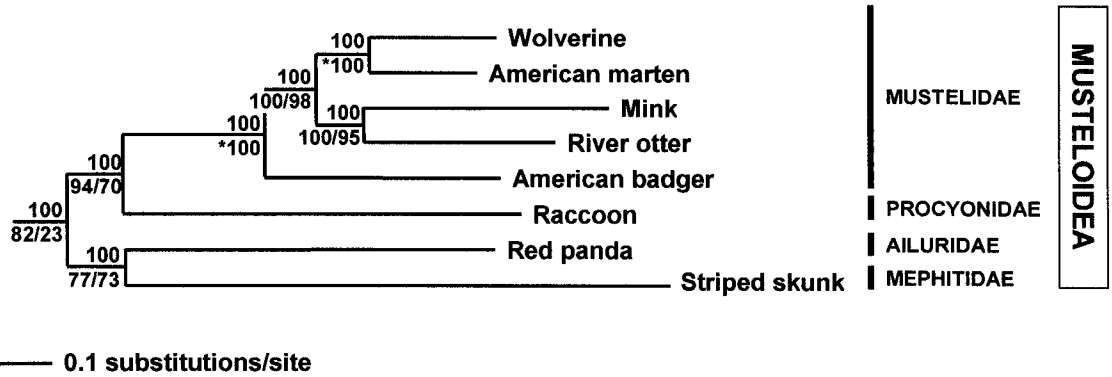


Figure 4-1 Maximum likelihood subtree of the Musteloidea clade, highlighting the relationships between Mustelidae, Procyonidae, Ailuridae and Mephitidae. The same topology was obtained using Bayesian, maximum likelihood, and RY maximum parsimony methods. Bayesian posterior probabilities are given above the branches, and parsimony bootstrap values (RY/four-state, * when equal) below.

Bibliography

Arnason, U., and B. Widegren. 1986. Pinniped phylogeny enlightened by molecular hybridizations using highly repetitive DNA. Molecular Biology and Evolution 3:356-365.

Bininda-Emonds, O. R. P., J. L. Gittleman, and A. Purvis. 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). Biological Reviews Of The Cambridge Philosophical Society 74:143-175.

Bryant, H. N., A. P. Russell, and W. D. Fitch. 1993. Phylogenetic relationships within the extant Mustelidae (Carnivora): appraisal of the cladistic status of the Simpsonian subfamilies. Zoological Journal of the Linnean Society 108:301-334.

Buckley, T. R. 2002. Model misspecification and probabilistic tests of topology: Evidence from empirical data sets. Systematic Biology 51:509-523.

Corbet, G. B., and J. E. Hill. 1991. A world list of mammalian species, 3rd edition. British Museum (Natural History) ; Oxford University Press, London, New York.

Dragoo, J. W., and R. L. Honeycutt. 1997. Systematics of mustelid-like carnivores. Journal of Mammalogy 78:426-443.

Ewer, R. F. 1973. The Carnivores. Weidenfeld and Nicolson, London.

Flynn, J. J., and M. A. Nedbal. 1998. Phylogeny of the Carnivora (Mammalia): Congruence vs incompatibility among multiple data sets. Molecular Phylogenetics and Evolution 9:414-426.

Flynn, J. J., M. A. Nedbal, J. W. Dragoo, and R. L. Honeycutt. 2000. Whence the red panda? Molecular Phylogenetics and Evolution 17:190-199.

Flynn, J. J., N. A. Neff, and R. H. Tedford. 1988. Phylogeny of the Carnivora. Pages 73-115 *in* The phylogeny and classification of the tetrapods. Vol. 2: Mammals (M. J. Benton, ed.). Clarendon Press, Oxford.

Goldman, N., J. P. Anderson, and A. G. Rodrigo. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49:652-670.

Larget, B., and D. Simon. 1999. Markov chains Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Molecular Biology and Evolution 16:750-759.

Ledje, C., and U. Arnason. 1996a. Phylogenetic analyses of complete cytochrome b genes of the order Carnivora with particular emphasis on the Caniformia. Journal of Molecular Evolution 42:135-144.

Ledje, C., and U. Arnason. 1996b. Phylogenetic relationships within caniform carnivores based on analyses of the mitochondrial 12S rRNA gene. Journal of Molecular Evolution 43:641-649.

Lewis, P. O. 2001. Phylogenetic systematics turns over a new leaf. Trends in Ecology & Evolution 16:30-37.

McKenna, M. C., and S. K. Bell. 1997. Classification of mammals : above the species level. Columbia University Press, New York.

Nowak, R. M. 1999. Walker's mammals of the world, 6th edition. Johns Hopkins University Press, Baltimore.

O'Brien, S. J., W. G. Nash, D. E. Wildt, M. E. Bush, and R. E. Benveniste. 1985. A molecular solution to the riddle of the giant panda's phylogeny. Nature 317:140-144.

Rannala, B., and Z. H. Yang. 1996. Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference. Journal of Molecular Evolution 43:304-311.

Sarich, V. M. 1973. The giant panda is a bear. Nature 245:218-220.

Schmidt-Kittler, N. 1981. Zur Stammesgeschichte der marderverwandten Raubtiergruppen (Musteloidea, Carnivora). Eclogae Geologicae Helvetiae 74:753-801.

Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16:1114-1116.

Simpson, G. G. 1945. The principles of classification and a classification of mammals. The American Museum of Natural History. Bulletins, 85, New York.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0b10. Sinauer Associates.

Vrana, P. B., M. C. Milinkovitch, J. R. Powell, and W. C. Wheeler. 1994. Higher level relationships of the arctoid Carnivora based on sequence data and "total evidence". Molecular Phylogenetics and Evolution 3:47-58.

Wayne, R. K., R. E. Benveniste, D. N. Janczewski, and S. J. O'Brien. 1989. Molecular and biochemical evolution of the Carnivora. Pages 465-494 *in* Carnivore behavior, ecology and evolution. Vol. 1 (J. L. Gittleman, ed.). Cornell University Press, Ithaca, NY.

Wilson, D. E., and D. M. Reeder. 1993. Mammal species of the world : a taxonomic and geographic reference, 2nd edition. Smithsonian Institution Press, Washington, D.C.

Wurster, D. H., and K. Benirschke. 1968. Comparative cytogenetic studies of the order Carnivora. Chromosoma 24:336-382.

Chapter 5

General discussion and conclusion

Rapid sequencing of mitochondrial DNA of carnivores

An important objective of this study was to take full advantage of modern molecular techniques such as PCR and high-throughput automated sequencing, as well as of more rapid computers and more efficient phylogenetic inference algorithms, to obtain a large molecular data set with which to revisit the phylogeny of caniform carnivores.

In Chapter 2, a methodology was developed to rapidly obtain sequence of homologous mitochondrial fragments across various carnivore species. A set of conserved primers was designed based on regions that appeared well conserved in the four species for which complete mitochondrial sequence was available at the outset of this study. These primers, amplifying large fragments of the mitochondrial genome, were used to obtain sequence from three species of bears. In turn, the new sequences were used to better assess conservativeness of newly designed internal primers.

The method has now been applied to a total of 29 other carnivores species (Davis et al., 2004; and Chapter 3, Table 3-2). As sequencing progressed, new, more specific primers

were designed to accommodate regions that were difficult to amplify or sequence in particular species. As a result, a collection of primers is now available to further facilitate the acquisition of additional data (Appendix A).

Phylogeny of Carnivora

A carnivore mitochondrial data set comprising 10 842 nucleotides, divided over 12 genes, for a total of 38 species, was used in a phylogenetic analysis to infer relationships between families of the suborder Caniformia of Carnivora (Chapter 3). Trees inferred using different methods of analysis were highly congruent and strongly supported (Figures 3-1 and 3-2). The monophyly of Arctoidea was clearly demonstrated, as well as, more importantly, it being constituted of three major monophyletic clades: Pinnipedia, Ursidae, and Musteloidea. Relationships within these three clades were also very well resolved and supported. Within Pinnipedia, the families Otariidae and Odobenidae were shown to form a clade, sister to Phocidae. Within Musteloidea, high support was also found for the sister relationship between true mustelids and procyonids, and between ailurids (red panda) and mephitids (skunks). Despite this high level of confidence obtained at most nodes, the tree was not completely resolved. Uncertainty remains about the relative position of the three major arctoid clades, effectively resulting in a polytomy (see Table 3-3).

At the beginning of this work, the only area of agreement that could be found upon a review of literature about the phylogeny of Carnivora, and Caniformia in particular, was a general acceptance of the monophyly of Arctoidea, a fact yet to be acknowledged in several classifications. Relationships between arctoid families have been so extensively debated, and hypotheses so varied and dependent on authors and type of data used, that they can only be reconciled in a polytomy (Chapter 1, Figure 1-1). Most relationships found here have been suggested at some point in time, however the evidence provided then was not strong enough to convincingly give a single hypothesis any preference. Therefore, what distinguishes the present study from prior publications on the phylogenetic relationships among Caniformia is the high level of support obtained, rather than the inferred relationships themselves. This is well demonstrated in the finding that Ailuridae (red panda) and Mephitidae (skunks) are sister taxa. This relationship has been suggested once (Flynn et al., 2000), based on an analysis of DNA sequence data from four genes, three mitochondrial and one nuclear. Perhaps because of the panoply of already existing hypotheses, the then newly proposed relationship did not appear to gain any prevalence over other possibilities.

Tests of competing hypotheses

Phylogenetic inference methods work at finding the best tree given a particular data set. A related question — how good is the best tree? — has always been more difficult to

answer. New statistical tests of competing evolutionary hypotheses allow an assessment of the significance of the results and of the explanatory power of a particular data set. Combined results of likelihood-based tests of topologies and Bayesian posterior probabilities showed that the mitochondrial data set significantly supports the sister relationship between the red panda and the skunks, over six previously suggested alternative hypotheses (Chapter 4). Conversely, it was clearly shown that the same data set does not have enough information to resolve the relationships among the three arctoid clades Pinnipedia, Ursidae, and Musteloidea (Chapter 3, Table 3-3).

A mitochondrial phylogeny

It is important to note that the phylogeny obtained here is based on a mitochondrial genealogy. The mitochondrial genome is a single hereditary unit whose gene tree may or may not reflect the species' evolutionary trajectory. As a gene tree, the mitochondrial phylogeny is a unique window into the evolutionary history of carnivores, a distinct line of evidence that must be compared to others equally valid. An important aspect of this study is that the agreement reached between different methods, as well as the high level of support obtained, is a good indication that the analysis may be converging toward the true mitochondrial tree. Therefore, most uncertainty within this unique line of evidence is removed, and problematic areas are clearly pinpointed for further improvement. This

should make comparisons with data of different sources a lot more significant and interesting.

Future work

The contributions and findings of this study can be built upon in two different ways.

Expansion of mitochondrial data set

Firstly, the methodology developed here can be used to expand the carnivore mitochondrial data set. Within Caniformia, the addition of key taxa such as the deep branching ursid *Ailuropoda melanoleuca* (giant panda), the putative mustelid genus *Mydaus* (stink badgers, shown to group with the skunks by Dragoo and Honeycutt, 1997), or the procyonid genus *Potos* (kinkajou, whose position makes the Procyonidae paraphyletic in the total evidence analysis of Vrana et al., 1994), may help in the resolution of the arctoid polytomy. Exhaustive taxon sampling within specific families should also provide insight into lower level, intra-familial relationships, as has been shown for the Phocidae (Davis et al., 2004). Moreover, the same strategy could be used to investigate relationships within Feliformia, a comprehensive molecular study of which is being called upon by recent work challenging traditional classifications (Gaubert and Veron, 2003; Veron et al., 2004).

Beyond phylogenetic questions restricted to the order Carnivora, deeper relationships between eutherian orders are debated (see Chapter 1) and the subject of a renewed interest stimulated by recent large-scale molecular studies (Madsen et al., 2001; Murphy et al., 2001a; 2001b; Arnason et al., 2002; Amrine-Madsen et al., 2003). The present data set has the potential to represent an important contribution to similar projects in the future.

Further exploration of carnivore phylogeny

A comprehensive view of the phylogeny of the Carnivora is dependent upon the exploration of different sources of data. The mitochondrial tree obtained here should be most useful in stimulating similar studies based on nuclear markers, Y-chromosome genealogies, and re-evaluations of the morphological data. Separate and combined analyses of these different markers are essential for revealing the full evolutionary history of the group.

Finally, once this is achieved, the phylogeny can be used to investigate evolutionary processes, and serve as a framework for investigations of other aspects of carnivore biology.

Bibliography

Amrine-Madsen, H., K.-P. Koepfli, R. K. Wayne, and M. S. Springer. 2003. A new phylogenetic marker, apolipoprotein B, provides compelling evidence for eutherian relationships. Molecular Phylogenetics and Evolution 28:225-240.

Arnason, U., J. A. Adegoke, K. Bodin, et al. 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. Proceedings of the National Academy of Sciences USA 99:8151-8156.

Davis, C. S., I. Delisle, D. B. Siniff, I. Stirling, and C. Strobeck. 2004. A phylogeny of the extant Phocidae inferred from complete mitochondrial DNA coding regions. Molecular Phylogenetics and Evolution (In press).

Dragoo, J. W., and R. L. Honeycutt. 1997. Systematics of mustelid-like carnivores. Journal of Mammalogy 78:426-443.

Flynn, J. J., M. A. Nedbal, J. W. Dragoo, and R. L. Honeycutt. 2000. Whence the red panda? Molecular Phylogenetics and Evolution 17:190-199.

Gaubert, P., and G. Veron. 2003. Exhaustive sample set among Viverridae reveals the sister-group of felids: the linsangs as a case of extreme morphological convergence within Feliformia. Proceedings of the Royal Society of London Series B-Biological Sciences 270:2523-2530.

Madsen, O., M. Scally, C. J. Douady, et al. 2001. Parallel adaptive radiations in two major clades of placental mammals. Nature 409:610-614.

Murphy, W. J., E. Eizirik, W. E. Johnson, Y. P. Zhang, O. A. Ryder, and S. J. O' Brien. 2001a. Molecular phylogenetics and the origins of placental mammals. Nature 409:614-618.

Murphy, W. J., E. Eizirik, S. J. O'Brien, et al. 2001b. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294:2348-2351.

Veron, G., M. Colyn, A. E. Dunham, P. Taylor, and P. Gaubert. 2004. Molecular systematics and origin of sociality in mongooses (Herpestidae, Carnivora). Molecular Phylogenetics and Evolution 30:582-598.

Vrana, P. B., M. C. Milinkovitch, J. R. Powell, and W. C. Wheeler. 1994. Higher level relationships of the arctoid Carnivora based on sequence data and "total evidence". Molecular Phylogenetics and Evolution 3:47-58.

Appendix A

Table A-1 Primer list. Primer names without parentheses are from the original work described in Chapter 2. Parentheses after primer names indicate that they were created as alternative primers, as indicated in Chapter 3.

Primer name[a]	Location[b]	Primer sequence (5' to 3')
mtDNA1 fragment	**16807-02382**	
Amplifying primers		
mtDNA1U	16807-00001	CAA ATG GGA CAT CTC GAT GGA CTA
mtDNA1L	02361-02382	CAG CTA TCA CCA GGC TCG TTA G
Internal primers		
mtDNA1U.int	00037-00058	TGG TGT CAT GCA TTT GGT AT
mtDNA1U.int2	00965-00985	CAC TGA AAA TGC CTA GAC GAG
mtDNA1U.int3	01427-01444	CAA ACT GGG ATT AGA TAC
mtDNA1L.int	01847-01862	GTG ACG GGC GGT GTG T
mtDNA1L.int2	01427-01444	GTA TCT AAT CCC AGT TTG
mtDNA2 fragment	**01915-03802**	
Amplifying primers		
mtDNA2U	01915-01940	GGA GAT AAG TCG TAA CAA GGT AAG CA
mtDNA2L	03782-03802	TCC TAC GAT GTT GGG TCC TTT
Internal primers		
mtDNA2U.int	02268-02286	CGA AAC CAG ACG AGC TAC C
mtDNA2U.int2	02628-02642	CTG GGC TAA TCT ATT
mtDNA2U.int3	03039-03058	GAA ATT GAC CTT CCC GTG AA
mtDNA2L.int	03283-03300	CCC CAG GGT AAC TTG TTC
mtDNA2L.int2	02835-02848	GTG TTT GCC GAG TT
mtDNA3 fragment	**03450-05409**	
Amplifying primers		
mtDNA3U	03450-03471	AAT CCA GGT CGG TTT CTA TCT A
mtDNA3L	05386-05409	ATC CTA TAT GGG CGA TTG ATG AGT
Internal primers		
mtDNA3U.int	03818-03831	CAA CCT ATC GCA GA
mtDNA3U.int2	04130-04146	CTA TCA GTC CTA CTA AT
mtDNA3U.int3	04637-04656	GAA ATA TGT CTG ACA AAA GA
mtDNA3L.int	05193-05209	ACG CCT TGT GTG ACT TC
mtDNA3L.int2	04738-04756	GGT AGC ACG AAG ATT TTT G
mtDNA3L.int3	04196-04217	GTC ATG TAG GGA AAA TTA GTC A
Alternative primers		
mtDNA3U(wf)	03485-03503	CCC AGT ACG AAA GGA CAA G
mtDNA3L(wf)	04929-04946	CTC GAA TCC AAT TCA GAT
mtDNA3U.int (af)	04041-04059	GAG CCT CAA ACT CCA AAT A
mtDNA3U.int2 (af)	04047-04066	CAA ACT CAA AAT ACG CCC TA
mtDNA3U.int2 (mu)	04088-04110	GCC CAA ACA ATC TCA TAC GAA GT
mtDNA3L.int2 (af)	04807-04824	CCA ACA TTT TCG GGG TAT
mtDNA3L.int3 (ly)	04249-04266	GCT CGA TTG GTT TCT GCT
mtDNA3L.int3 (mu)	04130-04152	CCG TTT ATT AGT AAG ACT GAT AG

Table A-1 Continued.

mtDNA4 fragment	**04813-061210**	
Amplifying primers		
mtDNA4U	04813-04837	CGA AAA TGT TGG TTT ATA CCC TTC C
mtDNA4L	06191-06210	TGC CAA GCT CTG TGG TGA AT
Internal primers		
mtDNA4U.int	05348-05361	AGG ACT AAA CCA AAC TCA AC
mtDNA4L.int	05914-05935	CAC TTA CTT AGG GCT TTG AAG G
Alternative primers		
mtDNA4L (as)	06094-06114	GGG AGA AGT AGA TTG AAG
mtDNA4U.int(ly)	05172-05188	CCC TTC CAC TTC TGA GT
mtDNA4L.int(wf)	05904-05923	GCT TTG AAG GCT CTT GGT C
mtDNA5 fragment	**0567-07776**	
Amplifying primers		
mtDNA5U	05967-05993	GGA CTG CAA GAA CAT ATC TCA CAT CAA
mtDNA5L	07751-07776	GGA GGA GGA CAT CCA TGT AGT CAT TC
Internal primers		
mtDNA5U.int	06797-06816	GCA ATA TCT CAA TAT CAA AC
mtDNA5U.int2	07172-07196	GAC ACA CGA GCT TAC TTC ACT TCA G
mtDNA5L.int	06980-06993	TGT CCG AAG AAT CA
Alternative primers		
mtDNA5U (pin)	05511-05536	TCA TCC ACA ACG ACA CTA TCA CTA TC
mtDNA5U (sk)	06054-06075	CCC ACG AAA ATT TAG TTA ACA G
mtDNA5L (sk)	07851-07871	GGT TAT GAC GTT GGC TTG AAA
mtDNA5U.int (as)	06618-06635	CCT CCA TAG TAG AAG CAG
mtDNA5U.int (ly)	06260-06280	TCA GCC ATT TTA CCT ATG TTC
mtDNA5U.int (mu)	06791-06816	CCT CCT GCA ATA TCA CAA TAC CAA AC
mtDNA5L.int (as)	07175-07193	AAG TAA AGT ATG CTC GTG T
mtDNA5L.int2 (as)	06626-06641	CTG CAC CTG CTT CTA C
mtDNA6 fragment	**07644-09496**	
Amplifying primers		
mtDNA6U	07644-07668	GCT CAT TTA TTT CAC TAA CAG CAG T
mtDNA6L	09475-09496	GGG CTA CAG CAA ATT CAA GGA T
Internal primers		
mtDNA6U.int	07849-07870	GGT TTC AAG CCA ATA TCA TAA C
mtDNA6U.int2	08334-08351	CAA GAA CTA AAG CCT GGA
mtDNA6U.int3	08861-08881	CCT TGA GAA GAA AAA TGA ACG
mtDNA6L.int	09241-09263	GGT AGA AAG TGA GCC AAG GAT GC
mtDNA6L.int2	08861-08881	CGT TCA TTT TTC TTC TCA AGG
mtDNA6L.int3	08268-08281	TCG TAA CTT CAA TAT CA
Alternative primers		
mtDNA6U (mu)	07070-07095	GAA CCT TTT GGC TAT ATA GGA ATA GT
mtDNA6U (mu2)	07157-07182	GTA GGG ATA GAC GTT GAC ACA CGA GC
mtDNA6U (ly)	07769-07794	CCT CCT CCA TAT CAC ACA TTT GAA GA
mtDNA6U (ses)	07415-07437	CTA TCA ATA GGA GCA GTC TTC GC
mtDNA6L (mu)	09553-09577	TAT GCA TGG GTT TGG TGG GTC ATT A
mtDNA6L (ly)	09780-09799	CCG TAT CGT AGT CCT TTT TG
mtDNA6L (ses)	09544-09564	GGT GAG TCA TTA GGT GTT ATC
mtDNA6U.int2 (as)	08478-08493	GCC ATT CCA GGA CGA C
mtDNA6U.int2 (mu)	08409-08431	CGC ATA CTA ATC TCG TCC GAA GA
mtDNA6U.int2 (rp)	08187-08206	GCC CTA CCT TCT CTA CGA AT

Table A-1 Continued.

mtDNA6L.int (mam)	09325-09346	CGG CTA GGG CTA TAG GTT GAA T
mtDNA6L.int3 (as)	08286-08304	GGT CTT CGT AGT CTG TGT A
mtDNA7 fragment	**09247-11135**	
Amplifying primers		
mtDNA7U	09247-09269	TTG GCT CAC TTT CTA CCT CAA GG
mtDNA7L	11109-11135	GTG GGG ATG ATG ATT TTT AGC ATT GTA
Internal primers		
mtDNA7U.int	09861-09877	CAC TCA AGC CTA GCA CC
mtDNA7L.int	10602-10620	GTA GTG CAA TTT CTA GGT C
mtDNA7L.int2	10060-10075	GTA AAG TAC ACG CCT A
Alternative primers		
mtDNA7U (ly)	09501-09521	CCA AGC CTA CGT TTT TAC CCT
mtDNA7L (ly)	11093-11116	GCA TTG CAG GAG GTT TAG GTT TTG
mtDNA7U.int (af)	09853-09874	GCC TTT TAT CAC TCA AGC CTA GC
mtDNA7U.int (r)	09920-09940	CCC TTA AAC CCC CTA GAA GT
mtDNA7U.int2 (as)	10535-10550	CAT AGG ATC AGC ACG C
mtDNA7L.int (as)	10539-10553	TAG GCG TGC TGA TCC
mtDNA7L.int (mu)	10605-10627	GGT AGG AGG AGT GCG ATT TCT AG
mtDNA7L.int (rp)	10770-10789	CCC TTA AAC CCC CTA GAA GT
mtDNA7L.int2 (as)	09900-09916	ATG CCT GTG GGT GGT CA
mtDNA7L.int2 (mu)	09867-09883	GGG GTT GGT GCT AGG CT
mtDNA8 fragment	**10770-12658**	
Amplifying primers		
mtDNA8U	10770-10792	CCA AAA CAA ATG ATT TCG ACT CA
mtDNA8L	12632-12658	GGT TCC TAA GAC CAA TGG ATT ACT TCT
Internal primers		
mtDNA8U.int	11192-11207	CAA CAG CTT ACA GCC T
mtDNA8U.int2	11725-11744	GTA AAA ATA CCT CTA TAC GG
mtDNA8L.int	12238-12254	ACA AAC AGT TCT CCG AT
mtDNA8L.int2	12041-12060	TAT TAA GGC TGT TGC TCC TA
mtDNA8L.int3	11725-11744	CCG TAT AGA GGT ATT TTT AC
Alternative primers		
mtDNA8U (alt)	10704-10724	GCA AGC CTA GCC TAC GAA TGA
mtDNA8U.int (as)	11437-11452	ACA TTC ACC GCC ACA G
mtDNA8U.int (mu)	11102-11123	AAT CTA CTA CAA TGC TAA AAA T
mtDNA8U.int (o)	11067-11088	CCA ATA CAT ACG GAA CAG ACT A
mtDNA8U.int2 (as)	12050-12065	CAG CCC TAA TAA TCG C
mtDNA8L.int (as)	12256-12274	ATG AAA ATG AGG CTA TTA C
mtDNA9 fragment	**12391-12414**	
Amplifying primers		
mtDNA9U	12391-12414	AAA CCA TCA TTC ACA CGA GAA AAC
mtDNA9L	13901-13925	GAG TTA GTA ATA GGG CTC AGG CGT T
Internal primers		
mtDNA9U.int	12674-12690	CTC CAA ATA AAA GTA AT
mtDNA9U.int2	13127-13140	GTA GGG ATC ATA TC
mtDNA9L.int	13883-13902	TTG GTA TAC GAC GTG TTG GC
mtDNA9L.int2	13364-13380	GGA TGA AGT CCG AAT TG
mtDNA9L.int3	13115-13135	GAT CCC TAC TCC TTC TCA GCC
mtDNA9L.int4	12799-12817	GGC GTA GGA GAC TGT AGT T

Table A-1 Continued.

Alternative primers		
mtDNA9U (af)	12475-12493	CTA GGA CCC ATC TAC TGT A
mtDNA9L (af)	13933-13956	CTG TAG GCA GCG GTT ATG GAT GTG
mtDNA9U.int (wf)	12457-12476	CTT AAC CCT AAA ATT ATT CT
mtDNA9U.int2 (af)	12799-12996	CAT GGT CTA TCA TAG AAT T
mtDNA9U.int2 (mu)	13409-13428	ACC CCT GTA TCA GCC CTA CT
mtDNA10 fragment	**13408-16671**	**Old: 13595-15517**
Amplifying primers		
mtDNA10U	13408-13431	TAC TCC TGT TTC AGC CCT ACT CCA
mtDNA10L	16647-16671	GCT GGT TTC TCG AAG CCT GGT GAT T
Internal primers		
mtDNA10U.int	13898-13915	ACC AAC GCC TGA GCC CTA
mtDNA10U.int2	14066-14081	GGA AGT ATT TTC GCA G
mtDNA10U.int3	14581-14600	TAT AAA GCC GCA ATC CCC AT
mtDNA10U.int4	15356-15374	CCA TCT TCT TTA TCT GCC T
mtDNA10U.int5	15733-15750	GAC TCA GAC AAA ATC CCA
mtDNA10L.int	16054-16067	AAT AGG CAT TGG CT
mtDNA10L.int2	15736-15754	GGA ATG GGA TTT TGT CTG A
mtDNA10L.int3	15422-15440	CCA ATG TTT CAT GTT TCT G
mtDNA10L.int4	14941-14959	CTT CTA AGC CTT CTC CCA T
mtDNA10L.int5	14605-14621	GGG TTT TTT AGT GAG GA
Alternative primers and fragments within mtDNA10		
mtDNA10U.int (wf)	13843-13864	CCT TAC AGG ATT CTA CTC CAA A
mtDNA10U.int2 (c)	14288-14307	CAC CGC CTC CCA TCA AAA AT
mtDNA10U.int3 (af)	15175-15193	TCT GCC TGA TGA AAC TTC G
mtDNA10U.int4 (af)	15457-15473	GCA ACC ATA GCC ACA GC
mtDNA10L.int (af)	16057-16077	TAG GAG TCA GAA TAG GCA TTG
mtDNA10L.int2 (af)	15883-15904	CGG GTT TGA TAT GTG GAG GGG T
mtDNA10L.int6 (c)	13949-13973	GAA TAT AAT TCG AGT GCT GTA GGC
mtDNA10 a fragment	**13595-15517**	
mtDNA10aU	13595-13617	GCA TTC TCA ACC TCA AGC CAA CT
mtDNA10aL	15491-15517	CTC AAA ATG ATA TTT GTC CTC ATG GTA
mtDNA10aU(mu)	13818-13839	GCC TAG CAC TTA CAG GAA TAC C
mtDNA10aL(mu)	14779-14799	GGC TAC TGA GCA GTA TCC TGA
mtDNA10 b fragment	**15278-00125**	
Amplifying primers		
mtDNA10bU	15278-15302	CCT TCT CAT CAG TAA CCC ACA TCT G
mtDNA10bU (alt)	15405-15430	CGG CTC CTA CAC ATT CAC AGA GAC AT
mtDNA10bL	00103-00125	ATT TGA CTG CGT CGA GAC CTT TA
mtDNA10bL (alt)	00085-00108	CCT TTA CGG TCA TAG CTG AGT CAT
mtDloop fragment	**16090-00987**	
Amplifying primers		
mtDloopU	16090-16120	CTA ACA TGA ATC GGA GGA CAA CCA G
mtDloopL	00965-00987	GGC TCA TCT AGG CAT TTT CAG TG

Table A-1 Continued.

Internal primers		
mtDloopU.int	00661-00673	CAA ACC CCC CTT A
mtDloopL.int	00700-00723	CAA GAC AAC CAT AAA TGT GCA TAA
mtDloopL.int2	00188-00208	TAT GTC CTG CGA CCA TTG ACT

[a] The letter codes U and L refer respectively to the upper (U, upper, heavy, coding for most genes) and lower (L, light, non-coding for most genes, except for ND6 and Gln, Ala, Asn, Cys, Tyr, Ser, Glu, and Pro tRNAs) strand of the mitochondrial DNA molecule.

[b] Location corresponds to nucleotide numbers from the harbor seal sequence. (*Phoca vitulina*, NC_001325, Arnason and Johnsson, 1992)

Bibliography

Arnason, U., and E. Johnsson. 1992. The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina*. Journal of Molecular Evolution 34:493-505.