Molecular Phylogenetics and Evolution
Vol. 13, No. 1, October, pp. 82–92, 1999
Article ID mpev.1999.0637, available online at http://www.idealibrary.com on IDEAL®

Rapid Radiation Events in the Family Ursidae Indicated by Likelihood Phylogenetic Estimation from Multiple Fragments of mtDNA

Lisette P. Waits,* Jack Sullivan,† Stephen J. O'Brien,‡ and R. H. Ward§

*Department of Fish and Wildlife Resources and University of Idaho, †Biological Sciences, Moscow, Idaho 83844-1136; ‡Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21702-1201; and §Department of Biological Anthropology, University of Oxford, Oxford, England

Received July 15, 1998; revised December 15, 1998

The bear family (Ursidae) presents a number of phylogenetic ambiguities as the evolutionary relationships of the six youngest members (ursine bears) are largely unresolved. Recent mitochondrial DNA analyses have produced conflicting results with respect to the phylogeny of ursine bears. In an attempt to resolve these issues, we obtained 1916 nucleotides of mitochondrial DNA sequence data from six gene segments for all eight bear species and conducted maximum likelihood and maximum parsimony analyses on all fragments separately and combined. All six single-region gene trees gave different phylogenetic estimates; however, only for control region data was this significantly incongruent with the results from the combined data. The optimal phylogeny for the combined data set suggests that the giant panda is most basal followed by the spectacled bear. The sloth bear is the basal ursine bear, and there is weak support for a sister taxon relationship of the American and Asiatic black bears. The sun bear is sister taxon to the youngest clade containing brown bears and polar bears. Statistical analyses of alternate hypotheses revealed a lack of strong support for many of the relationships. We suggest that the difficulties surrounding the resolution of the evolutionary relationships of the Ursidae are linked to the existence of sequential rapid radiation events in bear evolution. Thus, unresolved branching orders during these time periods may represent an accurate representation of the evolutionary history of bear species. © 1999 Academic Press

INTRODUCTION

The Ursidae, a relatively small carnivoran family of eight extant species, is classified within the Arctoidea and is thought to have originated some 15–20 million years ago (MYA) (Thenius, 1959; Kurtén, 1968). The earliest unequivocal bear, *Ursavus,* appears in the early Miocene as a descendant of the "bear-dogs" (Hemicyoninae; Crusafont and Kurtén, 1976; Martin, 1989), a transitionary group of late Oligocene carnivores. All extant bears, plus the extinct subfamily Agriotheriinae, are believed to be derived from the diversification of *Ursavus* during the Miocene (Kurtén, 1968; Thenius, 1982). Given this rapid radiation, it is not surprising that the taxonomic status of bears has been enigmatic and controversial. One long standing problem concerns the placement of the giant panda (*Ailuropoda melanoleuca*): while molecular data generally support inclusion of the giant panda within the Ursidae (O'Brien *et al.,* 1985; Nash and O'Brien, 1987; Goldman *et al.,* 1987; Wayne *et al.,* 1989; Hashimoto *et al.,* 1993), this is still being challenged at the molecular (Tagle *et al.,* 1988; Zhang and Shi, 1991) and morphological (Kitchener, 1994) levels.

In addition to the controversial taxonomic placement of the giant panda, the classification of the seven other bear species is unclear (Table 1). The spectacled bear (*Tremarctos ornatus*) is generally considered basal to the rest of the bears and has been alternatively placed in its own genus or subfamily. Fossil data suggest that the remaining six species (the ursine bears) originated in Eurasia during the early Pliocene (Kurtén, 1968), but the evolutionary relationships among these species are unresolved; paleontological (Kurtén, 1968; Thenius, 1982), cytological (Nash and O'Brien, 1987), immunological, DNA hybridization, and isozyme (O'Brien *et al.,* 1985; Goldman *et al.,* 1987) data have proven inconclusive. Analyses of mitochondrial DNA (mtDNA) sequence data have been problematic, with differing placement of the ursine species depending on the analysis and particular specimen used (Zhang and Ryder, 1993, 1994; Talbot and Shields, 1996a). As a result, taxonomic classifications (Table 1) range from placing each ursine species in its own genus (Eisenberg, 1981) to placing all six ursine species in the genus *Ursus* (Nowak, 1991).

Molecular sequence data have been particularly useful in the resolution of ambiguous phylogenetic relationships; however, the degree of phylogenetic resolution afforded by molecular sequences is critically dependent on several factors, including the time since divergence

1055-7903/99 $30.00

TABLE 1

Taxonomy of the Ursidae

Common name	Ewer (1973)	Eisenburg (1981)	Wozencraft (1989)	Nowak (1991)
Giant panda	*Ailuropoda melanoleuca*	*A. melanoleuca*	*A. melanoleuca*	*A. melanoleuca*
Spectacled bear	*Tremartos ornatus*	*Tr. ornatus*	*Tr. ornatus*	*T. ornatus*
Asiatic black bear	*Selenarctos thibetanus*	*S. thibetanus*	*Ursus thibetanus*	*U. thibetanus*
Sloth bear	*Melursus ursinus*	*M. ursinus*	*M. ursinus*	*Ursus ursinus*
Sun bear	*Helarctos malayanus*	*H. malayanus*	*H. malayanus*	*Ursus malayanus*
Polar bear	*Thalarctos maritimus*	*Th. maritimus*	*Ursus maritimus*	*U. maritimus*
Brown bear	*Ursus arctos*	*U. arctos*	*U. arctos*	*U. arctos*
American black bear	*Euarctos americanus*	*Ursus americanus*	*U. americanus*	*U. americanus*

and the average rate at which substitutions accumulate (Avise *et al.,* 1983), base frequencies (Lockhart *et al.,* 1994), distribution of rates among sites (e.g., Gaut and Lewis, 1995), and the nature of the topology being estimated (e.g., Yang, 1994a). Furthermore, the validity of the assumptions made by any phylogenetic analyses (either explicit or implicit) can impinge upon phylogenetic reliability and accuracy. In this paper, we use sequence fragments from several different mitochondrial genes to estimate the mtDNA gene tree for the extant bear species using maximum-likelihood analyses under an objectively determined model.

MATERIAL AND METHODS

Samples

Samples were obtained for a minimum of three individuals from each of the eight bear species, excluding the enigmatic Red panda, *Aliurus fulgens.* For each species, total genomic DNA was isolated from a = cell line (Sambrook *et al.,* 1989), b = serum (Ward *et al.,* 1991), c = blood sample (Vardenplas *et al.,* 1984), as follows: Giant panda (*Ailuropoda melanoleuca*)[a], spectacled bear (*Tremarctos ornatus*)[a], sloth bear (*Melursus ursinus*)[ab], American black bear (*U. americanus*)[c], Asiatic black bear (*U. thibetanus*)[c], sun bear (*U. malayanus*)[ab], European brown bear (*U. arctos arctos*)[c], Montana brown or grizzly (*U. arctos horribilis*)[c], Japanese brown (*U. arctos hokkaido*)[a], Kodiak brown bear (*U. arctos middendorffi*)[c], and polar bear (*U. maritimus*)[ab]. All samples from endangered species were obtained with appropriate CITES documentation.

Amplification and Sequencing

Sequences were obtained from the products of PCR amplification (Saiki, 1990) using solid-phase sequencing (Hultman *et al.,* 1989). Amplification was carried out in 25-μl volume reactions with 100–400 ng of genomic DNA using a *Thermus aquaticus* (*Taq*) polymerase kit (Perkin–Elmer–Cetus). The PCR temperature profile consisted of 30 cycles of 1 min at 94°C, 1 min at 42–50°C, and 1 min at 72°C followed by a final 5-min extension at 72°C. Published primer sets were used for the control region (Ward *et al.,* 1991), 16S rRNA (Cunningham *et al.,* 1992), and cytochrome *b* (Kocher *et al.,* 1989).

Primers were designed for NADH-5, NADH-4, and COII by aligning published mammalian sequences and choosing conserved regions. Primer sequences are listed below with reference to the 3′ nucleotide position on the harbor seal mitochondrial DNA reference sequence (Árnason and Johnsson, 1992), where L and H refer to the light and heavy strand, respectively. NAD5: L12673 (5′GGTGCAACTCCAAATAAA AGTA-3′), H12977 (5′-AGAATTCTATGATGG ATCATGT-′3′); NAD4: L12035 (5′-ATTCTCATCCAAACGCCATGAAG-3′), H12514 (5′-TTATTAGATTCACAATCTAAT-3′); and COII: L7930 (5′-TACATAACTTTGTCAAAGTTAA-3′), H8666 (5′-TCT CAATCTTTAACTTAAAAGGTT-3′); plus an internal sequencing primer L8146 (5′-GACGCACAAGAAG-TAGAGAC-3′). A giant panda-specific heavy strand primer was also developed for NADH-4—H12531 (5′-TAAAAAGATTTAACGTTTTAT-3′).

The control region was first sequenced for a minimum of three individuals per bear species; then the two most divergent individuals within each species were sequenced for the remaining five regions. However, due to amplification and sequencing difficulties, only one of the spectacled bear samples gave adequate sequence for COII and NADH-4. Hence, this species is represented by only a single individual in the analyses of these two regions and for all regions combined. The primary outgroup taxon was defined by sequence data obtained from the single harbor seal (*Phoca vitulina*) for all six regions.

Sequence Alignment and Domain Identification

All sequences for the four protein coding regions aligned perfectly, with no insertions or deletions. The 16S rRNA sequences were aligned by eye to the bovine secondary structure (Gutell and Fox, 1988), resulting in an aligned sequence of 485 nts, corresponding to positions 2911 to 3391 of the harbor seal reference, including four indels. Control region sequences were aligned using the CLUSTAL V program (Higgins and Sharp, 1989) and adjusted manually to preserve align-

ment of regions believed to be homologous throughout the entire Class Mammalia (Saccone *et al.,* 1991). This yielded an aligned sequence of 352 nts, corresponding to positions 16,388 to 16,711 of the harbor seal reference (Árnason and Johnsson, 1992). This sequence (see Appendix) includes 15 indels: 11 single-nt deletions (2 in the seal sequence, 8 in bear sequences and 1 in both bears and seal), 3 four-nt deletions in the seal sequence, plus one larger indel of repetitive sequence, which ranged in length from 5 to 59 nts depending on the species (positions 74 to 132).

Phylogenetic Analyses

Phylogenetic analyses were performed using a test version of PAUP* (4.0d61a; written by D. L. Swofford) and were restricted to aligned sequences, with all indels from the 16S rRNA and control region sequences omitted. We used an iterative search strategy proposed by Swofford *et al.* (1996), in which an initial parsimony search (equal weights, BANDB) is conducted to acquire a set of topologies. These were then used to evaluate the relative fit of alternative models and estimate model parameters for use in a heuristic search (random input order with 10 replicates, TBR branch swapping) under the likelihood criterion (see Frati *et al.,* 1997; Sullivan and Swofford, 1997; Sullivan *et al.,* 1997 for examples of this methodology). Four models of nucleotide substitution were examined: Jukes-Cantor (JC; Jukes and Cantor, 1969), Kimura two-parameter (K2P; Kimura, 1980), Hasegawa–Kishino–Yano (HKY85; Hasegawa *et al.,* 1985), and general time-reversible (GTR; Yang, 1994a). In addition, 4 models of among-site rate heterogeneity were examined: equal rates assumed at all sites, invariable sites plus a uniform rate at variable sites (I; e.g., Hasegawa *et al.,* 1985), rates at all sites assumed to follow a gamma distribution (G; Yang, 1994b), and a mixed-distribution model that combines the invariable sites and gamma-distributed rates model (I + Γ; Gu *et al.,* 1995; Waddell and Penny, 1996). In the models that employed a gamma distribution, we used the discrete approximation of the continuous gamma, with eight rate categories. Thus, a total of 16 alternative models of evolution were examined, each of which is a special case of the most general and parameter-rich GTR + I + Γ (see Swofford *et al.* [1996] and Sullivan and Swofford [1997] for description of models and comparisons of models assumptions). Nodal support was estimated using bootstrap analyses (Felsenstein, 1985). Five hundred replicates were conducted for the parsimony bootstrap analyses and 100 replicates were conducted for the likelihood bootstrap analyses. In the likelihood bootstrap analyses, the HKY + I + Γ model was used with transition ratio, proportion of sites invariable, and gamma-distribution shape parameter set to the values estimated from the original analyses, and base frequencies were recalculated for each replicate.

Phylogenetic trees were estimated separately for each of the six regions sequenced and for the combined data set of 1916 nts. The issue of whether or not to combine data from different genes is complex. Although we acknowledge that universally applicable tests for combinability do not exist (e.g., Sullivan, 1996), we agree with those who have suggested that the best estimate of phylogeny will not always derive from a combined data approach (e.g., Bull *et al.,* 1994; Mason-Gamer and Kellog, 1996), rather than those who view a total evidence approach as the only valid method of analysis (e.g., DeSalle and Brower, 1997); that nature is complex precludes such a dogmatic approach as the latter. Therefore, we tested the optimal trees for each of the fragments with the optimal tree estimated from the combined data set using the Kishino–Hasegawa Test (Kishino and Hasegawa, 1989). With the exception of the control region data set, the combined tree was not significantly worse than the optimum tree for any particular fragment. This suggests that, as would be expected for linked mitochondrial genes, each of the single region trees is an independent estimate of the mitochondrial gene tree and a combined analysis is warranted. Two individuals per species were included in each analysis, with the exception of the spectacled bear and the outgroup taxon (harbor seal), which were represented by only one individual. Alternate topologies were tested for significant differences using Kishino and Hasegawa's test.

RESULTS

Nucleotide Variability

The distribution of nucleotides observed to vary within bears for each of the six regions is summarized in Table 2 and Fig. 1. Overall, the proportion of nucleotides observed to vary ranges from 20% for 16S rRNA data to 46% for the NADH-5 (Table 2). As is commonly seen, when evaluated in terms of codon position (Fig. 1A), the nucleotide variability of the protein coding genes appears inversely related to the degree of selective constraints. For all four protein coding regions, the third position sites exhibited uniformly high levels of variation, with 72 to 80% of the nucleotide positions observed to vary among taxa, whereas only 21% of all first position sites and only 6% of all second position sites were observed to vary. The heterogeneity in overall nucleotide variability between protein-coding regions (Table 2) is due to differences in variability at first and second position sites, coupled with the differences in variability at second position sites. This likely results from the fact that each protein coding region has been subject to different selective constraints.

The proportion of nucleotides at which substitutions were observed was also examined separately for single-

TABLE 2

Summary of Model Parameters for the Fragment Data When Analyzed Separately and for the Combined Data

	COII	Cyt b	NADH-4	NADH-5	16S	CR	Comb.
No. of sites	453	255	223	243	480	260	1916
No. variable	162	94	84	111	94	97	641
Base Freq.	0.31457	0.27843	0.32332	0.35587	0.32323	0.24731	0.3087
	0.23245	0.27373	0.31278	0.25377	0.21698	0.26154	0.2497
	0.18256	0.16176	0.09081	0.09622	0.20750	0.18885	0.1650
T-ratio	45.606	8.702	14.804	36.618	5.804	11.362	15.886
a	Infinity	5.391	1.187	0.963	0.302	0.2837	1.440
p_{inv}	0.611	0.560	0.477	0.418	0.484	0.2435	0.583
p^1	0.248	0.424	0.090	0.872	0.256	0.0023	—

stranded and double-stranded regions of the 16S rRNA gene and for the variable (V) and conserved (C) domains of the control region (Fig. 1B). Considerable heterogeneity of nucleotide variation was observed within the 16S rRNA gene segment; three times more sites were observed to vary in the single-stranded loops than double-stranded stems. The location of the two variable domains and conserved domains in the control region is consistent with the location of the hypervariable and conserved segments postulated to occur in the mtDNA control region of all mammals (Saccone *et al.,* 1991). The heterogeneity between these domains is substantial, with five times more sites observed to vary in variable domains than in conserved domains. However, even the variable domains exhibit a smaller percentage of nonconstant sites than the third position sites of the protein-coding regions (Fig. 1). Hence, at least in bears, there is little support for the presumption that the average rate of interspecific nucleotide substitution in the control region is substantially greater than that of protein-coding regions. Overall, these data indicate that the probability of nucleotide substitution varies substantially within the 5′ end of the control region.

Phylogenetic Analyses

The best fit of the alternative models that we examined was the HKY + I + Γ model (Fig. 2). As expected, the most complex and parameter-rich model (GTR + I + Γ) fit the data somewhat better but this improvement in fit (as assessed by the χ^2 approximation of the null distribution) was only marginal ($\chi^2_{[4]} = 7.89$; $0.1 > P > 0.05$). The parameter estimates of this model for each of the fragments and for the combined data set are presented in Table 2.

Figure 3 summarizes the maximum likelihood topologies derived from the independent phylogenetic analysis of each single-region data set. The optimum trees for each region are different, and furthermore, none of the single-region trees was identical to the best tree for the combined data (Fig. 4). Only for the control region sequences, however, were the optimum topologies significantly better than the best trees for the combined data. Hence, despite their different appearance, the single-region topologies are not significantly different from the combined-data topologies (Fig. 4); and these single-region data sets fail to provide adequate discriminatory power to exclude alternative evolutionary hypotheses.

Three optimal trees differing only in placement of lineages in the brown bear/polar bear clade were found in the likelihood analysis of the combined data. These topologies indicate that polar bears may have arisen from within *U. arctos,* and the sun bear is sister taxon to this clade. There is weak support for a sister taxon relationship between the Asiatic and the American black bears, and the sloth bear is the basal ursine bear. The spectacled bear is sister to the ursine bears, and the giant panda is the most basal bear.

DISCUSSION

Perhaps one of the most interesting questions currently facing molecular systematists is how best to handle heterogeneous data from multiple genes. This issue will become increasingly important as data sets become increasingly large. In this case, we have sequences from six regions of the mitochondrial genome with potentially different rates and modes of evolution; thus, we were forced to use a relatively complex heterogeneous-rates model. However complex, this model (HKY + I + Γ) still treats all fragments as if they are evolving under the same Markov process. The validity of this assumption can be tested by allowing a unique HKY + I + Γ for each fragment and comparing the sum of the log likelihoods for each (lnL = −7455.988) to the score calculated for the combined data (lnL = −7675.316). In this case, there is a significant improvement in fit associated with allowing a unique process for each fragment ($\chi^2_{[24]} = 438.655$; $P < 0.001$); however, there is no current implementation of maximum likelihood that allows the partitioning of data in such a way during a tree search. Analyses using a better fit model are expected to provide a consistent estimate of phylogeny under a broader array of conditions than would be expected for analyses using a

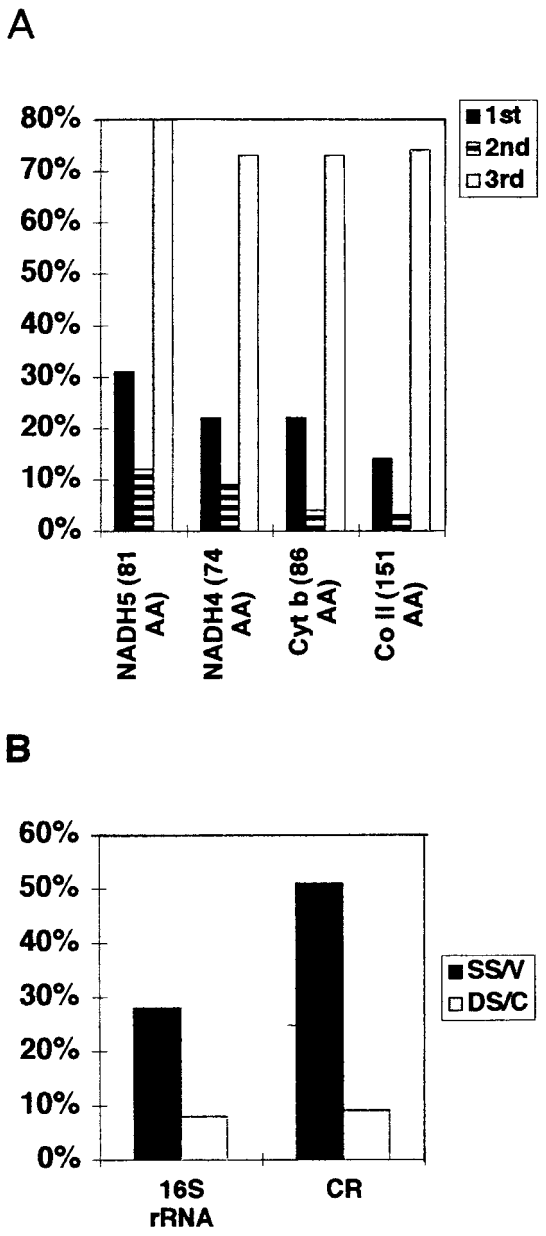

FIG. 1. Proportion of nucleotides observed to vary among the Ursidae mitochondrial DNA sequences. (A) Nucleotide variability observed in the four protein coding regions by first, second, and third codon position. Percentages represent the number of nucleotides observed to vary divided by the total number of nucleotides sequenced. NADH-5, NADH dehydrogenase subunit 5; NADH-4, NADH dehydrogenase subunit 4; Cyt b, cytochrome *b;* COII, cytochrome oxidase II. (B) Observed nucleotide variability in different domains of 16S rRNA and the control region. Single-stranded (SS), double-stranded (DS) for rRNA and conserved (C), variable (V) (Sacconne *et al.,* 1991) for control region. Insertions and deletions were omitted from these analyses.

simplified model but this may incur a cost in efficiency. That is, for sequences of a finite length, one may be no less likely to estimate the true tree using a simplified model rather than a complex, statistically better model. Indeed, there is evidence that maximum likelihood is robust to violations of the rate-heterogeneity model, as long as among-site rate variation is accommodated in some manner (J. Sullivan and D. L. Swofford; unpublished data). The trade-off between model complexity and performance of phylogenetic analyses is currently an active area of research.

Phylogenetic Analysis of the Combined Data Set

Given the ambiguities associated with analysis of the single-region data sets, our assessment of bear phylogeny will be based on topologies estimated from phylogenetic analyses of the combined data set. These analyses identified three optimal topologies (Fig. 4) that differed only in the placement of lineages within the brown bear and polar bear clade. There is high bootstrap support (93% ML, 100% MP) for the brown bear/polar bear clade, and this result is consistent with earlier conclusions of a close evolutionary relationship between the polar bear and the brown bear based on fossil data (Kurtén, 1964), protein allozyme data (Goldman *et al.,* 1987), mtDNA sequence data (Shields and Kocher, 1991), and the presence of fertile F1 hybrids (Gray, 1971). A paraphyletic relationship of brown bears and polar bears has been suggested based on mtDNA analyses of a more extensive geographic sampling of brown bears and polar bears (Cronin *et al.,* 1991; Talbot and Shields, 1996b; Waits *et al.,* 1998). In contrast, the suggestion of Zhang and Ryder (1993) that the polar bear is a recent offshoot of the Tremarctine lineage is rejected by our analysis.

The combined data analysis provided weak to marginal bootstrap support (30% ML, 61% MP) for the grouping of the sun bear as sister taxon to the brown bear/polar bear clade. This grouping was also suggested based on mtDNA sequence data from the entire Cyt b gene (Talbot and Shields, 1996a), but Zhang and Ryder (1994) suggested the Asiatic black bear is sister taxon to a brown/polar/sunbear clade based on analyses of 1185 bp of mtDNA sequence data from Cyt b, control region, 12S rRNA, and tRNA gene fragments. Some interpretations of the fossil data suggest that the Asiatic black bear and the American black bear are sister taxa (Thenius, 1959; Kurtén and Anderson, 1980). Our data are equivocal on this point as there is weak bootstrap support (38% ML, 36% MP) for a clade of black bears but we cannot reject a nonmonophyletic grouping of the two species. In fact parsimony analyses of our data are in agreement with Zhang and Ryder (1994) in suggesting that the Asiatic black bear is sister taxon to the sun/polar/brown bear clade (with a 58% bootstrap value; not shown).

Conversely, the sister taxa relationship between the

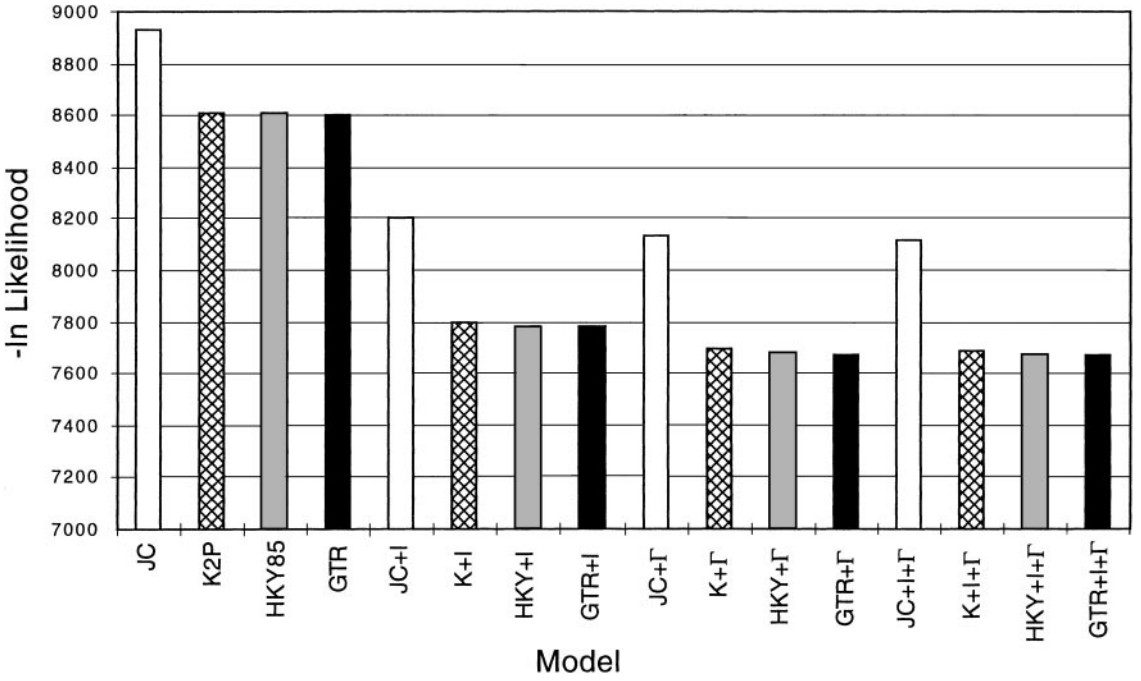

FIG. 2. Comparison and selection of models. Likelihood scores were calculated for each model using the tree that consistently had the highest or was tied for the highest likelihood scores across models. See text for model abbreviations. The HKY85 + I + Γ model provides the best compromise between goodness of fit (as indicated by likelihood score) and parameter economization.

black bear species is supported by a separate analysis of the entire sequence of the mtDNA Cyt b gene for all bear species (Talbot and Shields, 1996a). However, we analyzed their data using the same methodology as above (best fit model was GTR + I; $\chi^2_{[2]} = 3.2351$, $P > 0.05$), and their data set does not statistically reject the placement of the Asiatic black bear as sister to the polar/brown/sun bear clade ($P > 0.05$). Therefore, the alternative phylogenetic hypothesis regarding black bear species cannot be resolved using the molecular data currently available. The sloth bear appears to be the basal ursine in the combined data analyses as suggested by earlier mtDNA phylogenetic analyses (Zhang and Ryder 1993, 1994; Talbot and Shields, 1996a).

Based on the results from this study and earlier molecular analyses, we suggest that the difficulties surrounding the resolution of the four remaining Ursine bear species are linked to a rapid radiation and speciation event that occurred 2–3.5 mya based on fossil data (Kurtén and Anderson, 1980). During this time period, four ancestral bear lineages appear to have diverged and subsequently evolved into the extant taxa consisting of sun bear, Asiatic and American black bears, brown bear, and polar bear (Fig. 5). In the last 1 million years, the lineage leading to the brown and polar bears diverged into two morphologically and behaviorally distinct species. The resulting paraphyletic status of the brown bear in relation to the polar bear illustrates how rapidly morphological and behavioral changes can outstrip the divergence of molecular lineages into monophyletic clades and highlights the phenotypic plasticity of bear species.

Resolution of the Basal Lineages

Placement of the giant panda and spectacled bear, respectively, as the two basal bear lineages in the combined data set is consistent with nuclear DNA chromosome analyses (Wuster-Hill and Bush, 1980; Nash and O'Brien, 1987), protein allozyme data (Goldman *et al.,* 1987), and mtDNA sequence analyses (Zhang and Ryder, 1993, 1994; Talbot and Shields 1996a). Three resolutions of the basal lineages (rooting of the ingroup topology) are possible; the giant panda may be basal, the spectacled bear may be basal, or these two may form a monophyletic basal lineage. Talbot and Shields (1996a) found support for all three of these hypotheses. At first glance, our data seem to unequivocally resolve the giant panda as the basal lineage based on very strong bootstrap support in parsimony and moderately strong bootstrap support for likelihood (67% ML, 99% MP) (Fig. 4). However, neither of the suboptimal resolutions can be rejected based on the Kishino–Hasegawa test. The best tree with the spectacled bear basal had a likelihood score of −7678.83643 ($P = 0.24$), whereas the best tree with the spectacled bear and panda sister taxa had a likelihood score of −7678.91995 ($P = 0.22$).

Given the strong bootstrap support seen for the basal position of the giant panda (Fig. 4), it is surprising that

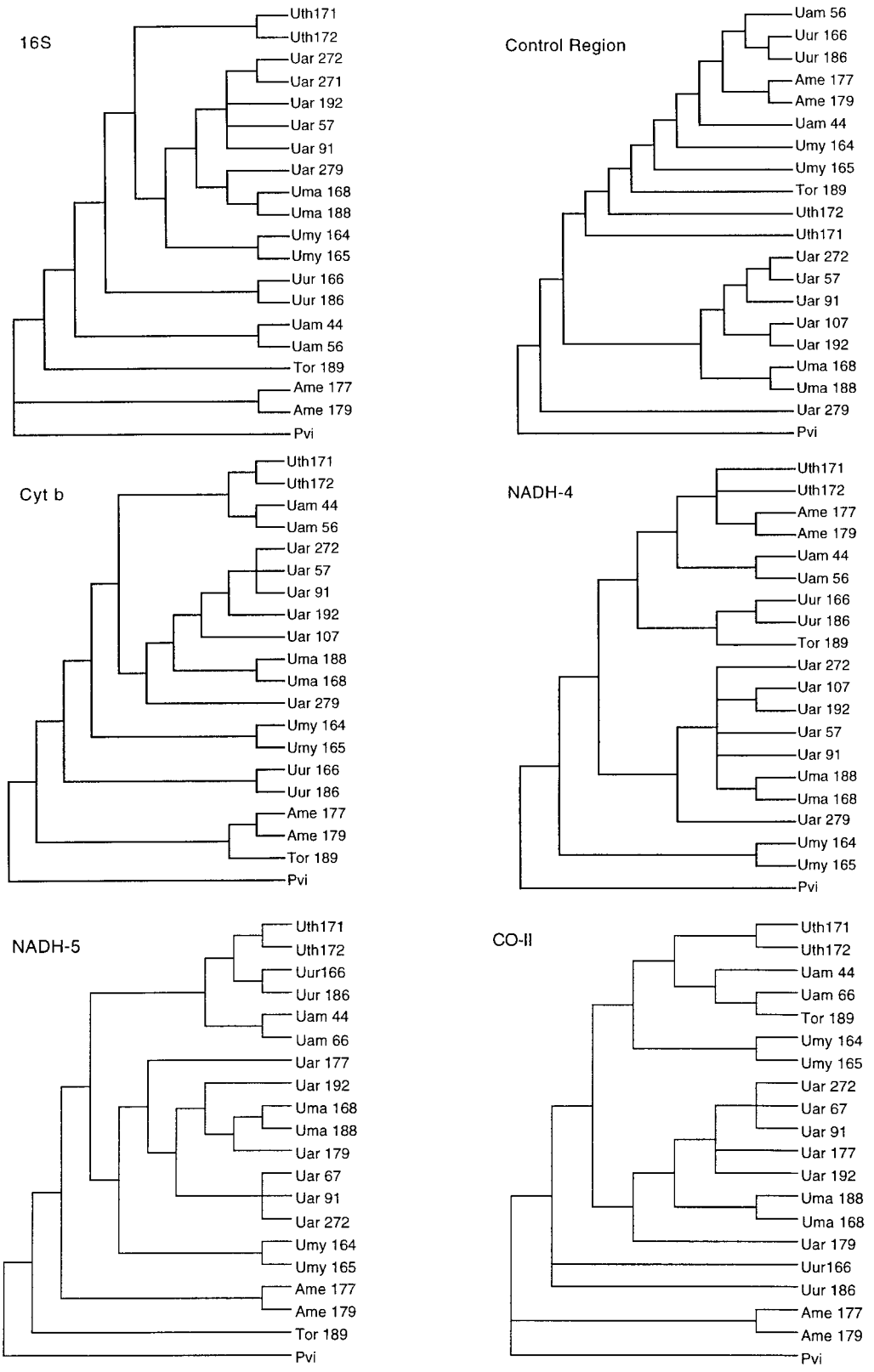

FIG. 3. Strict maximum likelihood consensus trees estimated for each single region using the HKY + I + Γ model (PAUP*). All six trees are drawn to the same scale with branch lengths defined by maximum likelihood estimates. Abbreviations: CO-II, cytochrome oxidase subunit II; Cyt b, cytochrome *b;* NADH-4, NADH dehydrogenase subunit 4; AA, amino acids; 16S, 16S ribosomal RNA; and NADH-5, NADH dehydrogenase subunit 5. Abbreviations: Uth, *Ursus thibetanus;* Uam, *U. americanus;* Uar, *U. arctos;* Uma, *U. maritimus;* Umy, *U. malayanus;* Uur, *U. ursinus;* Tor, *Tremarctos ornatus;* Ame, *Ailuropoda melanoleuca;* Pvi, *Phoca vitulina.*

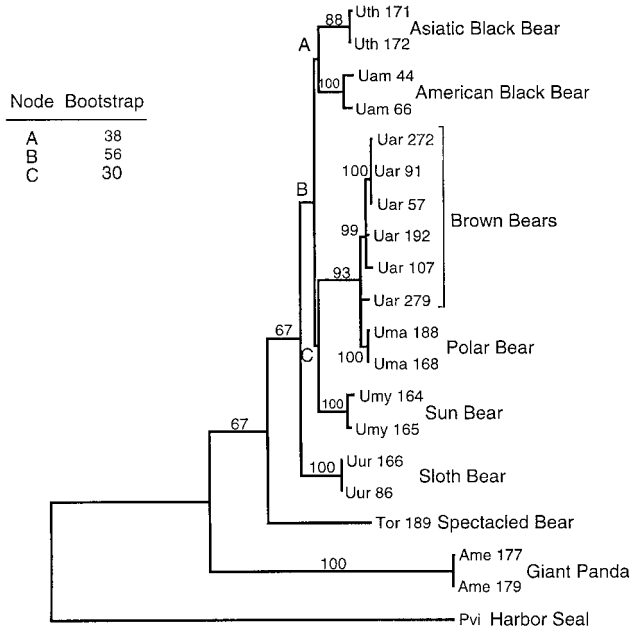

FIG. 4. Optimal maximum likelihood phylogeny (HKY + I + Γ) from the combined data analysis of 1916 nt of mtDNA sequence data for the Ursidae using the harbor seal as an outgroup. Bootstrap values from 100 maximum likelihood replicates are listed at each node. Abbreviations: Uth, *Ursus thibetanus;* Uam, *U. americanus;* Uar, *U. arctos;* Uma, *U. maritimus;* Umy, *U. malayanus;* Uur, *U. ursinus;* Tor, *Tremarctos ornatus;* Ame, *Ailuropoda melanoleuca;* Pvi, *Phoca vitulina.*

the KH tests could not distinguish among the alternative basal resolutions. The reason for this difficulty can be easily seen by examining the branch lengths on the best constrained trees. The two longest branches in the phylogeny are those leading to the outgroup and to the panda and long-branch attraction likely accounts for the extremely high parsimony bootstrap values. However, this does not necessarily imply that the basal position of the panda is artifactual. If two long-branch taxa (i.e., a basal taxon and an outgroup taxon) are on the same side of a short internal branch, methods that ignore process will have a higher likelihood of inferring the correct relationships because the length of the internal branch will be (perhaps strongly) overestimated (Wadell, 1995) and estimates of support for that node are expected to be inflated. This situation has been called the anti-Felsenstein Zone (Waddell, 1995). Thus, we conclude that the available molecular data suggest that a basal polytomy best represents the rapid radiation at the base of the ursid phylogeny (Fig. 5). It was not possible to address the problematic issue of the inclusion of the giant panda in the bear family because many of our primers would not work for taxa critical to the debate (red panda, raccoon).

Use of the Molecular Clock

Several studies have used the molecular clock hypothesis to attempt to date divergence within the ursids by using the 12 MY divergence date of the giant panda (Thenius, 1979; Wayne *et al.,* 1991) as a calibration. However, based on a likelihood-ratio test (Felsenstein, 1988), the molecular clock hypothesis can be rejected for our data ($d = 60.34$; $P < 0.001$), thus prohibiting us from using these data to estimate divergence times. Paradoxically, the giant panda lineage itself seems to be violating the clock most egregiously. This can be seen by using the linearized trees approach of Takezaki *et al.* (1995). We must prune both the giant panda and the spectacled bear from the tree for the data to fit the clock hypothesis ($\chi^2_{[14]} = 15.52118$; $0.5 > P > 0.1$). Talbot and Shields (1996a) reported that, as was the case with our data, the clock hypothesis could not be rejected for the ursines; however, they did not report a test of the clock

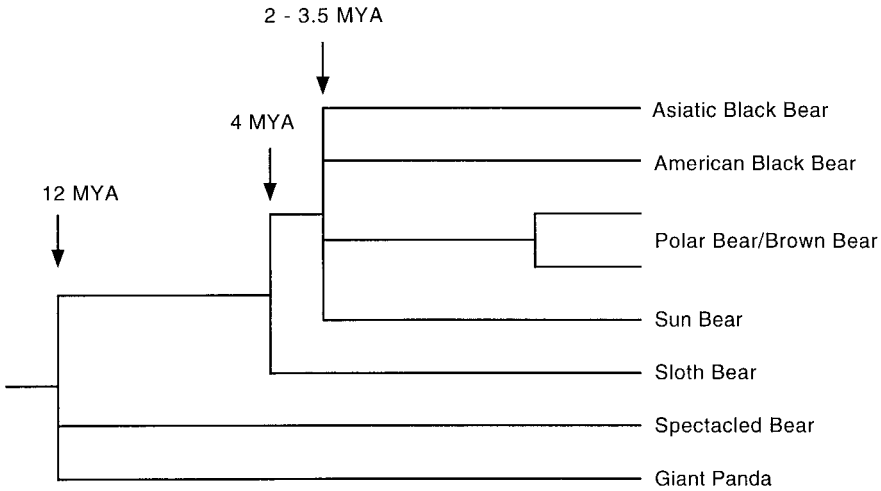

FIG. 5. Summary phylogeny of Ursidae divergences and radiations based on molecular data from this study, Talbot and Shield (1996), and Zhang and Ryder (1996). Divergence dates were obtained from fossil data (Thenius, 1979; Kurtén and Anderson, 1980; Wayne *et al.,* 1991).

for the entire data set. To assess the validity of the dates published by Talbot and Shields (1996a), we evaluated the fit of their data to the clock hypothesis in the same manner as above. As was the case with our data, the clock hypothesis can be rejected for the entire data set. Thus, the giant panda lineage violates the molecular clock assumption, and this lineage cannot be used as a reliable internal calibration to estimate divergence time for the remaining (clock-like) ursines as suggested by Talbot and Shields (1996a). Dating the ursine divergences using a molecular clock approach will require either molecular data for which the panda lineage does not violate the clock relative to the ursines or independent fossil data from an ursine divergence to calibrate the ursine clock.

Taxonomy and Evolution of Bears

If taxonomy is to reflect the hierarchy of evolutionary relationships, placement of the American black bear in *Ursus,* the same genus as the brown bear, seems unwarranted unless the sun bear and Asiatic black bear are also placed in this genus, as suggested by Nowak (1991). The genetic and morphological differentiation of the sloth bear supports separate placement in the genus *Melursus* (Ewer, 1973; Eisenberg, 1981; Honacki *et al.,* 1982; Corbet and Hill, 1991). The close genetic relationship of the brown bear and the polar bear reinforces recommendations that the polar bear genus designation *Thalarctos* be abandoned in favor of *Ursus* (Honacki *et al.,* 1982; Nowak, 1991). The phylogenetic distinctiveness and basal placement of the spectacled bear and giant panda support retention of the genus names suggested in Table 1.

Delineation of the evolutionary relationships of the extant bears provides a perspective on the processes that underlie bear evolution. Although *Ursavus,* the ancestral taxon, was strictly carnivorous, our phylogeny suggests that the extant bears arose by rapid radiation events that transformed a generalized carnivore into a series of ecomorphs (Martin, 1989). While each taxon follows the basic carnivore morphological pattern, the extant species are characterized by substantial reliance on herbivory and insectivory, despite long periods of evolutionary separation. The lineages that have diverged longest (giant panda, spectacled bear, and sloth bear) are the least carnivorous, while the youngest taxa (brown bear and polar bear) are the most carnivorous. By virtue of its colonization of the Arctic habitat, the polar bear is the sole exception to the general ursid deviation from true carnivory. Moreover, the selective pressure of adapting to the new environment rapidly led to the morphological features that distinguish the polar bear from brown bears, well before the molecular lineages could separate into distinct clades.

ACKNOWLEDGMENTS

We are grateful to the following who generously gave us samples for this study: S. Fain—Asiatic black bear and spectacled bear; S. Pääbo and M. Kohn—European brown bear; colleagues in US Fish and Wildlife Service, particularly C. Servheen, R. Smith, T. Thier, and V. Barnes—American black bears, grizzly, and Kodiak bears; D. Knight of the Interagency Grizzly Bear Study Team—grizzly bears; E. Ludwig—harbor seal. We benefited from their comments as well as from S. George's helpful critique. We thank David L. Swofford for allowing us to publish analyses conducted using test versions of PAUP*. The inception and completion of this study benefited immeasurably from the help and encouragement of M. and S. French of the Yellowstone grizzly foundation. Partial support came from a National Science Foundation doctoral fellowship (L.P.W.).

REFERENCES

Árnason, Ú., and Johnsson, E. (1992). The complete mitochondrial DNA sequence of the harbor seal, *Phoca vitulina. J. Mol. Evol.* **34:** 493–505.

Avise, J. C., Shapira, J. F., Daniel, S. W., Aquadro, C. F., and Lansman, R. A. (1983). Mitochondrial DNA differentiation during the speciation process in *Peromyscus. Mol. Biol. Evol.* **1:** 38–56.

Bull, J. J., Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L., and Waddell, P. J. (1994). Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* **43:** 368–372.

Corbet, G. B., and Hill, J. E. (1991). "A World List of Mammalian Species," Oxford Univ. Press, New York.

Cronin, M. A., Amstrup, S. C., Garner, G. W., and Vyse, E. R. (1991). Interspecific and intraspecific mitochondrial DNA variation in North American bears (Ursus). *Can. J. Zool.* **69:** 2985–2992.

Crusafont, M., and Kurtén, B. (1976). Bears and bear-dogs from the Vallesian of the Valles-Penedes basin, Spain. *Acta Zool. Fenn.* **144:** 1–29.

Cunningham, C. W., Blackstone, N. W., and Buss, L. W. (1992). Evolution of king crabs from hermit crab ancestors. *Nature* **355:** 539–542.

Desalle, R., and Brower, A. V. Z. (1997). Process partitions, congruence, and the independence of characters: Inferring relationships among closely related Hawaiian *Drosophila* from multiple gene regions. *Syst. Biol.* **46:** 751–764.

Eisenberg, J. F. (1981). "The Mammalian Radiations," Univ. of Chicago Press, Chicago.

Ewer, R. F. (1973). "The Carnivora," Cornell Univ. Press, Ithaca.

Frati, F., Simon, C., Swofford, D. L. (1997). Evolution of the mitochondrial cytochrome oxidase II gene in collembola. *J. Mol. Evol.* **44:** 145–156.

Felsenstein, J. (1985). Confidence limits on phylogeny: An approach using the bootstrap. *Evolution* **39:** 783–791.

Felsenstein, J. (1988). Phylogenies from molecular sequences: Inference and reliability. *Annu. Rev. Genet.* **22:** 521–565.

Gaut, B. S., and Lewis, P. O. (1995). Success of maximum likelihood phylogeny inference in the four-taxon case. *Mol. Biol. Evol.* **12:** 152–162.

Goldman, D., Giri, P. R., and O'Brien, S. J. (1987). Molecular genetic distance estimates among the Ursidae as indicated by one- and two-dimensional protein electrophoresis. *Evolution* **43:** 282–295.

Gray, A. P. (1971). "Mammalian Hybrids," Commonwealth Agricultural Bureaux, Slough.

Gu, X., Fu, Y.-X., and Li, W.-H. (1995). Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Mol. Biol. Evol.* **12:** 546–557.

Gutell, R. R., and Fox, G. E. (1988). A compilation of large subunit RNA sequences presented in a structural format. *Nucleic Acids Res.* **16:** r175–r313.

Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating the human-ape split by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22:** 160–174.

Hashimoto, T., Otaka, E., Adachi, J., Mizuta, K., and Hasegawa, M. (1993). The giant panda is closer to a bear, judged by alpha- and beta-hemoglobin sequences. *J. Mol. Evol.* **36:** 282–289.

Higgins, D. H., and Sharp, P. M. (1989). Fast and sensitive multiple region sequence alignments on a microcomputer. *CABIOS* **5:** 151–153.

Honacki, J. H., Kinman, K. E., and Koeppl, J. W. (1982). "Mammal Species of the World," Association of Systematic Collections, Lawrence.

Hultman, T., Stahl, S., Hormes, E., and Uhlén, M. (1989). Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. *Nucleic Acids Res.* **17:** 4937–4946.

Jukes, T. H., and Cantor, C. R. (1969). Evolution of protein molecules. *In* "Mammalian Protein Metabolism" (H. N. Munro, Ed.), pp. 21–132. Academic Press, New York.

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16:** 111–120.

Kishino, H., and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidae. *J. Mol. Evol.* **29:** 170–179.

Kitchener, A. C. (1994). A review of the evolution, systematics, functional morphology, distribution and status of the Ursidae. *Int. Zoo News* **242:** 4–24.

Kocher, T., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., and Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* **86:** 6196–6200.

Kurtén, B. (1964). The evolution of the polar bear, *Ursus maritimus. Acta Zool. Fenn* **108:** 1–30.

Kurtén, B. (1968). "Pleistocene Mammals of Europe," Aldine, Chicago.

Kurtén, B., and Anderson, E. (1980). "Pleistocene Mammals of North America," Cornell Univ. Press, Ithaca.

Lockhart, P. J., Steel, M. A., Hendy, M. D., and Penny, D. (1994). Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* **11:** 650–612.

Martin, L. D. (1989). Fossil history of the terrestrial Carnivora. *In* "Carnivore Behavior, Ecology, and Evolution" (J. L. Gittleman, Ed.), pp. 536–568. Cornell Univ. Press, Ithaca.

Mason-Gamer, R. J., and Kellogg, E. A. (1996). Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). *Syst. Biol.* **45:** 524–546.

Nash, W. G., and O'Brien, S. J. (1987). A comparative chromosome banding analysis of the Ursidae and their relationship to the carnivores. *Cytol. Cell Genet.* **45:** 206–212.

Nowak, R. M. (1991). "Walker's Mammals of the World," Johns Hopkins Press, Baltimore.

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

Saccone, C., Pesole, G., and Sbisa, E. (1991). The main regulatory region of mammalian mitochondrial DNA: Structure–function model and evolutionary pattern. *J. Mol. Evol.* **33:** 83–91.

Saiki, R. K. (1990). Amplification of genomic DNA. *In* "PCR Protocols" (M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, Eds.), pp. 13–20. Academic Press, San Diego.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Plainview.

Shields, G. F., and Kocher, T. D. (1991). Phylogenetic relationships of North American Ursids based on analysis of mitochondrial DNA. *Evolution* **45:** 218–221.

Sullivan, J. (1996). Combining data with different distributions of among-site rate variation. *Syst. Biol.* **45:** 375–380.

Sullivan, J., and Swofford, D. L. (1997). Are guinea pigs rodents? The importance of adequate models in molecular phylogenetics. *J. Mammal. Evol.* **4:** 77–86.

Sullivan, J., Markert, J. A., and Kilpatrick, C. W. (1997). Phylogeography and molecular systematics of the *Peromyscus aztecus* species group (Rodentia; Muridae) inferred using parsimony and likelihood. *Syst. Biol.* **46:** 426–440.

Swofford, D. L., Olsen, G. J., Waddell, P. J., and Hillis, D. M. (1996). Phylogenetic inference. *In* "Molecular Systematics" (D. M. Hillis, C. Moritz, and B. K. Mable, Eds.), 2nd ed., pp. 407–514. Sinauer, Sunderland, MA.

Tagle, D. A., Miyamoto, M. M., Goodman, M., Hofmann, O., Braunitzer, G., Goeltenboth, G. R., and Jalanke, H. (1988). Hemoglobins of the pandas: Phylogenetic relationships of carnivores as ascertained by protein sequence data. *Naturwissenschaften* **73:** 512–514.

Takezaki, N., Rhzetsky, A., and Nei, M. (1995). Phylogenetic test of the molecular clock and linearized trees. *Mol. Biol. Evol.* **12:** 823–833.

Talbot, S. L., and Shields, G. F. (1996a). A phylogeny of the bears (Ursidae) inferred from complete sequences of three mitochondrial DNA genes. *Mol. Phylogenet. Evol.* **5:** 567–575.

Talbot, S. L., and Shields, G. F. (1996b). Phylogeography of brown bears (*Ursus arctos*) of Alaska and paraphyly within the Ursidae. *Mol. Phylogenet. Evol.* **5:** 477–494.

Thenius, E. (1959). Ursiden phylogenes und biostratigraphie. *Z. Saugetierk.* **24:** 78–84.

Thenius, E. (1979). Zur systematischen und phylogenetischen Stellung des Bambusbaren Ailuropoda melanoleuca David (Carnivora: Mammalia). *Z. Saugetierk.* **44:** 286–305.

Thenius, E. (1982). Zur stammes und verbreitungageachichtlichen herkunfft des bambusbären. *Z. Geol. Wiss. Berlin* **10:** 1029–1042.

Waddell, P. (1995). "Statistical Methods of Phylogenetic Analysis, Including Hadamard Conjugations, LogDet Transforms, and Maximum Likelihood," Ph.D. dissertation, Massey Univ., Palmerston North, New Zealand.

Waddell, P., and Penny, D. (1996). Evolutionary trees of apes and humans from DNA sequences. *In* "Handbook of Symbolic Evolution" (A. J. Lock and C. R. Peters, Eds.), pp. 53–73. Clarendon, Oxford.

Waits, L. P., Talbot, S., Ward, R. H., and Shields, G. F. (1998). Mitochondrial DNA phylogeography of the North American brown bear and implications for conservation. *Cons. Biol.* **12:** 408–417.

Vardenplas, S., Wild, I., Grober-Rabie, A., Brebrer, K., Ricketts, M., Wallis, A., Bester, A., Boyd, C., and Mathew, C. (1984). Blot hybridization analysis of genomic DNA. *J. Med. Genet.* **21:** 164–168.

Ward, R. H., Frazier, B. L., Dew-Jager, K., and Pääbo, S. (1991). Extensive mitochondrial diversity within a single Amerindian tribe. *Proc. Natl. Acad. Sci. USA* **88:** 8720–8724.

Wayne, R. K., Benveniste, R. E., Janczewski, D. N., and O'Brien, S. J. (1989). Molecular and biochemical evolution of the Carnivora. *In* "Carnivore Behavior, Ecology, and Evolution" (J. L. Gittleman, Ed.), pp. 465–494. Cornell Univ. Press, Ithaca.

Wayne, R. K., VanValkenburgh, B., and O'Brien, S. J. (1991). Molecular distance and divergence time in carnivores and primates. *Mol. Biol. Evol.* **8:** 297–319.

Wuster-Hill, D. H., and Bush, M. (1980). The interrelationship of

chromosome banding patterns in the giant panda (*Aliuropoda melanoleuca*), hybrid bear (*Ursus middendorfi x Thalarctos maritimus*), and other carnivores. *Cytogenet. Cell Genet.* **27:** 147–154.

Yang, Z. (1994a). Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* **39:** 105–111.

Yang, Z. (1994b). Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J. Mol. Evol.* **39:** 306–314.

Zhang, Y., and Ryder, O. A. (1993). Mitochondrial DNA sequence evolution in the Arctoidae. *Proc. Natl. Acad. Sci. USA* **90:** 9557–9561.

Zhang, Y., and Ryder, O. A. (1994). Phylogenetic relationships of bears (the Ursidae) inferred from mitochondrial DNA sequences. *Mol. Phylogenet. Evol.* **3:** 351–359.

Zhang, Y., and Shi, L. M. (1991). Riddle of the giant panda. *Nature* **352:** 573.