Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 37 (2005) 815–831
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

# A molecular phylogeny of the Canidae based on six nuclear loci

Carolyne Bardeleben *, Rachael L. Moore, Robert K. Wayne

*Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095-1606, USA*

Received 9 February 2005; revised 8 July 2005; accepted 28 July 2005
Available online 5 October 2005

**Abstract**

We have reconstructed the phylogenetic relationships of 23 species in the dog family, Canidae, using DNA sequence data from six nuclear loci. Individual gene trees were generated with maximum parsimony (MP) and maximum likelihood (ML) analysis. In general, these individual gene trees were not well resolved, but several identical groupings were supported by more than one locus. Phylogenetic analysis with a data set combining the six nuclear loci using MP, ML, and Bayesian approaches produced a more resolved tree that agreed with previously published mitochondrial trees in finding three well-defined clades, including the red fox-like canids, the South American foxes, and the wolf-like canids. In addition, the nuclear data set provides novel indel support for several previously inferred clades. Differences between trees derived from the nuclear data and those from the mitochondrial data include the grouping of the bush dog and maned wolf into a clade with the South American foxes, the grouping of the side-striped jackal (*Canis adustus*) and black-backed jackal (*Canis mesomelas*) and the grouping of the bat-eared fox (*Otocyon megalotis*) with the raccoon dog (*Nycteruetes procyonoides*). We also analyzed the combined nuclear + mitochondrial tree. Many nodes that were strongly supported in the nuclear tree or the mitochondrial tree remained strongly supported in the nuclear + mitochondrial tree. Relationships within the clades containing the red fox-like canids and South American canids are well resolved, whereas the relationships among the wolf-like canids remain largely undetermined. The lack of resolution within the wolf-like canids may be due to their recent divergence and insufficient time for the accumulation of phylogenetically informative signal.


*Keywords:* Canidae; Phylogeny; Systematics; Nuclear

## 1. Introduction

The Canidae is a family of dog-like carnivores that includes 16 genera and 36 species (Nowak, 1999). Considerable interest in the evolutionary relationships of canids has resulted in analyses based on morphological data (Berta, 1987; Lyras and Van Der Geer, 2003; Tedford et al., 1995; Zrzavy and Ricankova, 2004) and molecular data, including G-banded karyotypes (Wayne et al., 1987a,b), DNA–DNA hybridization (Wayne et al., 1990), allozymes (Wayne and O'Brien, 1987), and mitochondrial DNA sequences (Geffen et al., 1992; Wayne et al., 1997). The most definitive molecular phylogenetic analysis of the Canidae to date utilized 2001 bp of sequence data from three mitochondrial genes (Wayne et al., 1997) and found four monophyletic groups within the Canidae (Table 1): (1) the wolf- and jackal-like canids; (2) the red fox-like canids; (3) the South American foxes; and (4) the maned wolf (*Chysocyon brachyurus*) and bush dog (*Speothos venaticus*). In addition, it was concluded that the gray fox (*Urocyon cinereoargenteus*), raccoon dog (*Nycteruetes procyonoides*), and bat-eared fox (*Otocyon megalotis*) are basal canids not closely associated with any of these monophyletic groups (Wayne et al., 1997). However, several phylogenetic issues remain unresolved including: (1) the branching order among the gray fox, raccoon dog and bat-eared fox; (2) the monophyly of the crab-eating fox (*Cerdocyon thous*) and small-eared dog (*Atelocynus mic-*

* Corresponding author. Fax: +1 310 206 3987.
*E-mail address:* carolyne@ucla.edu (C. Bardeleben).

1055-7903/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2005.07.019

Table 1
Taxa used in this study: species, abbreviation, and common name

| Species | Abbreviation | Common name |
|---|---|---|
| Red fox-like canids | | |
| *Alopex lagopus* | Ala | Arctic fox |
| *Fennecus zerda* | Fze | Fennec fox |
| *Vulpes corsac* | Vco | Corsac fox |
| *Vulpes macrotis* | Vma | Kit fox |
| *Vulpes vulpes* | Vvu | Red fox |
| Wolf-like canids | | |
| *Canis adustus* | Cad | Side-striped jackal |
| *Canis aureus* | Cau | Golden jackal |
| *Canis latrans* | Cla | Coyote |
| *Canis lupus* | Clu | Gray wolf |
| *Canis familiaris* | Cfa | Dog |
| *Canis mesomelas* | Cme | Black-backed jackal |
| *Cuon alpinus* | Cal | Dhole |
| *Lycaon pictus* | Lpi | Wild dog |
| South American canids | | |
| *Chrysocyon brachyurus* | Cbr | Maned wolf |
| *Speothos venaticus* | Sve | Bush dog |
| South American foxes | | |
| *Atelocynus microtis* | Ami | Small-eared dog |
| *Cerdocyon thous* | Cth | Crab-eating fox |
| *Pseudalopex griseus* | Pgr | Argentine gray fox |
| *Pseudalopex gymnocercus* | Pgy | Pampas fox |
| *Pseudalopex sechurae* | Pse | Sechuran fox |
| Putative basal canids | | |
| *Nycteruetes procyonoides* | Npr | Raccoon dog |
| *Otocyon megalotis* | Ome | Bat-eared fox |
| *Urocyon cinereoargenteus* | Uci | Gray fox |

*rotis*); (3) the relationship of the wild dog (*Lycaon pictus*) and side-striped jackal (*Canis adustus*) to the wolf-like canids; (4) the monophyly of the maned wolf and bush dog; and (5) the iterative or single appearance of the trenchant heel (a specialized dentition feature for cutting meat).

To better resolve these issues and to test previous hypotheses based on morphologic and molecular data, we sequenced portions of six nuclear genes focusing on non-coding regions that were located on different chromosomes. The Canidae consists of a group of species with a relatively recent evolutionary history and estimated divergence times between 0.3 and 12 MYA (Bininda-Emonds et al., 1999; Wayne et al., 1991). Although mitochondrial genes have the advantage of evolving rapidly, they essentially provide a single gene genealogy. Nuclear markers are attractive because multiple unlinked regions of the genome can be sampled and their slower rate of substitution reduces levels of homoplasy in the sequence data (Matthee et al., 2001; Prychitko and Moore, 2000; Springer et al., 1999, 2001). In addition, non-coding regions are more likely to accumulate indels which can provide phylogenetically useful information because they are relatively rare events and less likely to be reversible (Rokas and Holland, 2000). If multiple gene trees yield the same topology, it strengthens the support that the gene tree reflects the species tree (Miyamoto and Fitch, 1995; Penny and Hendy, 1986). If not, it suggests that additional data or analyses are required (Rokas et al., 2003). Further, combining data from multiple loci may reveal relationships not seen in analyzing individual loci and strengthen seemingly weak relationships (Chippindale and Wiens, 1994; Koepfli and Wayne, 2003; Mitchell et al., 2000; Olmstead and Swere, 1994; Slade et al., 1994). Finally, advances in methods of phylogenetic analyses have occurred since the previous mitochondrial DNA sequence analysis was performed, which allow us to perform additional analyses with maximum likelihood (ML) and Bayesian methods and compare tree topologies derived from the nuclear or mitochondrial data sets in a statistical framework. We analyzed the nuclear derived tree separately to obtain an estimate of canid phylogeny independent of the previously published mitochondrial phylogeny. We also combined the nuclear and mitochondrial data set to determine the resulting phylogeny.

## 2. Materials and methods

### 2.1. Samples, DNA isolation, PCR amplification, and sequencing

We obtained tissue or DNA samples from 23 species of Canidae, representing 14 genera (Table 1) from sources as previously described (Bardeleben et al., 2005). Samples from *Canis simensis, Lycalopex vetulus*, and *Pseudalopex culpaeus* included in Wayne et al. (1997) were no longer available. We refer to a species by its genus name only if the genus is monotypic. When we refer to a genus with multiple taxa, we refer to all the taxa in that genus, (e.g., *Pseudalopex* equals *Pseudalopex griseus*, *Pseudalopex gymnocercus*, and *Pseudalopex sechurae*). Seven arctoid carnivores including *Ursus americanus*, *Ailuropoda melanoleuca* (Institute of Zoology, London), *Odobenus rosmarus* (University of Alaska), *Mirounga augustirostris* (Brent Steward, Sea World), *Enhydra lutris*, *Lontra longicaudis*, and *Procyon lotor*, were used as outgroups. Sources for *E. lutris*, *L. longicaudis*, and *P. lotor* are as previously published (Koepfli and Wayne, 2003). For tissue or blood samples, total genomic DNA was isolated using standard phenol/chloroform extraction followed by precipitation with ethanol (Sambrook et al., 1989). Six regions of nuclear DNA consisting mainly of non-coding DNA were amplified by the polymerase chain reaction (PCR). The name, primers and chromosomal location for each locus are listed in Table 2. Some primers were modified to obtain an amplification product in the outgroups, including those used to amplify *CYPIA* in *A. melanoleuca, E. lutris*, and *L. longicaudis*, and the reverse primer used to amplify the 3′flank in *TRSP* in *E. lutris, L. longi-*

Table 2
Gene symbol, name, location, primers, and region amplified of nuclear loci used in this study

| Locus | Gene name | Location[a] | Primers | Region amplified[b] | Reference |
|---|---|---|---|---|---|
| *CHRNA1* | Cholinergic receptor, nicotinic alpha polypeptide 1 precursor | 2q24–q32 | F: 5′gaccatgaagtcagaccaggag3′<br>R: 5′ggagtatgtggtccatcaccat3′ | Intron 8 | Lyons et al. (1997) |
| *CYP1A1* | Cytochrome P-450 | 7 | (a) F: 5′ttggacctctttggagctgg3′<br>R: 5′tggttgatctgccactggtt3′ | Intron 3 to exon 6 | Venta et al. (1996) |
| | | | (b) F[c]: 5′gatttgacacagtcacaact3′<br>R[c]: 5′aagacgcaacgtcccttg3′ | | This study |
| *FES* | Feline sarcoma protooncogene | 15 | F: 5′ggggaactttggcgaagtgtt3′<br>R: 5′tccatgacgatgtagatggg3′ | Intron 14 | Venta et al. (1996) |
| *GHR* | Growth hormone receptor | 5 | F: 5′ccagttccagttccaaagat3′<br>R: 5′tgattcttctggtcaaggca3′ | Intron 9 to exon 10 | Venta et al. (1996) |
| *VTN* | Vitronectin | 17 | F: 5′agtgaggcctgggtaccc3′<br>R: 5′gaagaagtagacccgctccc3′ | Intron 4 | Jiang et al. (1998) |
| *TRSP* | Selenocyteine tRNA gene | 19q3 | (a) 5′ flank F: 5′gggcttctgaaagccgactt3′<br>R: 5′ccgcccgaaaggtggaattg3′ | 5′/3′ gene-flanking DNA, respectively | Bardeleben et al. (2005) |
| | | | (b) 3′flank F: 5′gcccggatgatcctcagtgg3′<br>R: 5′cactgtgtgccagcacctggc3′<br>R[d]: 5′gtgaaggggagagatcaaggacg3′ | | This study |

[a] Human chromosome.
[b] Generally primers are located in the exons flanking the introns; bp locations available on request.
[c] Used to amplify the outgroup taxa *Ailuropoda melanoleuca* and *Enhydra lutris*; to amplify *CYPIA1* in *L. longicaudis* the forward primer of (a) and the reverse primer of (b) was used.
[d] Used to amplify the 3′ flank of *TRSP* in the outgroup taxa *Enhydra lutris*, *Lontra longicaudis*, and *Procyon lotor*.

*caudis*, and *P. lotor* (Table 2). The conditions for PCR for *CHRNA1, CYP1A1, FES, GHR*, and *VTN* loci were 10 mM Tris–Cl, pH 9, 2.5 mM $MgCl_2$, 50 mM KCl, 200 μM each dNTP, 0.5 μM each primer, and 0.2 U *Taq* (Sigma) in a final reaction volume of 50 μL. The PCR program was 94 °C for 3 min, 30 cycles of 94 °C for 45 s, 54 °C for 45 s, 72 °C for 45 s, and a final extension of 72 °C for 5 min. Conditions for PCR for the *TRSP* locus have been previously described (Bardeleben et al., 2005). PCR products were fractionated on 1% agarose/TAE gels, bands of the expected size were excised, purified by UltraClean (BioOne), and sequenced using the originating PCR primers and the ABI PRISM BigDye Terminator Cycle Sequencing kit (PE Applied Biosystems). In some cases, it was necessary to clone the PCR product into the TOPO4 vector (Invitrogen) to obtain unambiguous sequence and at least five clones were sequenced. When possible, loci from at least two individuals for each species were sequenced. Species represented by only one individual include *Atelocynus*, *Fennecus*, and *P. sechurae*. Multiple attempts failed to amplify the *CYPIA* locus in *P. lotor* and consequently, this sequence is coded as missing in the data set. In addition, multiple attempts failed at obtaining unambiguous sequence for the first 190 nt of the *CYPIA* locus in *Atelocynus* and this region is coded as missing as well.

### 2.2. Phylogenetic analyses

All nuclear sequences except those for the *TRSP* loci were generated for this study. The *TRSP* sequences, which include the selenocysteine tRNA gene and its flanking region, are from a previous study (Bardeleben et al., 2005). Mitochondrial DNA sequences for cytochrome *b* (*cyt b*), cytochrome oxidase I (*COI*), and cytochrome oxidase II (*COII*) are from Wayne et al. (1997) except for the maned wolf *COII* sequence (Bardeleben et al., 2005). The *COI* sequences for *Alopex* and *Vulpes corsac*, and the *COI* and *COII* sequences for the outgroup taxa, *A. melanoleuca, E. lutris*, *L. longicaudis*, and *P. lotor*, were generated for this study using previously published primers (Wayne et al., 1997). All accession numbers are listed in Supplementary Table 1. If a heterozygous site was supported by sequencing on both strands and/or multiple sequencing runs, it was scored as a polymorphism (e.g., A and G = R, T and A = W, etc.). Sequences were aligned either visually or initially using Clustal X ver. 1.81 (Jeanmougin et al., 1998) followed by visual adjustment to minimize indels. In all but one analysis, each species was represented by one sequence; the other analysis explored the effect of multiple individuals or multiple alleles (from heterozygotes) on tree topology. This multiple species data set consisted of 50 sequences. Species were represented by two individuals with the following exceptions: *Atelocynus*, *P. sechurae*, and *Fennecus* were represented by one individual; and *Urocyon* and *P. griseus* were represented by three individuals. Heterozygous individuals were found in *Canis aureus*, *C. adustus*, *Canis mesomelas*, and *Otocyon*.

As the data matrix consists of closely related species with relatively few variable characters, we focused on character-based rather than distance based methods to maximize the informativeness of the data. Phylogenetic trees were generated by maximum parsimony (MP) and

maximum likelihood (ML) using PAUP*4.0b10 (Swofford, 2003) and by Bayesian inference using either Mr. Bayes 2.01 or Mr. Bayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). The data were analyzed in the following partitions: (1) each nuclear gene segment alone; (2) all nuclear gene segments combined; (3) three mitochondrial genes (*COI, COII*, and *cyt b*) combined; and (4) nuclear and mitochondrial sequences combined. For parsimony analyses of the nuclear data, all characters were weighted equally (unweighted). A data set in which indels were coded for phylogenetic information was used (Barriel, 1994). MP analysis of the mitochondrial DNA data set was done either with characters unweighted (UW-MP) or characters weighted based on the transition/transversion ratio of 8 estimated by averaging the ratios of all pairwise comparisons among the ingroup taxa or a value of 11 based on a ML estimation performed on an unweighted MP tree. Heuristic searches with 100 replicates of random stepwise addition and tree bisection–reconnection branch swapping was used. For the combined nuclear + mitochondrial data set, MP with implied weighting (IW-MP) ($K = 2$) was performed to down-weight homoplastic characters (Goloboff, 1993). Nodal support was evaluated by non-parametric bootstrapping (BS) using 1000 pseudoreplicates (Felsenstein, 1985). Branch support (Bremer, 1988, 1994) and partitioned branch support (PBS) (Baker and Desalle, 1997) was calculated using TreeRot (Sorenson, 1999) using the data set with gaps coded as missing. Assessing clade significance using branch support values is difficult because a particular value for branch support is data-dependent (Lee, 2000). Therefore, branch support is mainly used to assess support and conflict at nodes. Hidden support and conflict was calculated as described by (Gatesy et al., 1999). Briefly, hidden branch support (or hidden conflict) for a particular combined data set and a particular node is the difference between the branch support for that node in a simultaneous analysis and the sum of branch support values for that node from each data partition (locus) (Gatesy et al., 1999). To determine whether a data set contained more hidden support or hidden conflict, the sum of the branch support over all the nodes from a simultaneous analysis was compared to the sum of branch support over all nodes from each data partition.

ML and Bayesian analyses were carried out with a data set in which indels were coded as missing. ML analyses were carried out with the model of nucleotide substitution and parameters determined by ModelTest v.3.5 to best fit the data set using the AIC criterion (Posada and Crandall, 1998). For ML analyses, heuristic searches with 100 replicates of random stepwise addition and tree bisection–reconnection branch swapping was used. Nodal support in the ML trees for each single locus was evaluated using 100 psuedoreplicates with 10 replicates of random stepwise addition of non-parametric bootstrapping. Nodal support in the nuclear tree, mitochondrial tree and combined nuclear + mitochondrial tree was evaluated using a reduced effort bootstrap (MulTree option off) with five replicates of random stepwise addition (Debry and Olmstead, 2000). ML trees with topological constraints were tested for statistical significant differences using the Shimodaira-Hasegawa (SH) test with RELL bootstrapping (1000 replicates) implemented through PAUP* (Shimodaira and Hasegawa, 1999).

Bayesian analysis was used on the nuclear, mitochondrial, and combined nuclear + mitochondrial data sets. Metropolis-coupled Markov chain Monte Carlo (MCMCMC) phylogenetic reconstructions were conducted in MrBayes v3.1 (Ronquist and Huelsenbeck, 2003) with vague priors (as per the program's default) and model parameters estimated as part of the analyses. Three heated chains and a single cold chain were used in all MCMC analyses and runs were initiated with random trees, as per the program's default. Exploratory analyses were run for 500,000 (mitochondrial data set) or 1,000,000 generations. Two longer runs of 5,000,000 generations were carried out for the final analysis. Trees were sampled at intervals of 100 generations. Stationarity was achieved when ln likelihood values approached equilibrium, as determined by plotting the ln likelihood scores of the sampled trees against generation time. All trees sampled before reaching stationarity were discarded as "burn in" (Huelsenbeck et al., 2002). The remaining trees were used to generate a 50% majority-rule consensus tree with the percentage of trees recovering the node representing the node's posterior probability. Typically, the initial 25% of the sampled trees were discarded as burn in. We also viewed the tree created after discarding 50% of the sampled trees. Each analysis was run at least twice to compare for convergence determined by similar ln likelihood values in each run. Initial Bayesian analyses used the nucleotide substitution model determined by ModelTest. Exploratory runs were carried out to determine whether partitioning the data set and using the site-specific gamma (SSG) command would improve likelihood scores (Castoe et al., 2004). For the combined nuclear data set, the Bayesian analysis was carried out either without partitioning the data set or partitioning the data set into each locus and specifying the site-specific gamma command. Partitioning the data set improved likelihood scores ($\ln L$: $-13272$ vs. $-13304$, partitioned by locus and no partitioning, respectively). For the mitochondrial data set, the data was either analyzed using the General Time Reversible model of substitution with some invariable sites and with variable sites assumed to follow a gamma distribution (GTR+I+G) or partitioned according to codon position, and analyzed using the GTR model and a gamma parameter that is estimated for each of the three

positions (GTR+SSG3) (ln *L*: −17828 vs. −17265 for non-partitioned vs. three partitions) (Castoe et al., 2004; Danforth et al., 2003; Debry, 1999; Hall, 2001; Rodriguez et al., 1990). Further partitioning of the mitochondrial data set into each of the three genes did not greatly improve likelihood scores. For the combined nuclear + mitochondrial data set, the data were either analyzed using GTR+I+G or organized into four partitions with the nuclear sequences as one partition and the mitochondrial genes partitioned by codon position and analyzed using GTR and a gamma parameter that is estimated for each of the four partitions (GTR+SSG4) (ln *L*: −32,512 vs. −31,256 for non- partitioned vs. four partitions). The AIC (an information criterion) indicated that the model with the site –specific gamma produced a better fit to the data than the model without it (Akaike, 1974; Dumbacher et al., 2003).

Nodes with bootstrap (BS) values of over 70% in the MP and ML analyses and posterior probabilities (PP) of over 95% in the Bayesian analysis are considered to have strong support (Hillis and Bull, 1993; Huelsenbeck and Ronquist, 2001).

## 3. Results

### 3.1. Characteristics of nuclear loci

The sequence obtained from each locus ranged in length from 390 to 874 bp, inclusive of indels (Table 3). Among the individual loci, the percentage of non-coding DNA ranged from 49.9% (*VTN*) to 89.8% (*CHRNA1*). Within the ingroup, the number of variable sites per locus ranged from 31 to 74. Proportional to the length of the locus, *VTN* contained the fewest variable sites (7.8% of total sites) and *TRSP* the most (10%). On average, about 30% of the variable sites were parsimony informative (PI), ranging from 23.3% in *CHRNA1* (7 PI sites) to 38.3% in *TRSP* (28 PI sites) (Table 3). Nucleotide frequencies were generally homogenous within each

Table 3
Characteristics of loci with gaps coded as missing/ or gaps coded for phylogenetic content

| | Total (bp) | Var.[a] sites | PI[b] sites | Coding region[c] | | | In[d] | Non-coding region[e] | | | In[d] | PI[b] In |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | L[f] (bp) | Var.[a] sites | PI[b] sites | | L[f] (bp) | Var.[a] sites | PI[b] sites | | |
| *CHRNA1* | 390/359 | | | 36 | | | | 354/323 | | | | |
| Ingroup | | 32/37 | 7/7 | | 2 | 1 | 0 | | 30/35 | 6/6 | 5 | 0 |
| Outgroup | | 134/148 | 81/86 | | 7 | 3 | 0 | | 127/141 | 78/83 | 14 | 5 |
| *CYPIA* | 874/832 | | | 275 | | | | 599/557 | | | | |
| Ingroup | | 53/57 | 10/11 | | 15 | 4 | 0 | | 38/41 | 6/7 | 4 | 1 |
| Outgroup[g] | | 214/236 | 116/130 | | 53 | 29 | 0 | | 161/183 | 87/10 | 22 | 14 |
| *FES* | 483/454 | | | 163 | | | | 320/291 | | | | |
| Ingroup | | 37/41 | 12/13 | | 3 | 2 | 0 | | 34/38 | 10/11 | 4 | 1 |
| Outgroup | | 160/178 | 120/131 | | 23 | 4 | 0 | | 137/155 | 116/127 | 18 | 11 |
| *GHR* | 839/764 | | | 301 | | | | 538/463 | | | | |
| Ingroup | | 49/57 | 18/20 | | 7 | 3 | 0 | | 42/50 | 15/17 | 6 | 2[h] |
| Outgroup | | 160/186 | 92/106 | | 46 | 25 | 0 | | 114/140 | 67/81 | 26 | 14 |
| *TRSP* | 740/715 | | | 87 | | | | 653/628 | | | | |
| Ingroup | | 74/82 | 28/35 | | 0 | 0 | 0 | | 74/82 | 28/35 | 11 | 6(2[h]) |
| Outgroup | | 249/275 | 157/176 | | 0 | 0 | 0 | | 249/275 | 157/176 | 26 | 19 |
| *VTN* | 490/464 | | | 224 | | | | 266/240 | | | | |
| Ingroup | | 31/36 | 9/12 | | 7 | 2 | 0 | | 24/29 | 7/10 | 5 | 3 |
| Outgroup | | 132/143 | 73/80 | | 32 | 19 | 0 | | 100/111 | 54/61 | 11 | 7 |
| Combined nuclear[i] | 3816/3588 | 1049/1166 | 639/709 | 1086 | 161 | 80 | 0 | 2730/2502 | 888/1005 | 559/629 | 117 | 70 |
| Mt[i,j] | 2001/NA[k] | 893 | 765 | 2001 | 893 | 765 | 0 | 0 | 0 | 0 | 0 | 0 |
| Nuclear + Mt[i,j] | 5817/5589 | 1942/2059 | 1404/1474 | 3087 | 1054 | 845 | 0 | 2730/2502 | 888/1005 | 559/629 | 117 | 70 |

[a] Variable sites.
[b] Parsimony informative.
[c] Coding region: In the *TRSP* locus it encodes a tRNA, in the other five loci the coding region encodes protein.
[d] Indels.
[e] Non-coding region: In the *TRSP* locus the non-coding region consists of the 5′ and 3′ flanking regions, in the other five loci the non-coding region consists of introns.
[f] Length in basepairs.
[g] Did not amplify in *Procyon lotor*.
[h] In repetitive elements, sequencing of more individuals may revel they are homoplastic.
[i] Numbers calculated include the outgroup.
[j] Mt, mitochondrial.
[k] Not applicable.

locus among the 30 taxa (data not shown). A total of 35 indels were inferred within the ingroup in the six loci, ranging in number from 4 (*CYPIA* and *FES*) to 11 (*TRSP*) per locus and ranging in size from 1 nucleotide to 108 nucleotides. Twenty-two of the indels were autapomorphic and 13 were phylogenetically informative. There are 117 indels among the ingroup and outgroup, including 19 synapomorphic indels shared by all canids, including a 246 bp deletion in *CYPIA* and a 195 insertion in *GHR*. Interspecific uncorrected pairwise differences range from 0 to 4.2% within the ingroup (*CHRNA1*: 0–3.2%, *CYPIA*: 0–3.0%, *FES*: 0–2.8%, *GHR*: 0.1–2.4%, *TRSP*: 0.1–4.2%, and *VTN*: 0–3.4%). If the outgroup is included, the maximum pairwise distance increases in range from 11.1% (*GHR*) to 22.4% (*TRSP*).

The combined nuclear DNA data set consisted of 3816 nucleotides (3588 nucleotides if indels are coded for phylogenetic information) (Table 3). About one-third (28.5%) of this data set was derived from coding region (exons and *TRSP*) and the remainder (71.5%) from non-coding regions (introns and non-transcribed flanking region from *TRSP*). A majority of the variable and parsimony informative sites, 84.5 and 87%, respectively, and all indels occurred in the non-coding regions. Coding indels for phylogenetic information increased the number of PI sites from 84 to 98 within the ingroup.

### 3.2. Phylogenetic analyses of individual loci

Each individual locus was analyzed using unweighted MP and ML. In general, each locus produced gene trees with one to four nodes having bootstrap values over 70%. Tree statistics for each locus are summarized in Table 4 and trees are shown in Supplementary Fig. 1. Several groupings were repeatedly seen in the individual gene trees, including monophyletic groupings of *Alopex, Vulpes*, and *Fennecus* (red fox-like canids) (five loci), the wolf-like canids (five loci), *C. adustus* and *C. mesomelas* (three loci), *P. griseus* and *P. gymnocercus* (three loci), *C. lupus, C. latrans, C. aureus*, and *C. familiaris* (two loci), and *Nycteruetes* and *Otocyon* (two loci). These groupings are not always supported with strong bootstrap values (>70%), which is likely due to the low number of phylogenetically informative sites within each locus.

### 3.3. Phylogenetic analyses of the combined nuclear data

We have taken the "conditional combination" approach to assess the combinability of the data (Bull et al., 1993; Dequeiroz, 1993; Flynn and Nedbal, 1998; Huelsenbeck et al., 1996). The trees produced with the individual loci using both MP and ML did not contain strongly supported nodes that were conflicting among the loci, and therefore, we concatenated the six nuclear loci into one data set. Combining sequences from multiple loci and performing a simultaneous analysis may reveal hidden phylogenetic signal or conflict that in itself is useful information when constructing a phylogeny (Gatesy et al., 1999). The combined nuclear data set was analyzed using MP, ML, and Bayesian analyses. MP analyses yielded four equally parsimonious trees (Table 4). The four trees differed only with regard to the groupings of *V. vulpes* and *V. corsac* with each other and whether the sister grouping of *Otocyon* + *Nycteruetes* is basal to the red fox clade or alternatively, whether the

Table 4
Tree Statistics of maximum parsimony and maximum likelihood analyses

| Locus | MP[a] | | | | | ML[b] | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | # trees | Tree length | CI[c] | RI[d] | # of nodes >70% BS | Model[e] | *I*[f] | α[g] | Ts/Tv[h] | # trees | –lnL | # of nodes >70% BS |
| *CHRNA1* | 344 | 192 | 0.8295 | 0.9264 | 1 | TrNef + G | | 1.6900 | | 4 | 1498.09 | 0 |
| *CYPIA* | 72 | 272 | 0.8742 | 0.9283 | 2 | HKY + G | | 1.1382 | 3.5 | 1 | 2604.42 | 1 |
| *FES* | 30 | 223 | 0.8497 | 0.9385 | 3 | HKY + G | | 1.3187 | 3.0 | 1 | 1718.51 | 2 |
| *GHR* | 3 | 239 | 0.7974 | 0.8994 | 3 | K81uf + G | | 0.4455 | | 1 | 2363.24 | 3 |
| *TRSP* | 54 | 440 | 0.7021 | 0.8598 | 4 | TrN + I + G | 0.2565 | 0.8842 | | 1 | 3002.71 | 2 |
| *VTN* | 2 | 169 | 0.8667 | 0.9453 | 2 | HKY + G | | 1.5706 | 2.5 | 1 | 1572.50 | 2 |
| Nuclear combined | 4 | 1560 | 0.7801 | 0.8958 | 11 | K81uf + G | | 0.6396 | | 1 | 13267.43 | 12 |
| Mt[i] combined | 1 | 4013 | 0.3011 | 0.4759 | 9 | GTR+I+G | 0.5164 | 0.9822 | | 1 | 17791.72 | 10 |
| Nuclear + Mt[i] | 1 | 5460 | 0.3953 | 0.5904 | 10 | GTR + I+ G | 0.3901 | 0.3607 | | 1 | 32476.45 | 15 |

[a] Maximum parsimony (MP) tree statistics generated with the data set that coded indels according to Barriel (1994).
[b] Maximum likelihood (ML) tree statistics generated with the data set were indels were treated as missing.
[c] Consistency index (excluding uninformative characters).
[d] Retention index.
[e] Models of substitution were determined using ModelTest (Posada and Crandall, 1998): TrNef (Tamura–Nei with equal base frequency) (Tamura and Nei, 1993), HKY Hasegawa-Kishino-Yano (Hasegawa et al., 1985), K81uf (two transversion parameters with unequal base frequency) (Kimura, 1981), TrN (Tamura and Nei, 1993), GTR (General Time Reversible) (Rodriguez et al., 1990), I = invariant sites, G = gamma.
[f] Proportion of invariable sites were determined using ModelTest (Posada and Crandall, 1998).
[g] Gamma distribution shape parameter were determined using ModelTest (Posada and Crandall, 1998).
[h] Transition/transversion ratio were estimated in ModelTest (Posada and Crandall, 1998).
[i] Mitochondrial, unweighted.

topological relationships of these two species remain unresolved. The tree generated by ML (Fig. 1, phylogram in Supplementary Fig. 2) was nearly identical to the two most parsimonious trees that place *Otocyon* + *Nycteruetes* basal to the red fox clade; the only difference between these MP trees and the tree generated by ML was the grouping among *C. familiaris*, *C. latrans*, and *C. aureus*. We also note that the strict consensus tree of the MP analyses carried out with the data set in which indels were coded as missing was iden-

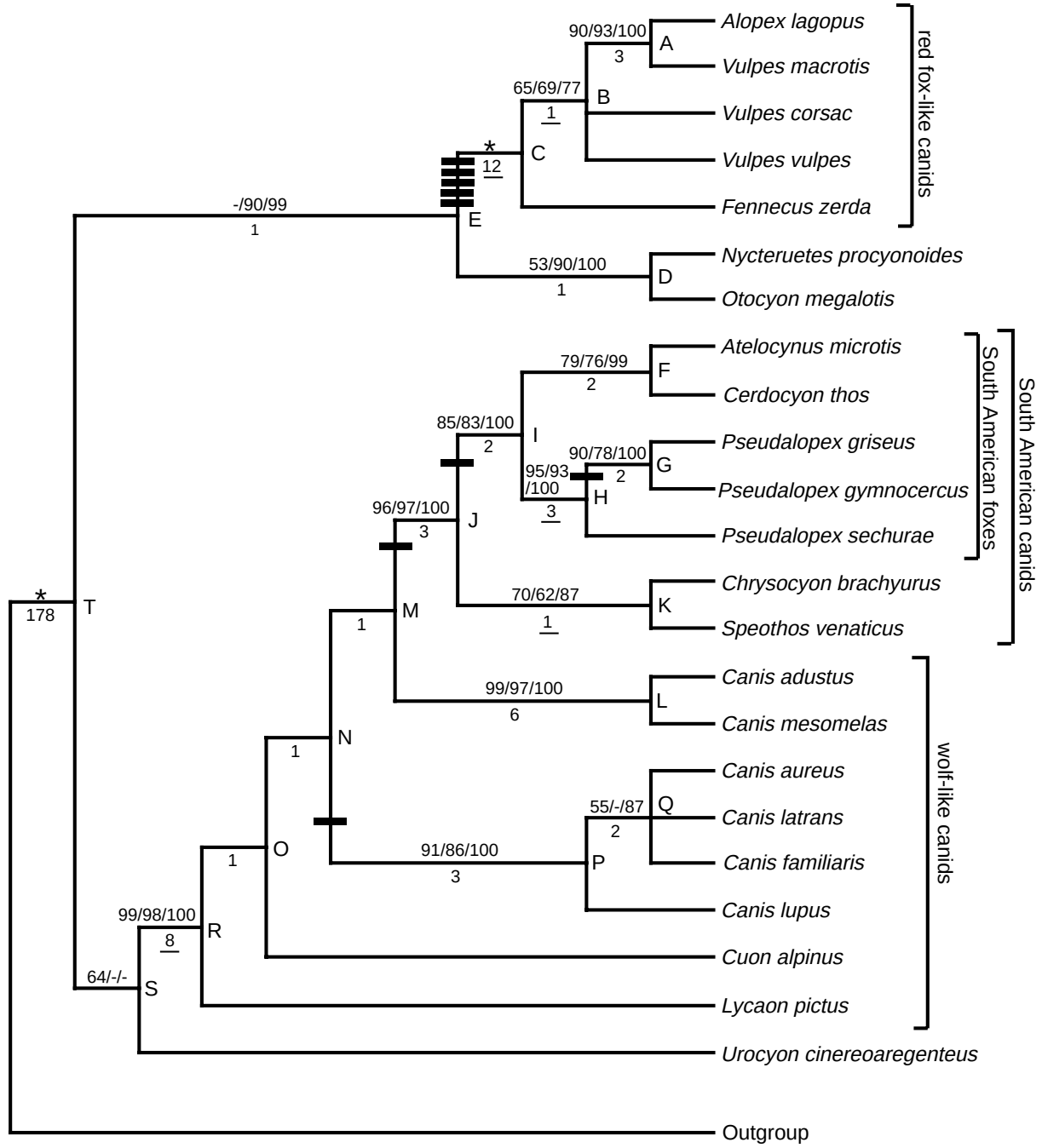


Fig. 1. Maximum likelihood (ML) tree based on the combined analysis of six nuclear loci (3816 bp) using the K81uf + G (two transversion parameters with unequal base frequency + gamma) model of sequence evolution. Bootstrap values (>50%) for maximum parsimony (MP) (out of 1000 pseudoreplicates), ML (out of 100 pseudoreplicates), and Bayesian posterior probability values (for the analysis in which the data set was not partitioned) are listed above the internodes, respectively. Nodes that receive 100% support from all analyses are indicated with an asterisk (*). Nodes are identified by letters. Branch support values are given below the internode. Values for nodes that do not show conflicting data as judged by partitioned Bremer support analysis are underlined. Parsimonious informative (PI) indels are indicated with a bar along the branch where they occur, phylogenetically informative indels that occur in repetitive elements are not included. The red fox-like clade, the South American fox clade, the South American canid clade, and wolf-like canids are indicated with brackets.

tical to the ML tree except in regard to the relationships among *C. familiaris*, *C. latrans*, and *C. aureus*. The differences between the tree generated by ML and the tree generated by Bayesian analysis in which the data are not partitioned are as follows: (1) *C. mesomelas* + *C. adustus*, the clade consisting of *C. familiaris*, *C. lupus*, *C. latrans*, and *C. aureus*, and the South American canid clade form a trichotomy in the Bayesian tree; and (2) *Urocyon* is not closely associated to the other canids in the Bayesian tree, whereas it is basal to the wolf-like and South American canids in the ML tree. In the Bayesian analysis with the data set partitioned by locus and using site-specific gamma values, the South American canids and the wolf-like canids are sister groups. Within the wolf-like canids clade, *Cuon* + *Lycaon* is a sister group to the clade consisting of *C. familiaris*, *C. lupus*, *C. latrans*, and *C. aureus*. The *C. mesomelas* + *C. adustus* clade is basal to the other wolf-like canids. In all other analyses, *Lycaon* is the basal taxon among the wolf-like canids. None of the alternative positions discussed above among the wolf-like canids are strongly supported nor is the placement of *Urocyon*. In general, all methods of analysis were consistent in the support given to a particular node (Supplementary Table 2). The red fox-like clade received strong support from all methods (100% BS, 100% PP), as did the relationship between *Alopex* and *V. marcrotis* as sister taxa (90–93% BS, 100% PP), whereas the basal placement of *Fennecus* to the other foxes received weak support. The red fox-like clade is supported by 14 synapomorphies including five indels. There is strong support for *Otocyon* and *Nycteruetes* as sister taxa from ML (90% BS) and Bayesian (100% PP) analyses and this pair of taxa share a total of three synapomorphies in two loci. The placement of *Otocyon* + *Nycteruetes* basal to the red fox-like clade receives high support from the ML and Bayesian analysis (90% BS, 99% PP). A clade of all South American canids consisting of *Atelocynus, Cerdocyon*, *Pseudalopex*, *Chrysocyon*, and *Speothos* is strongly supported (96–97% BS, 100% PP) by all analyses. This group shares two synapomorphies including an one nucleotide indel. Within the South American canid clade, a clade grouping all the South American foxes is supported (83–85% BS, 100% PP) and its members share three synapomorphies, including a 16 nucleotide deletion. *Pseudalopex* forms a monophyletic group (93–95% BS, 100% PP) within the South American fox clade, in which *P. griseus* and *P. gymnocercus* are sister taxa (78–90% BS, 100% PP) and share an indel and *P. sechurae* is basal to these taxa. The grouping of *Atelocynus* and *Cerdocyon* as sister taxa (76–79% BS, 99% PP) is moderately supported within the South American fox clade. *Chrysocyon* and *Speothos* joined as sister taxa has moderate support (62–70% BS, 87% PP) and they share one synapomorphic substitution. The South American canid clade falls within an unresolved clade of wolf-like canids, however, support in this region of the tree is weak. Within the clade of wolf-like canids, an association of *C. mesomelas* with *C. adustus* as sister taxa is strongly supported (97–99% BS, 100% PP) and this pair of taxa share four synapomorphies, including an indel (however, we note that this indel occurs in a poly-CT tract). A clade consisting of *C. familiaris*, *C. lupus*, *C. latrans*, and *C. aureus* is supported by all analyses (86–91% BS, 100% PP) and this group share four synapomorphies including one indel, however, there is little resolution within this clade. There is strong support that *Cuon* and *Lycaon* are related to the wolf-like canids, however, their branching order is uncertain.

#### 3.3.1. Intraspecific variation

Phylogenetic analyses of closely related species with short evolutionary histories may be confounded by hybridization and incomplete lineage sorting (Maddison, 1997; Nei, 1987; Pamilo and Nei, 1988). Consequently, we sequenced two to three individuals in 20 of the 23 ingroup species for all six nuclear loci. Intraspecific variation was low with uncorrected pairwise distances ranging from 0 to 0.8% across the six loci. Most intraspecific variation was due to point mutations; however, 10 indels (either within an individual or between individuals) were found, usually within homopolymer tracts or repeats (Supplementary Table 3).

We conducted MP analysis with the combined nuclear data set either with different individuals (or alleles) of a species or by including all sequenced individuals of the species. In some cases, changing the representatives of a species changed the number of most parsimonious trees found, ranging from 2 to 12. Generally, individuals from the same species grouped together with high (>70%) bootstrap support, including species in which intraspecific indels were inferred (tree not shown). Topological instability was noted in the *Pseudalopex* clade (only with the data set that coded indels as missing) and the clade including *C. latrans, C. lupus, C. familiaris*, and *C. aureus*. For example, one alternate *P. griseus* sequence dissolves the internal structure of the *Pseudalopex* clade, causing it to become a trichotomy. Likewise, *C. aureus* is either a member of a trichotomy with *C. latrans* and *C. familiaris* or is basal (74% BS) to *C. latrans, C. lupus*, and *C. familiaris*, depending on the individual used. This illustrates the usefulness of coding indels for resolving relationships among closely related species and also suggests that incomplete lineage sorting may confound relationships between *C. latrans, C. lupus, C. familiaris*, and *C. aureus*. More analyses with additional samples will be required to resolve this issue. Some of our samples, including those from *C. adustus, C. mesomelas, C. aureus*, and *V. corsac*, and two of the three individuals of *Otocyon* came from the same population. However, many samples are from zoos and therefore relationships between individual samples are unclear.

Future studies will address how individuals from across the geographic range may affect the resolution within terminal clades such as *Canis* and *Pseudalopex*.

#### 3.3.2. Branch support

To determine how much each of the six loci contributes to each node in the nuclear tree, a partitioned Bremer support (PBS) analysis was performed (Baker and Desalle, 1997). The PBS analysis indicates that the *TRSP* locus contributes the most phylogenetic signal (76.7%), whereas the *GHR* locus contributes more negative signal than positive (Supplementary Table 4). All other loci contribute positively to varying degrees (*CHRNA1*: 10.9%, CYPIA: 10.7%, *FES*: 7.4%, and *VTN*: 13.3%). Although the *GHR* gene tree does not contain well supported nodes in conflict with the nuclear tree, the fact that *GHR* contributes negatively in the simultaneous analysis indicates that this locus is not tracing the same phylogeny as the other loci. Possible reasons include differential lineage sorting (Pamilo and Nei, 1988), gene conversion (Rozas and Aguade, 1994) or selective pressure (Stewart et al., 1987). Removal of the exon portion of the *GHR* locus did not eliminate its negative contribution, suggesting that gene conversion and selective pressure are not the cause of conflict. Another possibility is that one rogue taxon may be responsible for the conflict. The position of *Urocyon* is uncertain (see Section 3.5). Removal of *Urocyon* from the data set eliminates the negative contribution of *GHR* and results in all loci contributing positively, with each loci contributing between 8–25% of the overall branch support (data not shown). The PBS analysis indicates that in 5 of the 19 nodes in the nuclear tree, there is no conflict among the six loci (underlined values in Fig. 1). Three of the five nodes received positive support from two to five loci (nodes C, H, R) and the other two were supported by a single locus (nodes B, K) (Supplementary Table 4). Although there is some conflict among the loci in most of the nodes, overall there is five times more hidden branch support than hidden conflict in the simultaneous analysis of the data set (Gatesy et al., 1999).

### 3.4. Mitochondrial DNA

Since the number and identity of taxa in the ingroup and the outgroup can affect the topology of a given tree and since the taxa used in the current study were not identical to that used in the original mitochondrial study, a modified version of the published phylogenetic tree derived from mitochondrial data (Wayne et al., 1997) was generated to make a direct comparison to the nuclear tree (Fig. 2). In addition, the original *COII Chrysocyon* sequence was replaced with a revised sequence (resolving some ambiguities we observed in the original autoradiographs).

The mitochondrial data set was analyzed using MP, ML, and Bayesian methods. In general, all methods of analysis agreed approximately in the level of node support (Supplementary Table 5). Strong support was found for the red fox-like clade and the relationships within this clade, the South American fox clade, the grouping of *Pseudalopex*, and the grouping of *C. familiaris*, *C. lupus, C. latrans*, and *C. aureus* (Fig. 2). Topological differences in the trees generated from the different analyses generally occurred in regions of the tree with low support, for example, within the South American fox clade and the wolf-like canid clade and the relationship among *Otocyon, Nycteruetes*, and *Urocyon* (Wayne et al., 1997). The main difference between the mitochondrial trees generated here and those from Wayne et al. (1997) is that *Chrysocyon* and *Speothos* no longer form a strongly supported clade that groups with the wolf-like canids as a result of the new *Chrysocyon COII* sequence. Depending on the analysis, *Chrysocyon* is grouped either with the South American foxes, the wolf-like canids or *Speothos*. *Speothos* is usually associated with the wolf-like canids. However, none of these associations receive high support.

### 3.5. Comparison of nuclear tree to mitochondrial tree

Although the mitochondrial data set is about 2/3 the size of the combined nuclear data set (2001 vs. 3588 characters) and contains six times as many PI sites (608 vs. 97 within the ingroup), the number of nodes resolved by each data set (nuclear: 19 and 20 nodes with MP and ML, respectively, mitochondrial: 22 nodes each with MP and ML) and the number of nodes with bootstrap values over 70% are similar (Table 4). As has been noted in many studies, the nuclear tree contains less homoplastic information than the mitochondrial tree (nuclear: CI = 0.7801, RI = 0.8958; mitochondrial: CI = 0.3011, RI = 0.4759) (Table 4) (Koepfli and Wayne, 2003; Matthee and Davis, 2001; Prychitko and Moore, 2000). A PBS analysis of the mitochondrial data indicates that there is no conflict among the three mitochondrial genes in 5 of the 22 nodes (underlined values in Fig. 2), compared to 5 of 20 nodes in the nuclear tree. Overall there is four times as much hidden support as hidden conflict in the mitochondrial data set. Therefore, in terms of the number of nodes that do not contain conflicting signal and the amount of hidden support over hidden conflict, the nuclear and mitochondrial data sets are similar.

Focusing on groups with high support, all phylogenetic analyses find four major clades in both the mitochondrial tree and the nuclear tree: the red fox-like clade, the South American fox clade, the *Pseudalopex* clade and a clade consisting of *C. aureus, C. familiaris, C. latrans*, and *C. lupus*. Mitochondrial DNA data are often considered to be better at resolving groupings between more recent species because of their more rapid

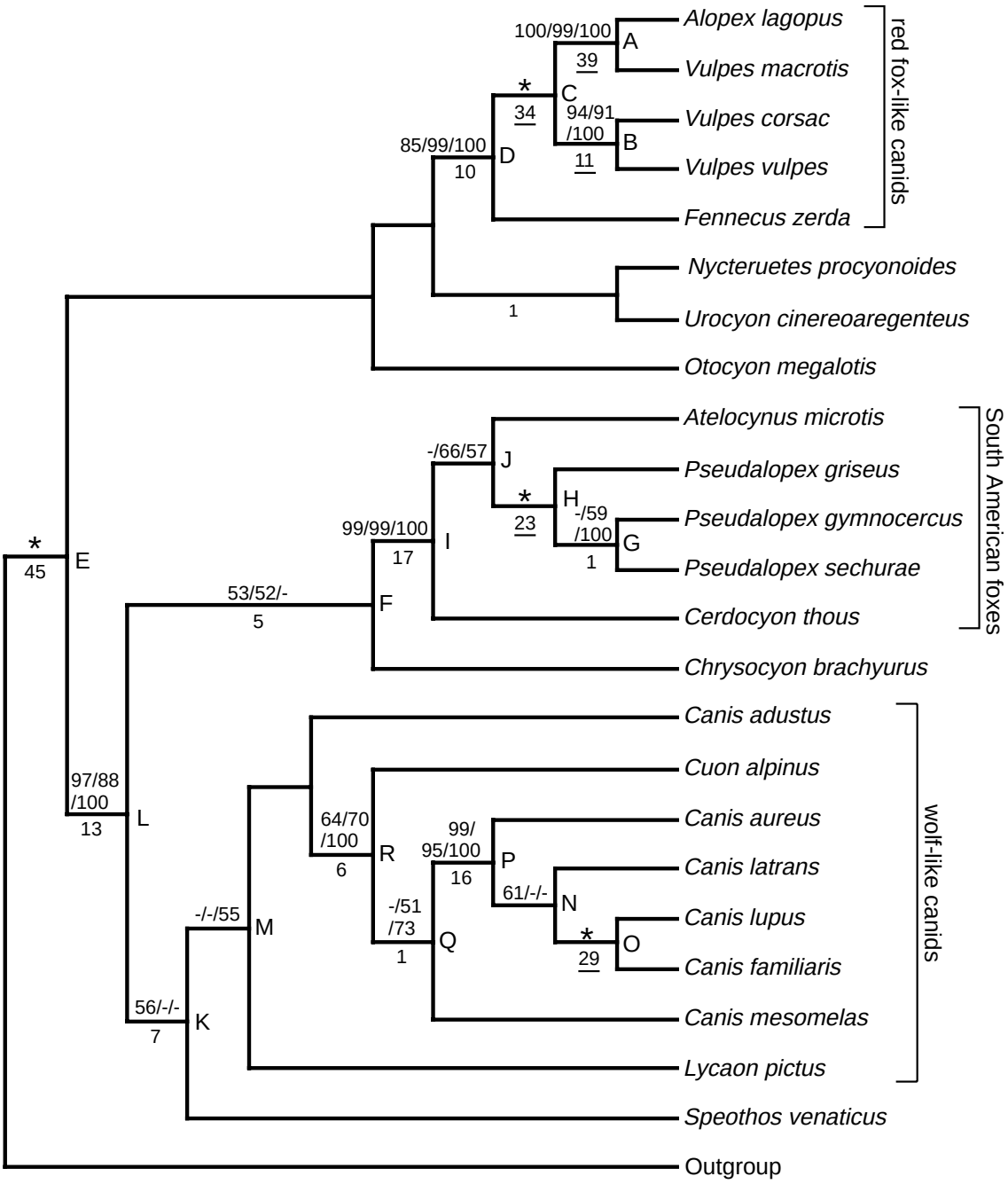


Fig. 2. Maximum likelihood (ML) tree based on the combined analysis of cytb, COI, and COII (2001 bp) using GTR+I+G (general time reversible + invariant sites + gamma) model of sequence evolution. Bootstrap values (>50%) for MP (out of 1000 pseudoreplicates), ML (out of 100 pseudoreplicates), and Bayesian posterior probability values (using the GTR+SSG3 model of sequence evolution) are listed above the internodes, respectively. Nodes are identified by letters. The figure is otherwise labeled as in Fig. 1. The red fox-like clade, the South American fox clade, and wolf-like canid clade are indicated with brackets.

rate of evolution (Miyata et al., 1982; Moritz et al., 1987). Accordingly, some of the relationships within the four main clades remain unresolved in the nuclear tree but receive high support from the mitochondrial tree. For example, the grouping *C. lupus* and *C. familiaris* as sister taxa, the grouping of *V. vulpes* and *V. corsac* as sister taxa and the placement of *Fennecus* basal to *Vulpes* + *Alopex* are all resolved and receive strong support in the mitochondrial tree but are not well resolved in the nuclear tree.

Some relationships differ between the nuclear tree and the mitochondrial tree, for example, the placement of *Chrysocyon* and *Speothos*. In the nuclear DNA phylogeny, all South American canids including *Chrysocyon* and *Speothos* form a clade with high support (Fig. 1). In the mitochondrial tree, *Chrysocyon* and

*Speothos* were either in a basal position to the wolf-like canids or *Chrysocyon* was weakly grouped with the South American foxes (Fig. 2). Constraining *Chrysocyon* and *Speothos* into a clade with the South American foxes on the mitochondrial tree is not significantly worse than the optimal ML tree (SH test, $P = 0.194$). Moving *Chrysocyon* and *Speothos* outside of the South American canid clade in the nuclear tree generates a significantly worse topology (SH test, $P < 0.05$). Therefore, with the current data a clade grouping all South American canids is favored. The grouping of *C. mesomelas* with *C. adustus* as sister taxa is strongly supported in the nuclear tree, but not inferred in the mitochondrial tree. Dissolving this clade in the nuclear tree generates a significantly worse topology than the optimal ML tree (SH test, $P < 0.05$) whereas grouping *C. adustus* with *C. mesomelas* as sister taxa in the mitochondrial tree did not produce a significantly worse tree (SH test, $P = 0.12$). Therefore, the current data supports the sister grouping of *C. adustus* with *C. mesomelas*.

As judged by nodal support, ML and Bayesian analyses of the nuclear data strongly support *Otocyon* and *Nycteruetes* as sister taxa and the placement of this group basal to the red fox-like canids. Mitochondrial trees generated with MP and weighing transitions over transversions by 8 or 11 also infer the same grouping and placement, but not with high bootstrap support. If the *Otocyon* + *Nycteruetes* clade is dissolved, the topology of the nuclear tree is marginally worse than the optimal topology generated by ML ($P = 0.052$), whereas there is no significant difference in the mitochondrial tree whether *Otocyon* and *Nycteruetes* are grouped as sister taxa or not ($P = 0.277$). In some of the original mitochondrial trees, *Otocyon* and *Nycteruetes* were inferred to be in basal positions of the canid tree (Wayne et al., 1997). In the nuclear tree, there was no significant difference in tree topology if the *Otocyon* + *Nycteruetes* sister group is placed basal to the red fox-like clade or basal to all other canids. However, if the sister grouping of *Otocyon* + *Nycteruetes* is dissolved and these taxa are then placed basal to all other canids, as suggested by the previously published mitochondrial tree, the resulting tree is significantly worse than the ML generated tree ($P < 0.05$). Constraining *Otocyon* and/or *Nycteruetes* + *Urocyon* to be basal to all other canids on the mitochondrial tree did not result in a significantly worse topology (SH test, $P = 0.192$–$0.267$). Therefore, the current data marginally suggests that *Otocyon* and *Nycteruetes* may be sister taxa, however, this relationship needs to be confirmed with additional data. There are insufficient data to determine whether *Otocyon* + *Nycteruetes* are associated with the red fox-like species or not closely associated with the other canids.

The placement of *Urocyon* within the mitochondrial and nuclear trees also differ. In the mitochondrial tree, *Urocyon* is either a sister to *Nycteruetes* and is basal to the red fox-like clade or alternatively, it is basal to all other canids. In the nuclear tree *Urocyon* is either basal to the wolf-like + South American canid clade or forms a trichotomy with the red fox-like clade + *Otocyon* + *Nycteruetes* and the wolf-like + South American canid clade. None of these placements receives high nodal support and therefore we cannot draw any conclusions in regard to the relationship of *Urocyon* to other canids. Because long-branch attraction is a greater problem with faster evolving sequences, the grouping of *Otocyon*, *Nycteruetes*, and *Urocyon* in the mitochondrial tree could be an artifact (Wayne et al., 1997).

Finally, an issue addressed in the previous mitochondrial DNA study (Wayne et al., 1997) was whether the trenchant heel, a complex morphological adaptation of the meat-processing tooth in the highly carnivorous canids (*Cuon, Lycaon*, and *Speothos*), evolved once or multiple times. Forcing *Cuon, Lycaon*, and *Speothos* in the nuclear tree into a single clade generated a significantly worse topology than the ML generated tree (SH test, $P < 0.05$) and consequently, the trenchant heel likely evolved at least twice.

### 3.6. Combined nuclear and mitochondrial data set

The fact that the nuclear tree and mitochondrial tree do not contain strongly conflicting nodes, except for one relationship within the *Pseudalopex* clade, suggests that it is reasonable to combine the two data sets. The topologies of the trees generated from the combined analysis from all methods were generally congruent and inferred three main clades: (1) the red fox-like canids plus *Otocyon* and *Nycteruetes*, (2) the South American canids and (3) the wolf-like canids (Fig. 3, phylogram Supplementary Fig. 3). The latter two clades formed sister groups. The relationships within the three clades are generally congruent with the following exceptions: (1) *Nycteruetes* is basal to *Otocyon* with unweighted MP (MP-UW), but they are grouped as sister taxa in all other analyses, (2) *Atelocynus* is basal to *Cerdocyon* with MP-UW and implied weighting (IW) MP, but grouped as sister taxa in ML and Bayesian analyses, and (3) the relationship among *Lycaon*, *Cuon*, *C. mesomelas*, and *C. adustus* varies according to the method of phylogenetic analysis. For example, in two of the analyses (MP-IW, Bayes-SSG4), *Canis* is monophyletic, with the *C. mesomelas* + *C. adustus* clade as sister to the clade consisting of *C. aureus, C. latrans, C. lupus*, and *C. familiaris*, and *Lycaon* as the most basal member of the wolf-like canids (Fig. 3). In two other analyses (ML, Bayes-GTR+I+G), *Cuon* is basal to a clade grouping *C. mesomelas* with *C. aureus, C. latrans, C. lupus*, and *C. familiaris* and the basalmost member among the wolf-like canids is either *Lycaon* (Bayes) or *C. adustus* (ML).

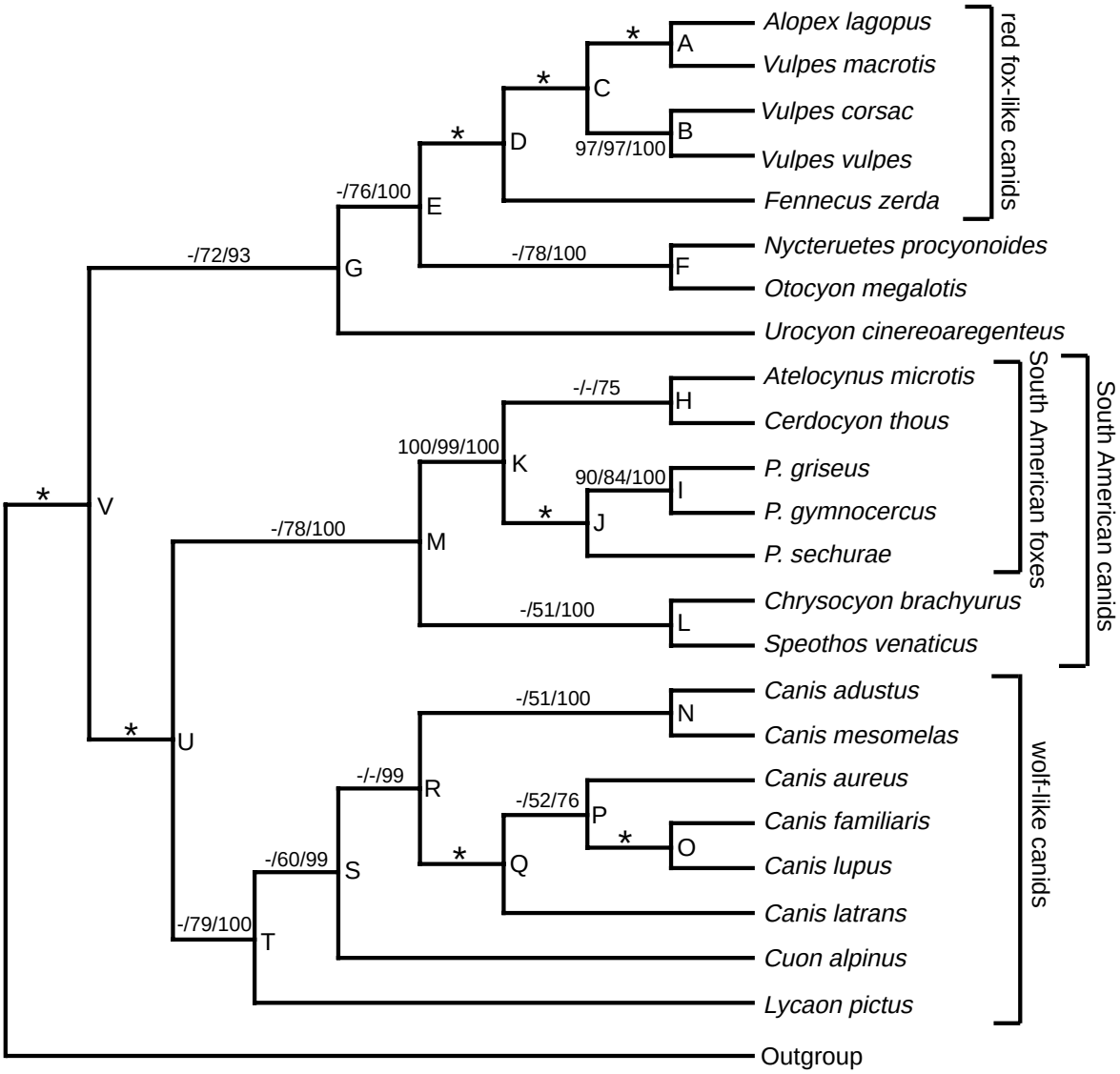


Fig. 3. Majority-rule consensus tree of 5 million MCMC generations of the Bayesian phylogenetic analysis of the combined nuclear and mitochondrial data sets (5817 bp) using the GTR+SSG4 model of sequence evolution. The initial 1.25 million generations were discarded as burn in. Bootstrap values (>50%) for MP-unweighted (out of 1000 pseudoreplicates), ML (out of 100 pseudoreplicates), and Bayesian posterior probability values from the GTR+SSG4 analysis are listed above the internodes, respectively. Nodes that receive 100% support from all analyses are indicated with an asterisk (*). Nodes are identified by letters. The red fox-like clade, the South American fox clade, the South American canid clade, and wolf-like canid clade are indicated with brackets.

In the combined analysis, many of the relationships that were strongly supported in the mitochondrial tree remain strongly supported in the combined nuclear + mitochondrial tree (nodes A, B, C, D , J, K, U, O, and Q in Fig. 3). Most of these nodes receive moderate to strong support in the nuclear tree as well (e.g., nodes A, C, D, J, K, U, and Q in Fig. 3). Combining the nuclear and mitochondrial data strengthened the support at six nodes (nodes B, I, J, K ,U, and Q in Fig. 3). Judging by UW-MP and ML bootstrap values, many of the relationships inferred from the nuclear tree that are different than those in the mitochondrial tree lose bootstrap support in the combined analysis (nodes E, F, H, M, L, and N in Fig. 3). This is likely due to conflicting phylogenetic signal between the nuclear and mitochondrial data sets. PBS analysis indicates that much of the conflict between the nuclear and mitochondrial data in the combined tree occurs at the aforementioned nodes (data not shown). Moreover, the addition of mitochondrial data causes higher levels of homoplasy, as indicated by the lower CI and RI values of the combined data set compared to the nuclear data set alone (Table 4), and results in lower bootstrap values as well (Matthee and Davis, 2001; Zharkikh and Li, 1995). Bayesian analysis often generated high posterior probability values but otherwise levels of support generally agreed (Table 5). Notably, some nodes that had low bootstrap support in MP-UW and ML analyses but high support from

Table 5
Bootstrap and posterior probability values from different methods of phylogenetic analyses on the nuclear + mitochondrial data set

| Node[a] | Method of analysis[b] | | | | |
|---|---|---|---|---|---|
| | MP-UW | MP-IW | ML | Bayes | Bayes SSG4 |
| A | 100 | 100 | 100 | 100 | 100 |
| B | 97 | 99 | 97 | 100 | 100 |
| C | 100 | 100 | 100 | 100 | 100 |
| D | 100 | 100 | 100 | 100 | 100 |
| E | — | 51 | 76 | 100 | 100 |
| F | — | — | 78 | 100 | 100 |
| G | — | — | 72 | 88 | 93 |
| H | — | — | — | 59 | 75 |
| I | 90 | 95 | 84 | 100 | 100 |
| J | 100 | 100 | 100 | 100 | 100 |
| K | 100 | 100 | 99 | 100 | 100 |
| L | — | 91 | 51 | 88 | 100 |
| M | — | 90 | 78 | 99 | 100 |
| N | — | 79 | 51 | — | 100 |
| O | 100 | 100 | 100 | 100 | 100 |
| P | — | — | 52 | — | 76 |
| Q | 100 | 100 | 100 | 100 | 100 |
| R | — | — | — | — | 99 |
| S | — | — | 60 | — | 99 |
| T | — | 72 | 79 | 99 | 100 |
| U | 100 | 100 | 100 | 100 | 100 |
| V | 100 | 100 | 100 | 100 | 100 |

[a] Refer to Fig. 3 for identity of nodes.
[b] MP-UW, maximum parsimony unweighted; MP-IW, MP with implied weighting ($k = 2$); ML, maximum likelihood using GTR+I+G; Bayes, Bayesian analysis using GTR+I+G; Bayes-SSG4, Bayesian analysis using GTR and gamma determined separately for each of four partitions consisting of the nuclear DNA, and the three codon positions in the mitochondrial DNA.

Bayesian analysis also receive high support from MP-IW analysis (nodes L, M, and N in Fig. 3) (Table 5). All these nodes were highly supported in the nuclear tree. Two nodes (R and S) exclusively received strong support from the Bayesian analysis. These define relationships within the wolf-like canids. Although posterior probability values are often higher than bootstrap values, their accuracy is also dependent on the model used in the analysis (Alfaro et al., 2003; Douady et al., 2003; Minin et al., 2003; Suzuki et al., 2002). Although the GTR-SSG4 model offers a better fit to the data, this may not be the appropriate model (Minin et al., 2003). Site-specific rate models have been criticized by some (Buckley et al., 2001, 2002), but also successfully used by others (Danforth et al., 2003). Because the two nodes (R and S) described above only receive high support from the Bayesian analysis using the GTR-SSG4 model, more data will be required to assess the validity of those relationships.

Taken together with the indel data, we interpret our data to indicate that the red fox-like clade and the relationships within it are well supported, at least with the taxa included in this study. The sister group of *Otocyon* + *Nycteruetes* is supported by ML and Bayesian analysis. The basal position of *Otocyon* + *Nycteruetes* to the red fox clade requires additional data to confirm or refute this placement. The monophyly of the South American canids, the South American foxes, *Pseudalopex*, and *Chrysocyon* + *Speothos* are strongly supported. The relationships within *Pseudalopex* are strongly supported (*P. sechurae* basal to *P. griseus* + *P. gymnocercus*), however, whether *Atelocynus* and *Cerdocyon* are sister taxa or not will require more data. Except for a monophyletic clade consisting of *C. familiaris, C. lupus, C. latrans*, and *C. aureus*, the close relationship of *C. familiaris* with *C. lupus* and the sister taxa grouping of *C. adustus* and *C. mesomelas*, the relationships among the wolf-like canids are not well resolved. Lastly, the relationship of *Urocyon* to the other canids remains uncertain.

## 4. Discussion

Although gene trees are often assumed to accurately reflect species trees, the stochastic nature of lineage sorting can result in differences between them (Pamilo and Nei, 1988). By comparing gene trees of different linkage groups, such as nuclear genes from different chromosomes or nuclear versus mitochondrial genes, areas of congruence can be identified and used as evidence for organismal history. The phylogeny we have generated using six independent nuclear loci generally agrees with the previous mitochondrial DNA phylogeny (Wayne et al., 1997), thus confirming inferred relationships. Novel relationships suggested by the nuclear tree include a clade of all South American canids, *C. mesomelas*, and *C. adustus* as sister taxa and *Otocyon* and *Nyctereutes* as sister taxa. Although the nuclear data did not provide many parsimony informative sites, they did provide phylogenetically informative indels that are especially informative because they are rare events (Rokas and Holland, 2000). Monophyletic groups that are supported by indels include: (1) the red fox-like canids; (2) the South American canids; (3) the South American foxes; (4) *P. gymnocercus* and *P. griseus*; (5) *C.lupus, C. latrans, C. familiaris*, and *C. aureus*; and (6) *C. adustus* and *C. mesomela*s. The mitochondrial and combined nuclear tree do not contain strongly conflicting nodes and three main clades are inferred: the red fox-like, South American and the wolf-like canids. Relationships within these clades are discussed below.

### 4.1. Red fox-like canids

This clade is well supported by all phylogenetic analyses and by five indels. The groupings within this clade of *Vulpes macrotis* plus *Alopex* and *V. vulpes* plus *V. corsac*, which together define a monophyletic group, is consistent with the Holoarctic clade suggested by Zrzavy and Ricankova (2004). The placement of *Fennecus* basal to

this clade is also consistent with previous studies (Geffen et al., 1992; Wayne et al., 1997; Zrzavy and Ricankova, 2004)

### 4.2. South American canids

The monophyly of all South American canids is highly supported in the nuclear tree and by most analyses of the combined nuclear + mitochondrial data set. This clade receives one of four highest branch supports in the nuclear tree and all members in this clade share a one base pair indel in *TRSP*. Moreover, there was some support for this clade with the combined analysis of morphological and mitochondrial DNA data (Wayne et al., 1997). These results suggest that the South American canids are likely a monophyletic group.

In the previous mitochondrial phylogeny (Wayne et al., 1997) *Chrysocyon* and *Speothos* were grouped as sister taxa and more closely related to the wolf-like canids than to the South American foxes. The replacement of the original *COII* sequence with the one generated from the current studies (Bardeleben et al., 2005; this study) caused the sister taxa grouping of *Chrysocyon* and *Speothos* to be dissolved in the mitochondrial tree and suggested weak support for an association with *Chrysocyon* as basal to the South American foxes or the wolf-like canids. However, the nuclear phylogeny supported the grouping of *Chrysocyon* and *Speothos* with moderate bootstrap support and a synapomorphic substitution. Further, a constrained ML mitochondrial tree with these two taxa grouped was no worse than the optimal tree. These taxa are also grouped by a synapomorphic indel of three nucleotides in *RPPH1*, a nuclear locus not used in this study due to the incompleteness of this data set (12 of 23 taxa in the ingroup) (Bardeleben et al., 2005). Morphological studies do not group these taxa as sister taxa, instead grouping *Speothos* with *Atelocynus*, or with *Cerdocyon* and *Atelocynus* (Lyras and Van Der Geer, 2003; Tedford et al., 1995 and refs within). Conflicts in relationships between morphological and DNA data may reflect the influence of ecological, behavioral or physiological factors on morphologic traits. Therefore, convergent evolution rather than shared ancestry may be responsible for some inferred relationships based on morphology. Non-coding DNA, such as that utilized in the current study, is expected to evolve in a neutral fashion and therefore is probably a more reliable estimate of inferred relationships. Thus, within the South American canid clade the molecular data suggests that two of the most morphologically divergent species, *Chrysocyon* and *Speothos*, shared a common ancestor.

Within the South American canid clade, the South American foxes form a clade that is strongly supported by all analyses and by a 16 nucleotide deletion in the *VTN* locus. Further, our previous molecular study found two indels in the *RPPH1* locus in the two sequenced representatives of this clade (*Cerdocyon* and *P. griseus*) (Bardeleben et al., 2005). Morphological data does not group *Cerdocyon* and *Atelocynus* with *Pseudalopex*, and instead *Cerdocyon* is grouped with *Nyctereutes*, and *Atelocynus* is grouped with *Speothos* (Tedford et al., 1995). Again, the conflict between the morphological and DNA data may be due to convergent evolution of some of the morphological characters that were used. In agreement with the mitochondrial DNA tree and contrary to the morphological data, the nuclear data do not support the placement of *Nyctereutes* with the South American foxes (Berta, 1987; Tedford et al., 1995; Wayne et al., 1997). Within the South American fox clade, *Pseudalopex* forms a monophyletic group. *P. griseus* and *P. gymnocercus* are sisters and share an indel and *P. sechurae* is basal to these taxa in the nuclear trees. The relationships within this clade, however, show some instability when different individuals are used in the nuclear data set (only if indels are coded as missing). The grouping of *P. sechurae*, *P. griseus*, and *P. gymnocercus* in a clade is consistent with morphological studies and those combining morphology and molecular data (Berta, 1987; Lyras and Van Der Geer, 2003; Tedford et al., 1995; Zrzavy and Ricankova, 2004).

The nuclear data provide moderate support for the monophyly of *Cerdocyon* and *Atelocynus*, whereas the support for this group decreased in the combined nuclear + mitochondrial tree. This relationship was suggested in the distance tree of mitochondrial DNA sequences (Wayne et al., 1997) but not well supported. Additional sequences and inclusion of additional South American foxes will be required to determine whether *Atelocynus* and *Cerdocyon* are monophyletic or paraphyletic.

### 4.3. Wolf-like canids

Although the nuclear + mitochondrial tree supports a clade consisting of *Lycaon, Cuon*, and *Canis*, the relationships among these genera are not well resolved. All analyses support *C. aureus*, *C. lupus, C. latrans*, and *C. familiaris* as a monophyletic group. In the mitochondrial tree, *C. aureus* is basal to the other three taxa in this clade whereas in the nuclear tree, the position of *C. aureus* was dependent on the individual representing the species. The grouping of *C. adustus* with *C. mesomelas* is highly supported, having the third highest branch support in the nuclear tree. *C. adustus* and *C. mesomelas* have been grouped based on morphological similarity with the South American foxes (Clutton-Brock et al., 1976). However, more recent studies assign these two taxa to the wolf-like canids (Bininda-Emonds et al., 1999; Wayne et al., 1997) and some evidence suggests that *C. adustus* and *C. mesomelas* do not group with the rest of *Canis* (Zrzavy and Ricankova, 2004).

The nuclear and mitochondrial data do not support *Cuon* and *Lycaon* as a monophyletic group, contrary to trees derived from morphological data (Tedford et al., 1995; Zrzavy and Ricankova, 2004) and instead weakly support *Lycaon* as the basal-most wolf-like canid. Both *Cuon* and *Lycaon* share with *Speothos* a modification of the meat-shearing carnassial teeth (upper P4 and lower M1), the trenchant heel, indicative of hypercarnivory (Ewer, 1973; Van Valkenburgh, 1990). The trenchant heel has been hypothesized to have evolved once and lost in the descendents of the *Speothos* and the *Cuon*/*Lycaon* lineages. Forcing a monophyletic grouping of the hypercarnivorous canids, *Lycaon*, *Cuon*, and *Speothos*, results in a significantly worse topology than the one generated by ML, supporting the hypothesis that the trenchant heel evolved more than once (Tedford et al., 1995); this is consistent with studies suggesting that hypercarnivores cannot easily reverse to a more generalized state and that among large fossil carnivorans, this seemingly complex character can readily evolve among closely related canids (Holliday and Steppan, 2004; Van Valkenburgh, 1991).

### 4.4. Otocyon, Nyctereutes, and Urocyon

The relationship among *Otocyon*, *Nyctereutes*, and *Urocyon* and the other canids is not clear. Analyses of nuclear, mitochondrial, and morphologic data have each resulted in different placements of these taxa within the canid phylogeny. For example, morphologic studies have grouped *Nyctereutes* with the South American canids (Berta, 1987; Lyras and Van Der Geer, 2003; Tedford et al., 1995). Molecular studies place *Nyctereutes* either with *Vulpes* or in a basal position in the canid tree with no closely related living species (Wayne and O'Brien, 1987; Wayne et al., 1997, 1987a,b). A supertree analysis places it as a basal member to the wolf-like and South American canid clade (Bininda-Emonds et al., 1999). *Otocyon* has been grouped with *Urocyon* based on morphology (Tedford et al., 1995), with *Fennecus* and *Urocyon* based on chromosomal data (Wayne et al., 1987a,b) or not closely related to any other taxa based on allozymes and the original analysis of the mitochondrial data (Wayne and O'Brien, 1987; Wayne et al., 1997). Multiple groupings have also been proposed for *Urocyon*. Morphologically it has been placed with *Vulpes* (Tedford et al., 1995) or judged not to be grouped with *Vulpes* (Lyras and Van Der Geer, 2003). Mitochondrial sequences or a combination of mitochondrial sequences with morphological data place *Urocyon* basal to all other canids (Wayne et al., 1997; Zrzavy and Ricankova, 2004).

In our study, ML and Bayesian analyses of the nuclear and nuclear + mitochondrial data suggests a sister grouping of *Otocyon* and *Nyctereutes*. The distance tree in the original mitochondrial analysis and the MP *COII* gene tree also suggested this grouping (Wayne et al., 1997; Zrzavy and Ricankova, 2004). However, this association is not strongly supported by MP analyses or the mitochondrial tree with the methods of analyses used in this study and more data will be required to confirm or refute this grouping. Some association of *Otocyon* and *Nyctereutes* with *Urocyon* was observed in the original mitochondrial study, but long-branch attraction was thought to be a confounding issue (Wayne et al., 1997). Long-branch attraction should be less of an issue with nuclear DNA due to its slower evolutionary rate. Our data are not sufficient to infer whether *Otocyon* + *Nyctereutes* should be basal to the red fox-like clade or basal to all other canids. However, based on sequence divergence, *Otocyon*, *Nyctereutes*, and *Urocyon* clearly are higher divergent lineages not closely related to other living canids.

## 5. Conclusions

In general, the conclusions drawn from the nuclear DNA analysis and combined nuclear + mitochondrial data set are consistent with those of the original mitochondrial study (Wayne et al., 1997) and of Zrzavy and Ricankova (2004) in their analysis of various data partitions of morphological and mitochondrial DNA data. In agreement with the previous studies, the current data support the following phylogenetic conclusions: (1) *Alopex, Fennecus*, and *Vulpes* form a monophyletic group; (2) the South American canids and the wolf-like canids are sister taxa; (3) the South American foxes are monophyletic; (4) *Pseudalopex* is monophyletic; and (5) the branching patterns within the red fox-like clade and the *Pseudalopex* clade are resolved topologically (with the taxa used in the current study). Differences include the following: (1) the placement of *Chrysocyon* and *Speothos* in the South American canids clade; (2) the grouping of *C. adustus* and *C. mesomelas* as sister taxa; and (3) the grouping of *Nyctereutes* and *Otocyon* as sister taxa. Several issues remain unresolved including whether *Atelocynus* is sister to *Cerdocyon*, the relationships within *Canis* and the relationship of *Canis* to *Cuon* and *Lycaon*, and the placement of *Urocyon* and *Nyctereutes* + *Otocyon*. However, in general, the resolution of the nuclear and combined nuclear-mitochondrial tree provides a surprising degree of consistency with past molecular studies and allows new insights into the evolutionary history canids. Our analysis, as well as many others (Johnson and Clayton, 2000; Koepfli and Wayne, 2003; Matthee and Davis, 2001) demonstrates how even a limited amount of nuclear DNA can help to resolve the relationships of recently radiated taxa.

## Acknowledgments

We are grateful to Klaus-Peter Koepfli for advice on phylogenetic analysis, his comments on the manuscript

and providing many of the DNA samples for the outgroup. We thank Nicholas Manoukis for his kindness and generosity for running our analyses on his Unix computers. We also thank Isabel Amorim do Rosario, Jennifer A. Leonard, Kristina D. Louie, Carles Vila, Chris Winchell, and Gary T. Shin for helpful comments and discussion, and C. Cicero, Museum of Vertebrate Zoology for tissue samples of small-eared dog and black bear, O. Ryder from CRES for DNA samples of bush dog and maned wolf, and L. Waits for DNA samples of black bear. We appreciate the thorough reading and helpful comments from two anonymous reviewers. The research was supported in part by an NSF grant to RKW.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2005.07.019.

## References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. AC-19, 716–723.

Alfaro, M.E., Zoller, S., Lutzoni, F., 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. 20, 255–266.

Baker, R.H., Desalle, R., 1997. Multiple sources of character information and the phylogeny of Hawaiian Drosophilids. Syst. Biol. 46, 654–673.

Bardeleben, C., Moore, R.L., Wayne, R.K., 2005. Isolation and molecular evolution of the selenocysteine tRNA (Cf TRSP) and RNase P RNA (Cf RPPH1) genes in the dog family, Canidae. Mol. Biol. Evol. 22, 347–359.

Barriel, V., 1994. Molecular phylogenies and how to code insertion/deletion events. Comp. Rend. L'Acad. Sci.—Ser. Iii, Sci. Vie 317, 693–701.

Berta, A., 1987. Origin, diversification and, zoogeography of the South American Canidae. Fieldiana Zool 39, 455–471.

Bininda-Emonds, O.R.P., Gittleman, J.L., Purvis, A., 1999. Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). Biol. Rev. (Cambridge) 74, 143–175.

Bremer, K., 1988. The limits of amino acid sequence data in Angiosperm phylogenetic reconstruction. Evolution 42, 795–803.

Bremer, K., 1994. Branch support and tree stability. Cladistics—Int. J. Willi Hennig Soc. 10, 295–304.

Buckley, T.R., Simon, C., Chambers, G.K., 2001. Exploring among-site rate variation models in a maximum likelihood framework using empirical data: effects of model assumptions on estimates of topology, branch lengths, and bootstrap support. Syst. Biol. 50, 67–86.

Buckley, T.R., Arensburger, P., Simon, C., Chambers, G.K., 2002. Combined data, Bayesian phylogenetics, and the origin of the New Zealand cicada genera. Syst. Biol. 51, 4–18.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., Waddell, P.J., 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42, 384–397.

Castoe, T.A., Doan, T.M., Parkinson, C.L., 2004. Data partitions and complex models in Bayesian analysis: the phylogeny of Gymnophthalmid lizards. Syst. Biol. 53, 448–469.

Chippindale, P.T., Wiens, J.J., 1994. Weighting, partitioning, and combining characters in phylogenetic analysis. Syst. Biol. 43, 278–287.

Clutton-Brock, J., Corbett, G.B., Hills, M., 1976. A review of the family Canidae with a classification by numerical methods. Bull. Br. Mus. Zool. 29, 119–199.

Danforth, B.N., Conway, L., Ji, S., 2003. Phylogeny of eusocial lasioglossum reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera: Halictidae). Syst. Biol. 52, 23–36.

Debry, R.W., 1999. Maximum likelihood analysis of gene-based and structure-based process partitions, using mammalian mitochondrial genomes. Syst. Biol. 48, 286–299.

Debry, R.W., Olmstead, R.G., 2000. A simulation study of reduced tree-search effort in bootstrap resampling analysis. Syst. Biol. 49, 171–179.

Dequeiroz, A., 1993. For consensus (sometimes). Syst. Biol. 42, 368–372.

Douady, C.J., Delsuc, F., Boucher, Y., Doolittle, W.F., Douzery, E.J.P., 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. Mol. Biol. Evol. 20, 248–254.

Dumbacher, J.P., Pratt, T.K., Fleischer, R.C., 2003. Phylogeny of the owlet-nightjars (Aves: Aegothelidae) based on mitochondrial DNA sequence. Mol. Phylogenet. Evol. 29, 540–549.

Ewer, R.F., 1973. The Carnivores. Cornell University Press, Ithaca, New York.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Flynn, J.J., Nedbal, M.A., 1998. Phylogeny of the Carnivora (Mammalia)-congruence Vs incompatibility among multiple data sets. Mol. Phylogenet. Evol 9, 414–426.

Gatesy, J., O'grady, P., Baker, R.H., 1999. Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics 15, 271–313.

Geffen, E., Mercure, A., Girman, D.J., Macdonald, D.W., Wayne, R.K., 1992. Phylogenetic relationships of the fox-like canids: mitochondrial DNA restriction fragment, site and cytochrome beta sequence analyses. J. Zool. 228, 27–39.

Goloboff, P.A., 1993. Estimating character weights during tree search. Cladistics 9, 83–91.

Hall, B.G., 2001. Phylogenetic Trees Made Easy: A How -to Manual For Molecular Biologists. Sinauer, Sunderland, Mass.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21, 160–174.

Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.

Holliday, J.A., Steppan, S.J., 2004. Evolution of hypercarnivory: the effect of specialization on morphological and taxonomic diversity. Paleobiology 30, 108–128.

Huelsenbeck, J.P., Bull, J.J., Cunningham, C.W., 1996. Combining data in phylogenetic analysis. Trends Ecol. Evol. 11, 152–158.

Huelsenbeck, J.P., Larget, B., Miller, R.E., Ronquist, F., 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Syst. Biol. 51, 673–688.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., Gibson, T.J., 1998. Multiple sequence alignment with clustal X. Trends Biochem. Sci. 23, 403–405.

Jiang, Z., Priat, C., Galibert, F., 1998. Traced orthologous amplified sequence tags (TOASTs) and mammalian comparative maps. Mamm. Genome 9, 577–587.

Johnson, K.P., Clayton, D.H., 2000. Nuclear and mitochondrial genes contain similar phylogenetic signal for pigeons and doves (Aves: Columbiformes). Mol. Phylogenet. Evol. 14, 141–151.

Kimura, M., 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Natl. Acad. Sci. USA 76, 454–458.

Koepfli, K.-P., Wayne, R.K., 2003. Type I STS Markers are more informative than Cytochrome *b* in phylogenetic reconstruction of the mustelidae (Mammalia: Carnivora). Syst. Biol. 52, 571–593.

Lee, M.S.Y., 2000. Tree robustness and clade significance. Syst. Biol. 49, 829–836.

Lyons, L.A., Laughlin, T.F., Copeland, N.G., Jenkins, N.A., Womack, J.E., 0'Brien, S.J., 1997. Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. Nature Genetics 15, 47–56.

Lyras, G.A., Van Der Geer, A.A.E., 2003. External brain anatomy in relation to the phylogeny of Caninae (Carnivora: Canidae). Zool. J. Linnean Soc. 138, 505–522.

Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.

Matthee, C.A., Burzlaff, J.D., Taylor, J.F., Davis, S.K., 2001. Mining the mammalian genome for artiodactyl systematics. Syst. Biol. 50, 367–390.

Matthee, C.A., Davis, S.K., 2001. Molecular insights into the evolution of the family Bovidae: a nuclear DNA perspective. Mol. Biol. Evol. 18, 1220–1230.

Minin, V., Abdo, Z., Joyce, P., Sullivan, J., 2003. Performance-based selection of likelihood models of phylogeny estimation. Syst. Biol. 52, 674–683.

Mitchell, A., Mitter, C., Regier, J.C., 2000. More taxa or more characters revisited: combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insecta: Lepidoptera). Syst. Biol. 49, 202–224.

Miyamoto, M.M., Fitch, W.M., 1995. Testing species phylogenies and phylogenetic methods with congruence. Syst. Biol. 44, 64–76.

Miyata, T., Hayashida, H., Kikuno, R., Hasegawa, M., Kobayashi, M., Koike, K., 1982. Molecular clock of silent substitutions: at least a sixfold preponderance of silent changes in mitochondrial genes over those in nuclear genes. J. Mol. Evol. 19, 28–35.

Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mtDNA: relevance for population biology and systematics. Annu. Rev. Ecol. Syst 18, 269–292.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Nowak, R.M., 1999. Walker's Mammals of the World. Johns Hopkins University Press, Baltimore.

Olmstead, R.G., Sweere, J.A., 1994. Combining data in phylogenetic systematics—an empirical approach using three molecular data sets in the Solanaceae. Syst. Biol. 43, 467–481.

Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5, 568–583.

Penny, D., Hendy, M., 1986. Estimating the reliability of evolutionary trees. Mol. Biol. Evol. 3, 403–417.

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

Prychitko, T.M., Moore, W.S., 2000. Comparative evolution of the mitochondrial cytochrome *b* gene and nuclear beta-fibrinogen intron 7 in woodpeckers. Mol. Biol. Evol. 17, 1101–1111.

Rodriguez, F., Oliver, J.L., Marin, A., Medina, J.R., 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142, 485–502.

Rokas, A., Holland, P.W.H., 2000. Rare genomic changes as a tool for phylogenetics. Trends Ecol. Evol. 15, 454–459.

Rokas, A., Williams, B.L., King, N., Carroll, S.B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425, 798–804.

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

Rozas, J., Aguade, M., 1994. Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of Drosophila. Proc. Natl. Acad. Sci. USA 91, 11517–11521.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning a Laboratory Manual 2nd ed. vols 1– 3. Cold Spring Harbor Laboratory Press, New York.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Slade, R.W., Moritz, C., Heideman, A., 1994. Multiple nuclear-gene phylogenies—application to pinnipeds and comparison with a mitochondrial DNA gene phylogeny. Mol. Biol. Evol. 11, 341–356.

Sorenson, M.D., 1999. TreeRot, version 2. Boston University.

Springer, M.S., Amrine, H.M., Burk, A., Stanhope, M.J., 1999. Additional support for Afrotheria and Paenungulata, the performance of mitochondrial versus nuclear genes, and the impact of data partitions with heterogeneous base composition. Syst. Biol. 48, 65–75.

Springer, M.S., Debry, R.W., Douady, C., Amrine, H.M., Madsen, O., De Jong, W.W., Stanhope, M.J., 2001. Mitochondrial versus nuclear gene sequences in deep-level mammalian phylogeny reconstruction. Mol. Biol. Evol. 18, 132–143.

Stewart, C.-B., Schilling, J.W., Wilson, A.C., 1987. Adaptive evolution in the stomach lysozymes of foregut fermenters. Nature 330, 401–404.

Suzuki, Y., Glazko, G.V., Nei, M., 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. USA 99, 16138–16143.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). version 4. Sinauer Associates.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.

Tedford, R.H., Taylor, B.E., Wang, X., 1995. Phylogeny of the Caninae (Carnivora: Canidae): The living taxa. American Museum Novitates 0, 1–37.

Van Valkenburgh, B., 1990. Carnivore dental adaptations and diet: a study of trophic diversity in guilds. In: Gittleman, J.L. (Ed.), Carnivore Behavior, Ecology and Evolution. Cornell University Press, Ithaca, New York, pp. 410–436.

Van Valkenburgh, B., 1991. Iterative evolution of hypercarnivory in canids (Mammalia: Carnivora): Evolutionary interactions among sympatric predators. Paleobiology 17, 340–362.

Venta, P.J., Brouillette, J.A., Yuzbasiyan-Gurkan, V., Brewer, G.J., 1996. Gene-specific universal mammalian sequence-tagged sites: application to the canine genome. Biochem. Genet. 34, 321–341.

Wayne, R.K., O'Brien, S.J., 1987. Allozyme divergence within the Canidae. Syst. Zool. 36, 339–355.

Wayne, R.K., Benveniste, R.E., Janczewski, D.N., O'Brien, S.J., 1990. Molecular and biochemical evolution of the carnivora. In: Gittleman, J.L. (Ed.), Carnivore Behavior, Ecology and Evolution. Cornell University Press, Ithaca, New York, pp. 465–494.

Wayne, R.K., Geffen, E., Girman, D.J., Koeppfli, K.P., Lau, L.M., Marshall, C.R., 1997. Molecular systematics of the Canidae. Syst. Biol. 46, 622–653.

Wayne, R.K., Nash, W.G., O'Brien, S.J., 1987a. Chromosomal evolution of the Canidae. I. Species with high diploid numbers. Cytogenet. Cell. Genet. 44, 123–133.

Wayne, R.K., Nash, W.G., O'Brien, S.J., 1987b. Chromosomal evolution of the Canidae. II. Species with low diploid numbers. Cytogenet. Cell. Genet. 44, 134–141.

Wayne, R.K., Van Valkenburgh, B., O'Brien, S.J., 1991. Rates of molecular evolution on carnivores and primates. Mol. Biol. Evol. 8, 297–319.

Zharkikh, A., Li, W.-H., 1995. Estimation of confidence in phylogeny: the complete-and-partial bootstrap technique. Mol. Phylogenet. Evol. 4, 44–63.

Zrzavy, J., Ricankova, V., 2004. Phylogeny of recent Canidae (Mammalia, Carnivora): relative reliability and utility of morphological and molecular datasets. Zool. Scr. 33, 311–333.