Mitochondrial DNA coding region sequences support the phylogenetic distinction of two Indian wolf species

ARTICLE *in* JOURNAL OF ZOOLOGICAL SYSTEMATICS AND EVOLUTIONARY RESEARCH · APRIL 2007

Impact Factor: 1.91 · DOI: 10.1111/j.1439-0469.2006.00400.x

Accepted on 24 September 2006
J Zool Syst Evol Res doi: 10.1111/j.1439-0469.2006.00400.x

[1]*Centre for Cellular and Molecular Biology, Hyderabad, India;* [2]*Estonian Biocentre and Department of Evolutionary Biology, Tartu University, Tartu, Estonia*

Mitochondrial DNA coding region sequences support the phylogenetic distinction of two Indian wolf species

R. K. Aggarwal[1], T. Kivisild[2], J. Ramadevi[1] and L. Singh[1]

Abstract

Two small endangered populations of Indian wolves were recently shown to be distant from other wolf and dog mtDNA lineages characterized so far. None of the inner branches in the tree of canid species based on partial hypervariable D-loop sequences were, however, statistically supported by the data raising the question whether the two Indian wolf lineages represent two new species, occupying an intermediate position between *Canis latrans* and *C. lupus* or have diverged from the sub-species of *C. lupus* due to isolation and drift. Here we report complete D-loop, cytochrome b, and 16S rRNA sequences data for 23 additional wolves from India analysed in the context of other canid species. Extended analyses of D-loop data and partial sequences of 16S rRNA showed highly reticulated pattern and were unable to resolve unambiguously the phylogenetic relationship of Indian wolves among other canid species. The phylogenetic reconstructions of cytochrome b sequences, however gave significant statistical support for the inner branches supporting genetic distinction of the two Indian wolf lineages within themselves as well as from all other wolves of the world, including individuals belonging to subspecies *C. lupus chanco* and *C. lupus pallipes* to which the two Indian wolf populations have been traditionally assigned. Their genetic differentiation relative to worldwide variation of wolves supports the suggestion to treat them as separate wolf species, *C. himalayensis* and *C. indica*.

Key words: Canid evolution – mitochondrial DNA – Indian wolves – 16S rRNA – cytochrome b – D-loop

Introduction

The grey wolf (GW), *Canis lupus*, is believed to be the most widely distributed terrestrial mammal that ever lived on earth at one time point (Goldman 1944), originally inhabiting major parts of the Northern hemisphere (Fig. 1). From being omnipresent and abundant, it has assumed endangered species status in many countries (Mech 1970; Ellegren et al. 1996). During the last century, based on the variation in physical features, behavioural aspects and geographical distribution, up to 32 sub-species of GW have been described, many of which are now believed to be extinct (Hall 1981; Macdonald 2001). Ten extant subspecies can be characterized in Eurasia (Fig. 2), the status of many of these is, however, contested, e.g. subspecies *C. l. campestris* Dwigubski, 1804, *C. l. chanco* Hodgson, 1847 and *C. l. desertorum* Bogdanov, 1882 are suggested to be synonyms of subspecies *C. l. lupus* Linnaeus, 1758 (Bush 1993), making the GW taxonomy highly debatable pending validation and possible revisions in the light of accumulating genetic data. Presently, only six subspecies, viz. *C. l. lupus* Linnaeus, 1758, *C. l. campestris* Dwigubski, 1804, *C. l. albus* Kerr, 1792, *C. l. tundrarum* Miller, 1937, *C. l. lycaon* Schreber, 1775 and *C. l. nubilus* Schreber, 1775 are believed to survive globally, and only the first two of these in Asia (Macdonald 2001). In recent years, many molecular genetic studies have been carried out to understand the origin and evolution of Canidae family (Roy et al. 1994; Wayne et al. 1997 and cf. therein; Tsuda et al. 1997; Bininda-Emonds et al. 1999) and the population dynamics of isolated GW populations (Wayne and Jenks 1991; Gottelli et al. 1994; Ellegren et al. 1996). Similar studies for resolving the largely debatable GW taxonomy at the subspecies level are generally still lacking. However, a high-resolution genomic analysis of dog breeds showed the potential of molecular studies to capture the structure of genetic differentiation even if it is of very recent origin (Parker et al. 2004). In contrast, studies based on mitochondrial DNA (mtDNA) variation, employing only short stretches of the D-loop region, have revealed high level of homogeneity of all wolf and dog breeds on the continental or regional scales (Wayne et al. 1992; Vila et al. 1999). The majority of the wolf and dog D-loop sequences cluster together in five major closely related lineage clusters (Vila et al. 1997).

Little is known about the origin and evolutionary history of wolves in South Asia. Generally, as per the conventional taxonomy, it is believed that the small number of wolves found in present-day India represent local variants of two widely spread sub-species of *Canis lupus*. One of these wolf populations is found only in the upper Trans-Himalayan region of India across the two northernmost states of Himachal Pradesh and Jammu and Kashmir, which as per the available census numbers ~350 individuals (Fox and Chundawat 1995). This Himalayan wolf (HW) population is considered to be a part of the relatively better known Tibetan wolf, *C. l. chanco*, which is found throughout central Asia with its range extending into Tibet, China, Manchuria and Mongolia (Fox and Chundawat 1995) (Fig. 2). On the other hand, the second wolf population estimated to be numbered less than 1500 individuals (http://www.wolf.org/wolves/learn/basic/populations/fall99insert.asp), is found throughout the arid/semi-arid plains of peninsular India. These wolves, commonly called Indian GW, are believed to represent the second sub-species in India, i.e. *C. l. pallipes* that is also found in the middle Eastern countries like Iran and Israel (Shahi 1982). The wolf in India are accorded schedule '1' endangered species status under the Indian Wildlife Protection Act of 1972 and of CITES, and in the last few decades federal efforts were initiated for wolf protection and conservation by setting up wildlife preserves in the country and captive breeding programmes in few national zoological parks. Despite their endangered status, no studies were made to understand the population genetics of the Indian wolves and their relationship with wolves from other parts of the world. In

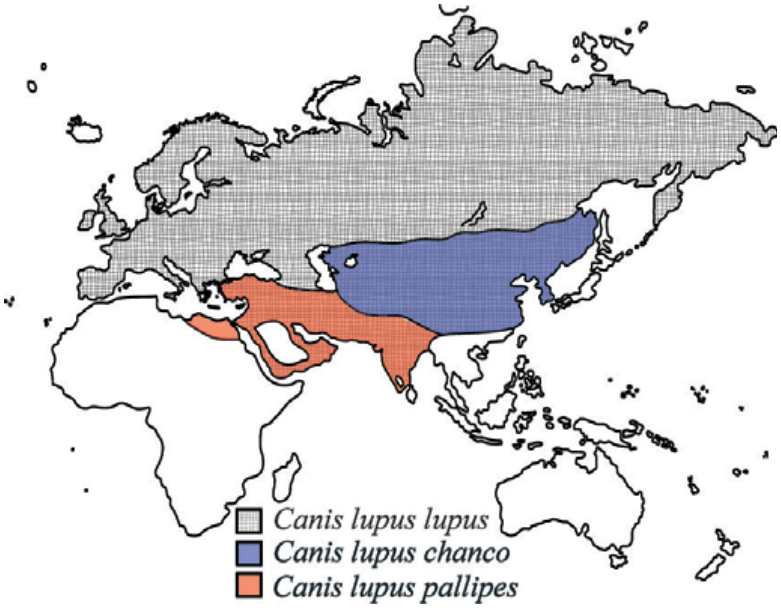

Fig. 1. Worldwide map showing the distribution of wolf sub-species, *C. lupus*, and *C. l. pallipes* and *C. l. chanco*

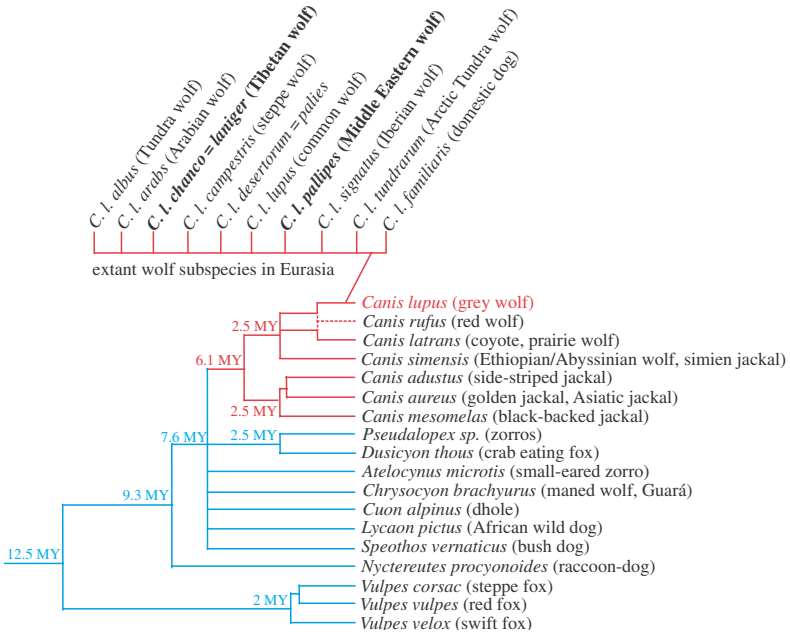

Fig. 2. Schematic tree of canid phylogeny. Divergence times, shown in millions of years, are given as in Bininda-Emonds et al. (1999)

studies performed elsewhere, only six samples were referred as Indian wolf (Tsuda et al. 1997), five of which originated from Afghanistan, and the source of the only one sample used in studies by Vila et al. (1997, 1999) is indicated to be from India. Therefore, it becomes imperative to understand the genetic make-up of wolf populations in India and their exact taxonomic status with other worldwide wolf populations to have a long-term and effective plan in place for their continual survival and conservation.

Recently, Sharma et al. (2004) compared the variation of 440 bp of mtDNA D-loop in the two Indian wolf populations within the context of worldwide variation among wolves. This study showed that both HW and GW clustered separately from other extant wolf and dog sequences, consistent with the earlier suggestions that both of them might represent independent wolf species – *Canis laniger* and *Canis pallipes* (cf. Sharma et al. 2004). However, the exact phylogenetic position of the two Indian wolf populations remained unsettled because statistical support was sufficient only for the external, species-specific branches of the phylogenetic tree. The clade including all wolf and dog sequences except HW and GW, for example, had received only 50% of bootstrap support from maximum likelihood analysis suggesting that in another half of the trees either HW or GW lineages were intermingled with those common to wolf and dog. A likely reason for the low phylogenetic resolution is that a short stretch of the hypervariable

D-loop region that was used in this study is an inadequate tool to solve splits at species level because of its shallow saturation curve. Furthermore, the neighbour-joining tree of D-loop sequences presented in the study by Sharma et al. (2004) included canid species whose sequence data were not reported in the study nor could be retrieved from the GenBank for the reported bp range (for example, data for maned wolf were completely absent and for simian jackal only partial sequences could be retrieved). Neither did the study by Sharma et al. (2004) explicitly report the analyses or data for cytochrome b sequences that were used to draw conclusions about the coalescent times of the splits between HW, GW and wolf lineages. To test whether the Indian GW and HW populations are indeed phylogenetically distant from other *C. lupus* subspecies we have analysed three regions of mtDNA, including cytochrome b, 16S RNA and an extended analyses of the D-loop region.

Materials and Methods

Study sample

The HW population is estimated to be of ~350 animals in the wild, spread over an area of 70 000 km^2 encompassing the Trans-Himalayan region of Jammu and Kashmir and Himachal Pradesh (Fox and Chundawat 1995). Apart from these animals in the wild, there are a total of only 21 live animals in four of the Zoological Parks of India, most of which are captive bred from few wild-caught animals from the Trans-Himalayan region (Anonymous 2002). In the present study we analysed 18 of these animals available in the Padmaja Naidu Himalayan Zoological Park (PNHZP), Darjeeling, West Bengal. In addition, we analysed representative samples of related canids (Table 1) available in the Nehru Zoological Park (NZP), Hyderabad, Andhra Pradesh, and Sakkarbaug Zoo, Junagarh, Gujarat. These included GW, wild dog and Indian jackal, originally caught from wild from three widely separated geographical regions of middle India.

The Institutional Ethics Committee as well as the Central Zoo Authority of India that regulates the research on endangered animals has approved the study.

DNA isolation

From each animal, ~10 ml blood was drawn using vacutainers having EDTA as anti-coagulant, transported to the laboratory at room temperature, and stored at 4°C till further use. The total genomic DNA was isolated following the lysis-freeze fracture method (Aggarwal et al. 1994).

DNA amplification and sequencing for nucleotide diversity/DNA polymorphisms

All samples were analysed for nucleotide diversity across three domains of the mitochondrial genome. These included part of cytochrome b gene (also complete gene for few representative samples), complete hypervariable d-loop control region and part of 16S rDNA. In each case, the target sequences were PCR amplified and sequenced using an automated DNA sequencer (ABI-Prism3700; Applied Biosystems, Foster City, California, USA) as described earlier (Aggarwal et al. 2003; Shanker et al. 2004).

Cytochrome-b gene

A ~332-bp conserved region of cytochrome-b gene towards the amino terminus of the protein was amplified and directly sequenced using primers L14841 and H15149 described by Kocher et al. (1989) for all the samples. In addition, complete cytochrome-b gene sequence was determined for one sample each of HW, GW and Indian jackal. For the purpose, the target gene was amplified and sequenced using the conserved flanking primers L-14724, H-15915 and internal primers L-15162, L-15513 and H-15149 (Irwin et al. 1991).

D-loop control region

For each of the sample, the complete d-loop region (Fig. 3), encompassing 5′ hypervariable control region-1 (CR-I), repetitive region and 3′ relatively conserved control region-II (CR-II) was amplified and sequenced using the primers described by Tsuda et al. (1997). Initially the ~1300 bp region was amplified using primers DL-14 and DH-7, which was then sequenced using two additional internal primers DH-6 and DH-10.

Partial d-loop sequences were also amplified and sequenced using the primers DLH-16340 and Thr-L15926 described by Vila et al. (1999), which in addition to the part CR-1 region of D-loop also span the proline t-RNA^{Pro}, and phenylalanine t-RNA^{Phe} genes.

16S rDNA

Part of 16S rDNA was amplified and sequenced using the primers 16Sar-L and 16Sbr-H (Palumbi et al. 1991) for all the 27 samples analysed in the study.

Table 1. Details of samples used and DNA sequencing analysis done in the present study

Sample/species	*n*	Status	Origin	Source
Himalayan wolf (HW)	18	Captive bred from few wild-caught animals	Trans-Himalayan region of India	Padmaja Naidu Himalayan Zoological Park, Darjeeling (West Bengal)
Gray wolf (GW)	5[1]	Wild caught	Central peninsular India	Nehru Zoological Park, Hyderabad (Andhra Pradesh); Sakkarbaug Zoo, Junagarh (Gujarat)
Jackal (J), [*C. aureus*]	2	Wild caught	Central India	
Wild dog (WD) [*Cuon alpinus*]	2	Wild caught	Central India	
DNA sequencing analysis done				
Mitochondrial domain	Samples analysed (*n*)			
D-loop control region (~1300 or more bp)	HW (18); GW (5); J (2)			
Partial cytochrome-b gene (~350 bp)	HW (18); GW (3); J (2); WD (2)			
Complete cytochrome-b gene (1140 bp)	HW (2); GW (2); J (1)			
16S rRNA gene (560 bp)	HW (18); GW (5); J (2); WD (2)			

[1]Samples originated from three different states of peninsular India.

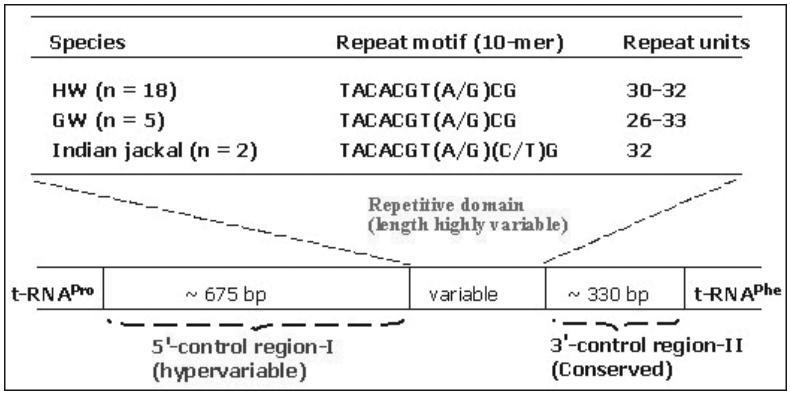

Fig. 3. Mitochondrial d-loop control region organization. Upper panel shows the sequence polymorphism seen in the repetitive domain of d-loop in canids from India (present study)

EMBL database sequences for comparative analysis

The GenBank database (http://www.ncbi.nlm.nih.gov) was searched for similar sequences using BLAST, and sequences described for wolf and other related canid species were retrieved for use as reference in phylogenetic comparisons. The details of the reference sequences, viz. EMBL accession number, source species, sequence length and location, and their original contributors are given in Supplementary Tables 1 and 2. The 90 D-loop sequences retrieved for comparative analysis represented the global wolf populations across all the continents. In contrast, the reference sequences for cytochrome-b and 16S rDNA comparisons were mainly from the related canid genera/species.

Phylogenetic analysis

The sequences obtained in the study were aligned with corresponding reference sequences from the EMBL database using the ClustalW program (http://www.ebi.ac.uk/clustalw/). The aligned sequences were manually checked for gaps and edited at the ends to avoid missing information in the compared taxa/entries. Diversity statistics of the compared sequences were obtained using Dna-SP ver. 3.51 (http://www.ub.es/dnasp/). The nucleotide diversity was used to derive genetic distance estimates as gamma-corrected Kimura 2-parameter (Kimura 1980), and also phylogenetic relationships using various analytical routines available in the software package PHYLIP 3.6 (http://evolution.genetics.washington.edu/phylip.html), Phylo_Win (http://pbil.univ-lyon1.fr/software/phylowin.html), and Tree-Puzzle (http://www.tree-puzzle.de/). The phylogenetic analysis using character-state sequence data was also carried out using the maximum-parsimony (MP) and maximum-likelihood (ML) approaches. Support for nodes found on the shortest tree was assessed by bootstrap analysis (Felsenstein 1985). Median-Joining networks were generated using NETWORK program (http://www.fluxus-engineering.com).

The HW-GW pair-wise distance values based on cytochrome-b sequence data were used to derive the estimate of evolutionary time separating the two, using the conservative sequence divergence rates reported for canid mtDNA divergence (Wayne et al. 1997).

Results and discussion

Approximately 2.0–2.2 kb mtDNA was sequenced encompassing its three domains, parts of cytochrome-b and 16S rRNA genes and complete D-loop control region (Fig. 3) for 23 wolves (18 HW, 5 GW) and four samples from related canids (Table 1; GenBank: AY291406–AY291430, AY289946–AY289997). In case of dhole (*Cuon alpinus*), only cytochrome-b and 16S rRNA gene regions were sequenced. In addition, complete cytochrome-b gene was also sequenced from one sample each of HW, GW and jackal (GenBank: AY291431–AY291433). Reference sequences mainly of different wolf populations/subspecies and related canid species were retrieved from the EMBL database (Tables S1 and S2). The sequence polymorphism data analysed are summarized in Table 2 and Tables S3–S8.

D-loop control region

Analysis of the complete D-loop region (~1300 bp) revealed high level of homogeneity within HW and GW samples. When excluding the repetitive domain, all 18 HW individuals representing 85% of the total individuals available across four zoological parks in India shared the same haplotype, suggesting that they derive from a single female ancestor, possibly only a few generations ago. This finding is consistent with the data of Sharma et al. (2004) wherein all the eight individuals defined only one haplotype HW-C over the sequenced 440 bp fragment that matched with the corresponding domain of our HW haplotype. On the other hand, the nine museum samples reported by Sharma et al. (2004) were derived from this haplotype by 1.3 mutations on average corresponding to an age of the last common female ancestor of the Himalayan wolves only about 30,000 years ago when using the mutation rate of 9.2% per million years (Sharma et al. 2004).

Similarly the five Indian grey wolves analysed by us revealed the presence of only two closely related haplotypes. One of these haplotypes, accounting for four individuals matched again within the stretch of 440 bp with haplotype IW-B that was common throughout the Indian GW samples analysed by Sharma et al. (2004). The fifth Indian GW in our study represented a new haplotype, derived from IW-B by a single transition at np 144 of the D-loop sequence.

Phylogenetic analysis of the D-loop region revealed a high level of polymorphism and homplasy within and between the canid species. Using a stretch of 326 bp that has been commonly sequenced for multiple canid species revealed highly reticulate pattern between different species (Fig. 4a). Even though both Indian HW and GW lineages clustered separately from other common lineages within *C. lupus* species the exact branching order remained undetermined. Extended maximum likelihood analysis of 673 bp of the D-loop involving *C. aureus*, *C. latrans*, HW, GW and different sub-species of *C. lupus* failed to give either statistical support for a monophyletic clade of the species (Fig. 4b). A weakly supported (53%) clade that included all extant *C. lupus*, GW and HW lineages also hosted both Indian jackals that were submitted to the tree analysis. Later, in the branching order, a branch (with 55% support) holding all *C. lupus* lineages except HW and GW adds support to the idea that *C. l. chanco* and

Table 2. Summary statistics of DNA polymorphisms analysed in the study

Sequence parameters	Mitochondrial domain used for taxonomic comparisons							
	Complete Cyto-B gene	Partial Cyto-B gene	16S rRNA gene	D-loop (CR-I) 233 bp	D-loop (CR-I) 545 bp	D-loop (CR-I) 651 bp	D-loop (CR-I) 675 bp	D-loop region-II
Sequences compared	5	35	27	80	50	33	24	11
Sequence size compared (bp)	1140	332	554	233	545	651	675	294
Sequence size without gaps/'N' (bp)	1132	280	554	210	505	608	656	252
Polymorphic nucleotides	232	77	32	35	72	73	75	22
Singleton polymorphic nucleotides	178	34	–	14	33	39	44	16
Informative nucleotides	54	43	32	21	39	34	31	6
Base composition (average, %)								
A	28.9	28.2	26.7	30.5	26.9	27.3	27.2	33.6
C	27.6	23.0	20.3	24.0	27.1	26.5	25.6	21.4
G	14.2	16.6	21.7	12.3	14.9	15.3	16.3	11.3
T	29.4	32.2	31.3	33.2	31.1	30.9	30.9	33.7
Observed transition/transversion ratios (over all sites/sequences/pair-wise comparisons)								
Average	4.15	8.70	11.65	19.57	9.89	7.60	6.96	14.20
Minimum	2.80	0.00	5.70	0.00	0.00	0.00	0.00	0.00
Maximum	18.75	26.00	11.00	15.00	23.00	24.00	25.00	12.00
Observed substitution statistics								
Over all sites/sequences/pair-wise comparisons								
No. haplotypes	5	18	6	36	39	28		10
No. average substitutions	106.0 ± 66.0	12.35 ± 9.57	7.21 ± 0.93	4.97 ± 0.15	9.65 ± 0.83	11.15 ± 1.7		5.53 ± 1.5
Nucleotide diversity (%)	9.39 ± 4.87	4.37 ± 1.15	1.30 ± 0.50	2.37 ± 0.04	1.91 ± 0.04	1.83 ± 0.03		2.19 ± 0.03
Ks (%)	57.2 ± 48.6	19.2 ± 16.5						
Ka (%)	1.49 ± 0.79	0.77 ± 0.80						
HW versus GW and wolf/dog populations from elsewhere								
No. average substitutions				14.1 ± 1.6	20.7 ± 1.9	21.3 ± 1.9		7.1 ± 1.6
Minimum–maximum substitutions				11–16	17–24	18–25		5–9
Median (mode) substitutions				14 (16)	20 (19)	21 (21)		7 (7)
GW versus HW and wolf/dog populations from elsewhere								
No. average substitutions				13.4 ± 1.3	17.3 ± 1.6	18.0 ± 2.0		5.1 ± 1.6
Minimum–maximum substitutions				10–17	14–21	14–21		4–9
Median (mode) substitutions				13 (13)	17 (17)	18 (19)		5 (4)
Between different wolf/dog populations excluding HW/GW from India								
No. average substitutions				4.5 ± 2.23	7.4 ± 3.2	7.6 ± 3.4		3.1 ± 1.3
Minimum–maximum substitutions				0–10	0–15	0–15		0–5
Median (mode) substitutions				4 (4)	8 (7)	8 (7)		3 (3)

Ks: synonymous substitutions; Ka: non-synonymous substitution.

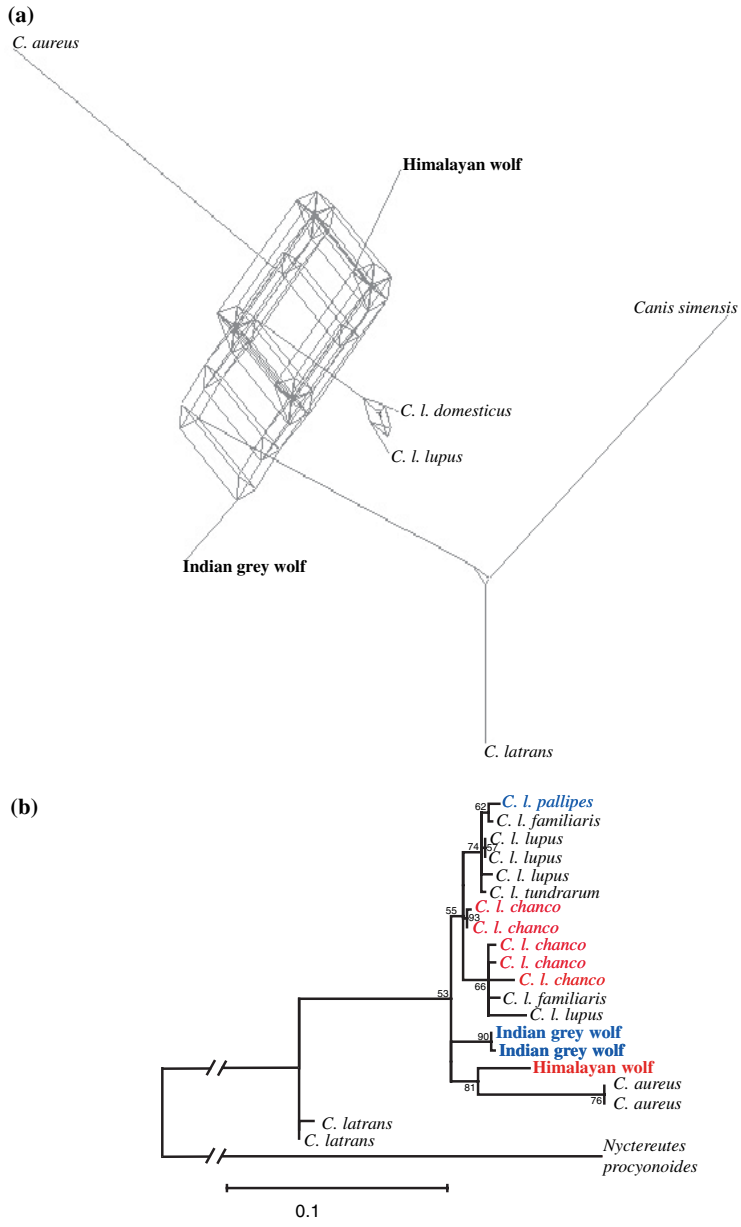

Fig. 4. Phylogenetic reconstruction showing genetic relationship of HW and GW with other canids from India and elsewhere based on 326 bp mtDNA corresponding to D-loop CR-1 + tRNA($^{Pro/Phy}$) sequences. (a) Median-joining network ($\varepsilon = 2$). (b) A quartet-puzzling tree. Support for the internal branches is shown in percentage. Branch lengths are computed using Kimura 2 parameter model with gamma distributed rate heterogeneity with parameter alpha estimated from the data

C. l. pallipes cannot be accepted as monophyletic wolf sub-species, at least as far as mtDNA is concerned. But this support, as in case with the analyses reported by Sharma et al. (2004), is statistically not robust. Furthermore, the clock-like tree was rejected in a likelihood ratio test at the 5% significance level.

Importantly, however, the comparative analysis of D-loop polymorphism across different lengths of its CR-1 region and varying numbers of reference wolf sequences/haplotypes (viz. 545 bp/44 sequences; 651 bp/27 sequences; and 675 bp/19 sequences) showed Indian wolves (HW and GW) to be genetically most distant and forming a completely separate clade without any overlap with the haplotypes seen in the other wolves of the world (Fig. 4, Figs S1–S3). Moreover, the HW and GW clade was found always to be basal to the other major clade comprising all other wolf haplotypes and closest to the jackal, one of the closest ancestral canid species, suggesting them to be the derivatives of a more ancient independent wolf radiation.

Significantly, both HW and GW were found to carry D-loop polymorphism that was not seen in both of their corresponding purported sub-species of *C. lupus*, i.e. *chanco* and *pallipes* respectively. Indian HW analysed in the study, which has been referred in published literature as Tibetan/Chinese wolf *C. lupus chanco*, differs from the latter sub-species by 24 nucleotides over 651 bp of the CR-I region (Table S7). These data conclusively show that HW found in India is genetically distinct from the more abundant and widely distributed Tibetan wolf *chanco*. Similarly, the GW samples analysed in the study neither carried the nucleotide polymorphism found diagnostic of *pallipes* from Afghanistan (Tsuda et al. 1997), nor showed any similarity with the wolf

haplotype 'w12' reported for the only *pallipes* sample from India by Vila et al. (1997, 1999). Moreover, *C. lupus pallipes* reported in these studies did not have any of the GW-specific polymorphisms. These findings thus suggest that the GW from India is not the *pallipes* found in Middle East and Central Asia. The apparent discrepancy with respect to the only sample from India analysed earlier (Vila et al. 1999) having a very different haplotype (w12) can be explained by the possibility of wolf–female dog hybridization, a phenomenon that has been observed and reported earlier. Incidentally, this reported 'w12' haplotype of the Indian wolf was also closest to those of the dogs. It may be noted that the GW samples used in the present study, though few in numbers ($n = 5$), were unrelated wild-caught animals from geographically different parts of central peninsular India.

The repetitive domain of the d-loop control region was found to be highly variable for all the samples of HW, GW and Indian jackal sequenced in the study. The variation was seen both in the number of repeat motifs, as well as due to extensive intra-individual heteroplasmy for specific nucleotides of the repeat motifs (Fig. 3; AY289973–AY289997). The 10-mer repeat motifs of HW and GW were similar to what had been reported earlier (Tsuda et al. 1997) for dogs and wolf, but unlike their results, varied in number of tandem units (30–32 for HW and 26–33 for GW). In comparison with HW and GW, the repeat motif for Indian jackal, though of same length and sequence, had two variable nucleotides in its motif.

Cytochrome-b polymorphism

Sequencing of the 332 bp region of cytochrome-b revealed four transitions (sequence divergence = 1.24%) between HW and GW. In comparison, HW differed for eight transitions (3 A > G, 1 G > A, 4 T > C; sequence divergence = 2.48%) and GW for four transitions (2 A > G, 2 T > C) from the holarctic wolf (Table S3). The data thus show the three wolf types to be genetically different from each other with HW to be most divergent. The data from complete cytochrome-b gene (Table S4) also revealed similar divergence between HW and GW, with HW being closer to jackal. Using a stretch of 696 bp for which comparable data were available for *C. latrans* (Wayne et al. 1997), HW and GW, we estimated the K-2 corrected divergence rate at the third codon position of the cytochrome b gene as 16.5% (SE 3.1%) as the divergence of the coyote and wolf/dog lineages. Assuming a minimum age for this split as 1 Myr (million years), the separation times of HW and GW lineages from those of common wolves and dogs are estimated as 630 (SE 135) and 270 (SE 80) KYA (thousand years ago).

Maximum likelihood tree based on 709 bp of cytochrome b sequences clustered HW and GW in between the wolf/dog clade and other *Canis* species (Fig. 5). The distinction of the clade including two wolf and two dog sequences received a statistical support of 92%. The clock-like tree could not be rejected ($p = 0.905$). The support of HW outgrouping the cluster including GW and wolves/dogs also received a relatively high statistical support (90%). Similarly, high support was obtained for clustering of HW, GW and wolves (87%) and the clade including all species in *Canis* family (91%). Likewise, in case of D-loop sequences, *Canis aureus* did not appear more distant from wolves than *C. latrans* and *C. simensis* (viz. Fig. 2), even though the support for this clade excluding *C. mesomelas* was rather weak (52%).

16S rDNA polymorphism

Relatively low variation was observed among the samples over the ~550 bp of the 16S rDNA sequenced in the study. Only 32 nucleotides were found to be polymorphic and informative.

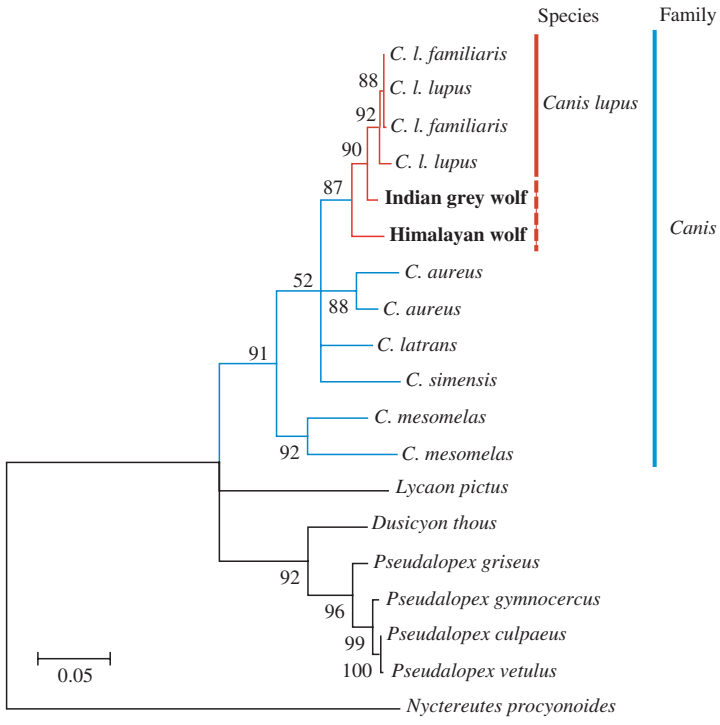

Fig. 5. Quartet puzzling tree based on canid 708 bp of canid cytochrome b sequences

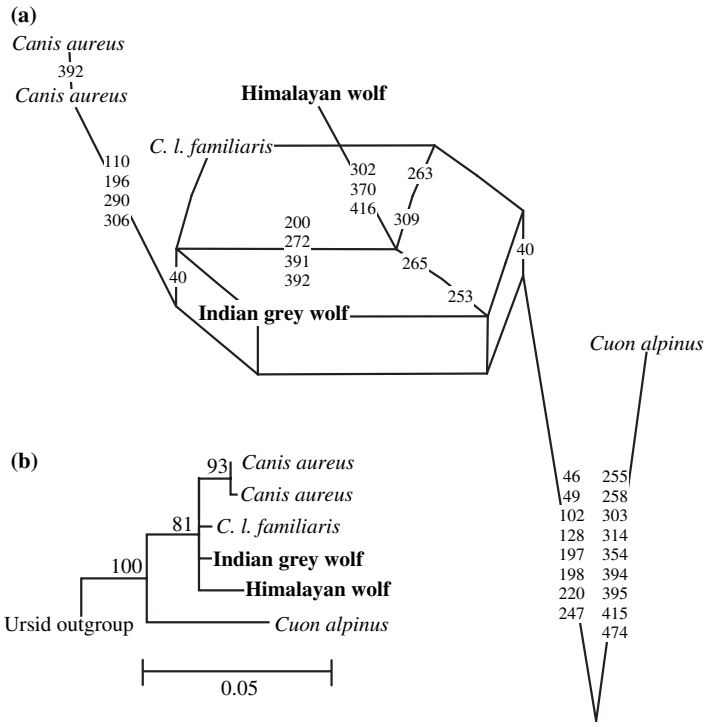

Fig. 6. The position of Himalayan and Indian grey wolves in the canid phylogeny based on 554 bp of their 16S rRNA sequences. (a) Median joining network. The variable positions are shown on branches connecting the sequences and correspond to the numbering from 1 to 554 in the stretch of nucleotide positions 1938 to 2491 in the complete sequence of the domestic dog (Kim et al. 1998). (b) Quartet puzzling tree. Support for the internal branches is shown in percentage. Branch lengths are computed using Kimura 2 parameter model with gamma distributed rate heterogeneity with parameter alpha estimated from the data set as 0.12 (SE 0.02). Clock-like tree could not be rejected in the likelihood ratio test on a critical significance level 41.62%. Mutation rate 3.8×10^{-9} substitutions per site per year was estimated by assuming 7.6 MYA split between *Canis* and *Cuon* genera (see Fig. 2). Applying this rate to the observed distances within *Canis* species corrected for gamma distribution and transition/transversion ratio [estimated from the canid data as 8.5 (SE 2)] yields a date of 2.2 (SE 0.7) Myr for the split between *Canis aureus* and the *Canis lupus* sequences, 2.4 (SE 0.9) Myr for the split between Himalayan wolf and *C. l. familiaris*, and 700 (SE 430) KYA (thousand years ago) for the split between Indian grey wolf and *C. l. familiaris*

EMBL database search did not reveal 16S sequences from wolf or coyote populations for reference, and thus the phylogenetic analysis was limited to HW, GW, dhole, dog and jackal sequences, using ursids as an outgroup (Fig. 6). Consistent with D-loop and cytochrome b data, HW appeared as more distant from the domestic dog (nine substitutions, K-2 corrected distance 0.0165, SE 0.0054) than the GW (four substitutional differences, $d = 0.0073$, SE 0.0037). However, the jackal sequences showed lower distance from dog (7.5 substitutions, $d = 0.0137$, SE 0.005) than HW. The dhole and dog sequences showed 24 differences ($d = 0.0451$, SE 0.0093). The clock-like tree could not be rejected ($p = 0.187$). Assuming 7.6 MYA (million years ago) split between dogs and dholes, the splits between dog and HW, GW, and jackal would be dated as 2.8 (SE 0.9), 1.2 (SE 0.6) and 2.3 (SE 0.8) MYA respectively.

Possible origin of HW and GW of India

Our data, consistent with earlier observations (Aggarwal et al. 2003; Sharma et al. 2004), strongly suggest HW and GW to represent the two oldest wolf lineages that are distinct from the third visibly more recent lineage that gave rise to all other global wolf populations. In all the phylogenetic reconstructions, both the lineages specific to Indian wolves appear as genetically distinct completely separate clade(s) with no overlap with any of the haplotypes representing wolves from the rest of the world. These results suggest that Indian wolves represent a relic ancestral lineage of wolf that evolved for a long time in isolation from the evolutionary lineage that is represented by all other wolves from outside India. In the light of our results and current understanding of canid systematics (Vila et al. 1997; Wayne et al. 1997), we hypothesize that the wolves of India represent a pre-wolf, post-jackal, ancestral canid radiation that migrated to India sometime in early Pleistocene about 1–2 MYA and since then evolved independently without any interaction with the evolving wolf lineage(s) in other parts of the world. This in turn suggests that the Indian subcontinent had been one major centre of wolf/canid evolution, an inference that is in agreement with the school of taxonomists that ascribes an Asian origin to the wolf-like canids (Bush 1993).

The data presented here demonstrate Indian wolves to be genetically unique, supporting the proposal to assign them to new species/sub-species status, *C. himalayaensis* and *C. indica*, as suggested previously (Aggarwal et al. 2003). These can be briefly described as follows.

Canis himalayensis resembles in its outer appearance (morphological features), social/reproductive behaviour to *Canis lupus chanco/laninger* (an Eurasian wolf found in China, Mongolia and South-East Asia), but is a genetically

distinct and ancestral species. HW is found in the Trans-Himalayan terrains of India across the states of Himachal Pradesh, Jammu & Kashmir. Similarly, *Canis indica* resembles *Canis lupus pallipes* (an Eurasian wolf found all across Middle East Asian countries) in its outer appearance (morphological features), social/reproductive behaviour, but the Indian GW (*C. indica*) is relatively smaller in size and genetically distinct from *C. l. pallipes*. It is found in the arid, semi-arid peninsular plains of Central India.

Conclusions

The present study aimed at understanding the genetic/taxonomic status of Indian GW and HW relative to other wolf-like canids from the rest of the world based on diversity in their mitochondrial DNA. The study thus suggest that: (1) HW is distinct from the purported similar wolf from China, *C. lupus chanco*, and represents the most ancient wolf lineage ever recorded and analysed using molecular means; (2) similarly, GW is distinct from the purported similar wolves from South-East Asia, *C. lupus pallipes*, and probably also represents the other ancient wolf lineage from India; (3) the Indian subcontinent may have been one of the ancient centres of wolf origin and diversification.

Acknowledgements

We thank the Central Zoo Authority, India, Chief Conservators of Forests (West Bengal, Andhra Pradesh and Gujarat), Ministry of Environment and Forests, Government of India, for necessary support and permissions; Mr M. Banerjee, ex-Curator NZP, Hyderabad, Dr B.R. Sharma, Curator PNHZP, Darjeeling, Curator, Sakkarbagh Zoo, Junagarh, and their teams of veterinary doctors and animal keepers, for helping with the collection of samples and Mehar Sultana for primer synthesis. The study was carried out under the Indian Government agency (CZA, DBT, CSIR)-supported mega project LACONES on conservation of endangered animals of India. TK was supported by an Estonian Science Foundation Grant (5574).

Résumé

Des séquences de la région codante de l'ADNmt supportent la distinction phylogénétique de deux espéces de loops indiennes

Deux petites populations en danger de loups indiens ont été démontrées récemment comme étant distantes des autres lignées mitochondriales de loups et de chiens caractérisées jusqu'à maintenant. Aucunes des branches internes de l'arbre des canidés basé sur le séquençage d'une partie de la boucle hypervariable D n'étaient cependant statistiquement soutenues par les données, soulevant la question à savoir si les deux lignées de loups indiens représentent deux nouvelles espèces, occupant une position intermédiaire entre *C. latrans* et *C. lupus*, ou bien ont divergé de la sous-espèce *C. lupus* à cause d'isolement et de dérive génétique. Nous rapportons ici les séquences de la boucle-D en entier, du cytochrome b et ARNr 16S de 23 loups additionnels d'Inde analysées dans le contexte d'autres espèces de canidés. Des analyses extensives des données de la boucle-D et des séquences partielles de l'ARNr 16S révèlent un modèle hautement réticulé et étaient incapables de résoudre sans ambiguïté la relation phylogénétique des loups indiens parmi les autres espèces de canidés. La reconstruction phylogénétique des séquences du cytochrome b cependant donne un soutien statistique significatif pour les branches internes, soutenant une distinction génétique des deux lignées de loups indiens entre-elles ainsi que des autres loups du monde, y compris les individus appartenant aux sous-espèces *C. lupus chanco* et *C. lupus pallipes* où les deux populations de loups indiens ont été traditionnellement assignées. Leur différentiation génétique relative aux autres loups mondiaux supporte la suggestion de les traiter comme des espèces de loups séparées, *C. himlayensis* et *C. indica*.

References

Aggarwal RK, Majumdar KC, Lang JW, Singh L (1994) Generic affinities among crocodilians as revealed by DNA fingerprinting with a Bkm-derived probe. Proc Natl Acad Sci U S A **91:**10601–10605.

Aggarwal RK, Ramadevi J, Singh L (2003) Ancient origin and evolution of the Indian wolf: evidence from mitochondrial DNA typing of wolves from Trans-Himalayan region and peninsular India. Genome Biol **4:**P6.

Anonymous (2002) Inventory of Animals in Indian Zoos 2000–2001. Central Zoo Authority, New Delhi, India, 314 pp.

Bininda-Emonds OR, Gittleman JL, Purvis A (1999) Building large trees by combining phylogenetic information: a complete phylogeny of the extant Carnivora (Mammalia). Biol Rev Camb Philos Soc **74:**143–175.

Bush RHE (1993) The Wolf Almanac. The Lyons Press, New York, 241 pp.

Ellegren H, Savolainen P, Rosen B (1996) The genetical history of an isolated population of the endangered grey wolf *Canis lupus*: a study of nuclear and mitochondrial polymorphisms. Philos Trans R Soc Lond B Biol Sci **351:**1661–1669.

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

Fox JL, Chundawat RS (1995) Wolves in the Trans-Himalayan region of India: the continued survival of a low density population. In: Carbyn LN, Fritts SH, Seip DR (eds), Ecology and conservation of wolves in a changing world. Proceedings of the second North American Symposium on wolves (Edmonton, Alberta, Canada, 25–27 August 1992) Canadian Circumpolar Institute, University of Alberta, Canada, pp. 95–104.

Goldman EA (1944) The wolves of the North America, Part-II. In: Young SP, Goldman EA (eds), Classification of Wolves. New York/American Wildlife Institute, Washington, D.C., pp 389–636.

Gottelli D, Sillero-Zubiri C, Applebaum GD, Roy MS, Girman DJ, Garcia-Moreno J, Ostrander EA, Wayne RK (1994) Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. Mol Ecol **3:**301–312.

Hall ERE (1981) The Mammals of North America. John Wiley & Sons, New York, 1181 pp.

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of mammals. J Mol Evol **32:**128–144.

Kim KS, Lee SE, Jeong HW, Ha JH (1998) The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. Mol Phylogenet Evol **10:**210–220.

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

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc Natl Acad Sci U S A **86:**6196–6200.

Macdonald DE (2001) The New Encyclopedia of Mammals. Oxford University Press, Oxford, pp 42–43.

Mech LD (ed.) (1970) The wolf: the ecology and behaviour of an endangered species. The Natural History Press, New York.

Palumbi SR, Martin A, Romano S, McMillan WO, Stice L, Grabowski G (1991) The Simple Fool's Guide to PCR, version 2. Zoology Department, University of Hawaii, Honolulu, 46 pp.

Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, Malek TB, Johnson GS, DeFrance HB, Ostrander EA, Kruglyak L (2004) Genetic structure of the purebred domestic dog. Science **304:**1160–1164.

Roy MS, Geffen E, Smith D, Ostrander EA, Wayne RK (1994) Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. Mol Biol Evol **11:**553–570.

Shahi PS (1982) Status of grey wolf (*Canis lupus pallipes*) in India: a preliminary survey. J Bombay Nat His Soc **79:**493–502.

Shanker K, Ramadevi J, Chaudhury BC, Singh L, Aggarwal RK (2004) Phylogeography of olive ridley turtles (*Lepidochelys olivacea*) on the east coast of India: implications for conservation theory. Mol Ecol **13:**1899–1909.

Sharma DK, Maldonado JE, Jhala YV, Fleischer RC (2004) Ancient wolf lineages in India. Proc R Soc Lond B Biol Sci **271**(Suppl. 3)**:**S1–S4.

Tsuda K, Kikkawa Y, Yonekawa H, Tanabe Y (1997) Extensive interbreeding occurred among multiple matriarchal ancestors during the domestication of dogs: evidence from inter- and intraspecies polymorphisms in the D-loop region of mitochondrial DNA between dogs and wolves. Genes Genet Syst **72:**229–238.

Vila C, Savolainen P, Maldonado JE, Amorim IR, Rice JE, Honeycutt RL, Crandall KA, Lundeberg J, Wayne RK (1997) Multiple and ancient origins of the domestic dog. Science **276:**1687–1689.

Vila C, Amorim IR, Leonard JA, Posada D, Castroviejo J, Petrucci-Fonseca F, Crandall KA, Ellegren H, Wayne RK (1999) Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. Mol Ecol **8:**2089–2103.

Wayne RK, Jenks SM (1991) Mitochondrial DNA analysis implying extensive hybridization of the endangered red wolf *Canis rufus*. Nature **351:**565–568.

Wayne RK, Lehman N, Allard NW, Honeycutt RL (1992) Mitochondrial DNA variability of the grey wolf: genetic consequences of population decline and habitat fragmentation. Conserv Biol **6:**559–569.

Wayne RK, Geffen E, Girman DJ, Koepfli KP, Lau LM, Marshall CR (1997) Molecular systematics of the Canidae. Syst Biol **46:**622–653.

Authors' addresses: Dr. Ramesh K. Aggarwal (correspondence), Scientist EII, Centre for Cellular and Molecular Biology, Uppal Road, Tarnaka, Hyderabad 500 007, India. E-mail rameshka@ccmb.res.in; Ms Jetty Ramadevi, Centre for Cellular and Molecular Biology, Uppal Road, Tarnaka, Hyderabad 500 007, India. E-mail: makara92000@yagoo.co.in; Dr Lalji Singh, Director, Centre for the Cellular and Molecular Biology, Uppal Road, Tarnaka, Hyderabad 500 007, India E-mail: lalji@ccmb.res.in; Dr. Toomas Kivisild: The Henry Wellcome Building, University of Cambridge, Fitzwilliam Street, Cambridge, CB2 1QH, UK. E-mail: tk331@cam.ac.uk

Supplementary Material

The following supplementary material is available for this article online:

Fig. S1. NJ phylogenetic tree showing relationship of HW and GW with 44 other wolf sequences/haplotypes from elsewhere and 2 dog entries. Polymorphism based on ~545 bp of d-loop control region-I. Indian jackal is used as outgroup. Note the Indian wolf haplotypes appearing as distinct clades closer to the jackal and basal to the one comprising all other global wolf haplotypes, indicating them to be representing genetically unique ancient lineages of wolf. Distance measure: Corrected K 2-P. Reference taxa are as per codes in Table S2. Numerical values near the nodes are bootstrap values (only values > 50% are shown)

Fig. S2. ML phylogenetic tree showing relationship of HW and GW with 27 other wolf sequences/haplotypes from elsewhere and 2 dog entries. Polymorphism based on ~651 bp of d-loop control region-I. Indian jackal is used as outgroup. Note the Indian wolf haplotypes appearing as distinct clades closer to the jackal and basal to the one comprising all other global wolf haplotypes, indicating them to be representing genetically unique ancient lineages of wolf. Distance measure: Corrected K 2-P. Reference taxa are as per codes in Table S2. Numerical values near the nodes are bootstrap values (only values > 50% are shown)

Fig. S3. ML phylogenetic tree showing relationship of HW and GW with 19 other wolf sequences/haplotypes from elsewhere and one dog entry. Polymorphism based on complete 675 bp of d-loop control region-I. Note the Indian wolf haplotypes appearing as distinct clades closer to the jackal and basal to the one comprising all other global wolf haplotypes, indicating them to be representing genetically unique ancient lineages of wolf. Indian jackal is used as outgroup. Distance measure: Corrected K 2-P. Reference taxa are as per codes in Table S2. Numerical values near the nodes are bootstrap values (only values > 50% are shown)

Table S1. Reference mitochondrial cytochrome-b and 16S RNA gene sequences retrieved from EMBL database for use in the phylogenetic analysis

Table S2. Reference mitochondrial D-loop control-region sequences/haplotypes retrieved from EMBL database for use in the phylogenetic analysis

Table S3. DNA substitutions observed in the partial (~332 bp) mitochondrial cytochrome-b gene sequences of Indian wolves and other reference canid species

Table S4. DNA substitutions observed in the complete mitochondrial cytochrome-b gene sequences of Indian wolves and three other canids from India and elsewhere

Table S5. DNA Substitutions and deletions (–) in 233 bp mtD-loop CR-I region of Indian wolves and other wolf populations around the globe (sample codes are as per Table S2)

Table S6. DNA substitutions and deletions (–) in 545 bp mtD-loop CR-I region of Indian wolves and other wolf sequences/haplotypes from around the globe. Sequence of Indian jackal, which is used as outgroup taxa, is also included. For sample codes, see Table S2

Table S7. DNA substitutions and deletions (–) in 651 bp mtD-loop CR-I region of Indian wolves and other wolf sequences/haplotypes from around the globe. Sequence of Indian jackal, which is used as outgroup taxa, is also included. For sample codes, see Table S2

Table S8. DNA substitutions and deletions (–) in 296 bp mtD-loop CR-II region of Indian wolves and other wolf sequences/haplotypes. Sequence of Indian jackal, which is used as outgroup taxa, is also included. For sample codes, see Table S2

This material is available as part of the online article from http://www.blackwell-synergy.com

