

This information has not been peer-reviewed. Responsibility for the findings rests solely with the author(s).

Deposited research article

Ancient origin and evolution of the Indian wolf: evidence from mitochondrial DNA typing of wolves from Trans-Himalayan region and Pennisular India

Ramesh K Aggarwal, J Ramadevi and Lalji Singh

Addresses: Centre for Cellular and Molecular Biology, Hyderabad 500 007, India

Correspondence: Ramesh L Aggarwal. Email: rameshka@ccmb.res.in

Posted: 29 April 2003

Received: 22 April 2003

Genome Biology 2003, 4:P6

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2003/4/6/P6>

This is the first version of this article to be made available publicly.

© 2003 BioMed Central Ltd

comment

reviews

reports

deposited research

referenced research

interactions

information



deposited research

AS A SERVICE TO THE RESEARCH COMMUNITY, GENOME **BIOLOGY** PROVIDES A 'PREPRINT' DEPOSITORY TO WHICH ANY ORIGINAL RESEARCH CAN BE SUBMITTED AND WHICH ALL INDIVIDUALS CAN ACCESS FREE OF CHARGE. ANY ARTICLE CAN BE SUBMITTED BY AUTHORS, WHO HAVE SOLE RESPONSIBILITY FOR THE ARTICLE'S CONTENT. THE ONLY SCREENING IS TO ENSURE RELEVANCE OF THE PREPRINT TO GENOME **BIOLOGY**'S SCOPE AND TO AVOID ABUSIVE, LIBELLOUS OR INDECENT ARTICLES. ARTICLES IN THIS SECTION OF THE JOURNAL HAVE **NOT** BEEN PEER-REVIEWED. EACH PREPRINT HAS A PERMANENT URL, BY WHICH IT CAN BE CITED. RESEARCH SUBMITTED TO THE PREPRINT DEPOSITORY MAY BE SIMULTANEOUSLY OR SUBSEQUENTLY SUBMITTED TO GENOME **BIOLOGY** OR ANY OTHER PUBLICATION FOR PEER REVIEW; THE ONLY REQUIREMENT IS AN EXPLICIT CITATION OF, AND LINK TO, THE PREPRINT IN ANY VERSION OF THE ARTICLE THAT IS EVENTUALLY PUBLISHED. IF POSSIBLE, GENOME **BIOLOGY** WILL PROVIDE A RECIPROCAL LINK FROM THE PREPRINT TO THE PUBLISHED ARTICLE.



Ancient origin and evolution of the Indian wolf: evidence from mitochondrial DNA typing of wolves from Trans-Himalayan region and Pennisular India

Ramesh K. Aggarwal, J. Ramadevi, Lalji Singh

Centre for Cellular and Molecular Biology, Hyderabad 500 007, India

Address for Correspondence:

Ramesh K. Aggarwal, Ph.D.

Scientist

Centre for Cellular & Molecular Biology

Uppal Road, Tarnaka

Hyderabad 500007

India

Ph: 00-91-40-27192643

Fax: 00-91-40-27160591

E-mail: rameshka@ccmb.res.in

Abstract

The two wolf types found in India are represented by two isolated populations and believed to be two sub-species of *Canis lupus*. One of these wolf, locally called Himalayan wolf (HW) or Tibetan wolf, is found only in the upper Trans-Himalayan region from Himachal Pradesh to Leh in Kashmir and is considered to be *C. lupus chanco*. The other relatively larger population is of Indian Gray wolf (GW) that is found in the peninsular India and considered to be *C. lupus pallipes*. Both these wolves are accorded endangered species status under the Indian Wildlife Protection Act. In 1998 for the first time in India, we initiated molecular characterization studies to understand their genetic structure and taxonomic status. Since then, we have analyzed the genetic variability in 18 of the total of 21 HW samples available in Zoological parks along with representative samples of GW, wild dogs and jackals. Our study of mitochondrial DNA diversity across three different taxonomically informative domains i.e., cytochrome-B gene, 16S rDNA and hypervariable d-loop control region revealed HW to be genetically distinct from the GW as well as from all other wolves of the world, including *C. lupus chanco* from China. Most importantly, d-loop haplotypic diversity revealed both HW and GW from India to be significantly diverse from other wolf populations globally and showed that these represent the most ancient lineages among them. Phylogenetic analysis revealed the Indian wolves as two independent lineages in a clade distinct and basal to the clade of all wolves from outside of India. Conservative estimate of evolutionary time-span suggests more than one million years of separation and independent evolution of HW and GW. We hypothesize that Indian wolves represent a post-jackal pre-wolf ancestral radiation that migrated to India about 1-2 mya and underwent independent evolution without contamination from other wolf like canids. The study thus, suggests that Indian subcontinent had been one major center of origin and diversification of the wolf and related canids. Further, the significant degree of genetic differentiation of HW from GW and of these two from other wolves, suggest the interesting possibility of them to be new wolf species/sub-species in evolution that may need to be described possibly as *C. himalayaensis* and *C. indica* (or as *C. lupus himalayaensis* and *C. lupus indica*), respectively. Thus for the first time, the study reveals new ancient wolf lineages in India and also highlights the need to revisit the origin, evolution and dispersion of wolf populations in Asia and elsewhere. Simultaneously, it increases the conservation importance of Indian wolves warranting urgent measures for their effective protection and management, especially of the small HW population that at present is not even recognized in the canid Red List.

Introduction

The gray wolf, *Canis lupus*, is believed to be the most widely distributed terrestrial mammal that ever lived on earth at one time point [1], originally inhabiting major parts of the Northern hemisphere. From being omnipresent and abundant, it has assumed endangered species status in many countries. In most parts of Southern, Western and Northern Europe and Americas, major gray wolf populations are lost and their present distribution is highly fragmented [2]. In England and Japan, it has already become extinct while in many other parts it is on the verge of extinction [3]. The wolves are now restricted to few large forested areas in Eastern Europe, a few isolated mountain ranges in the Mediterranean, mountains and semi-desert areas of middle-East and the wilderness areas of North America, Russia and China, with largest concentrations in Russia followed by Canada and Alaska [cf. 4]. Therefore, serious efforts have been initiated in Europe and America for protection and conservation of the existing wolf populations in recent years.

In the last century, up to 32 sub-species of gray wolf have been described by wolf taxonomists, based on the variations in physical features, behavioral aspects and geographical distribution, many of which are now believed to be extinct [4, 5]. Status of many of these subspecies is contested eg., subspecies *campestris*, *chanco* and *desertorum* are suggested to be synonymous of subspecies *lupus* [6], making the gray wolf taxonomy highly debatable pending validation and revision. Presently, only six subspecies viz., *lupus*, *campestris*, *albus*, *tundaram*, *lycaon* and *nubilus*, are believed to be surviving globally, and only the first two being in Asia [4]. In recent years, though many molecular-genetic studies were carried out to understand the origin and evolution of canidae family [7, 8 and cf. therein; 9] and also on the population dynamics of isolated, specific gray wolf populations [3, 10-12], similar studies to resolve the largely debatable gray wolf taxonomy at the subspecies level are generally lacking. The two detailed studies on the latter aspect based on mitochondrial genome variations (mtDNA d-loop haplotyping) show little genetic partitioning on the continental or regional scales [13, 14].

In Indian subcontinent, wolf is known to exist for a longtime but not much is known about its origin and evolutionary history. Generally, as per the conventional taxonomy, it is believed that the small numbers of wolves found in today's India, are two different sub-species of *Canis lupus* which are represented by geographically isolated broadly non-overlapping (allopatric) populations. 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 [15]. This Himalayan wolf (HW) population, adapted to the cold environment, is considered to be representing the extant population (direct relatives) of the

relatively better known Tibetan wolf, *Canis lupus chanco*, which is found throughout central Asia with its range extending into Tibet, China, Manchuria and Magnolia [15]. On the other hand, the second wolf population estimated to be less than 1500 (<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 gray wolf (GW), are believed to represent the second sub-species in India, i.e., *C. lupus pallipes* that is also found in the middle-Eastern countries like Iran and Israel [16]. The wolf in India are accorded schedule '1' endangered species status under the Indian Wildlife Protection Act of 1972 and of CITES, and in last few decades federal efforts were initiated for wolf protection and conservation by setting up wildlife preserves in the country and captive breeding programs in few National Zoological Parks.

Since early 1970s, to help the conservation efforts few studies were also initiated to understand the ecology and population biology of wolves, but all of these were on GW of peninsular India (Wildlife Institute of India website). The only detailed account on HW population is by Fox and Chundawat [15] and its earlier references in the wildlife surveys of snow leopard [17]. There is essentially no study attempted to understand the genetics of Indian wolves and their relationship with wolves from other parts of the world. Perusal of the published literature on molecular-genetic analysis of wolf and other canids, reveals only six analyzed samples referred as Indian wolf, five of which originated from Afghanistan [9] and the source of the only one sample used in studies by Wayne group [14, 18] is indicated to be from India. Under this dismal scenario and ever increasing threat to the remaining small wolf populations in India, it becomes imperative to understand their genetic make-up and exact taxonomic status between themselves and with other world-wide wolf populations to have a long-term and effective plan in place for their continual survival and conservation.

Many features of mtDNA viz., high mutation rate than nuclear genome; uniparental (maternal) inheritance, high copy number and lack of recombination, makes it the biological material of choice for mirroring the evolutionary past of a species. Moreover, the mtDNA diversity can be efficiently used to infer both, relatively recent as well as ancient evolutionary events by selectively analyzing selection-pressure independent (d-loop control region) or dependent (essential functional genes like cytochrome-B) domains. These attributes made mtDNA analysis a powerful tool in the molecular systematics pursuits, which in recent years has extensively and successfully been utilized to understand the population history, phylogeography and taxonomic relationships of a large number of animal species. Our present day understanding of origin, evolution and taxonomy of wolf and related canid species is largely based on mtDNA polymorphism analysis [8, 9, 14, 18].

It is in this background; in 1998 we initiated molecular characterization studies on the highly endangered HW to understand their genetic structure and taxonomic status relative to the GW and other wolf-like canids from Pennisular India and wolf populations from rest of the world. Our detailed mitochondrial DNA polymorphism studies on a large number of HW and representative samples of GW, jackal and wild dog of India followed by comparative genomics with information available for world-wide wolf populations have led to some novel and startling findings about the origin and evolution of wolves of India. Significantly, our studies suggest that: a) both the Indian wolf populations are genetically unique within themselves and both of them from all other wolf populations recorded world-wide; b) HW is distinct from the purported similar wolf from China, *C. lupus chanco*, and represent the most ancient wolf lineage ever recorded and analyzed using molecular means; c) similarly, GW is distinct from the purported similar wolves from South-East Asia, *C. lupus pallipes*, and probably also represent the other ancient wolf lineage from India; d) each of the two wolf populations from India thus represents a case of new wolf species in evolution waiting to be re-christened; e) Indian subcontinent as one ancient center of wolf origin and diversification, wherein the two wolf populations represent wolf/canid lineages derived from an ancestral wolf radiation that underwent evolution independent from other world wolf populations for the last 1-2 mya. Thus, our results increase the biological significance of the Indian canid populations and warrants urgent and effective conservation measures for their protection and management.

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 km² encompassing the Trans-Himalayan region of Jammu and Kashmir and Himachal Pradesh [15]. Apart from these animals in the wild, there are a total of only 21 live animals in four of the Zoological Parks of India [19]. Most of these animals are captive bred from few wild caught animals from the Transhimalayan region. In the present study we analyzed 18 of these animals available in the Padmaja Naidu Himalayan Zoological Park, Darjeeling, West Bengal. In addition to these animals, we analyzed wild caught representative samples of related canids from middle India viz. gray wolf (originally caught from three widely separated geographical regions in middle India), wild dog and Indian jackal, available in the Nehru Zoological Park, Hyderabad, Andhra Pradesh and Sakkarbaug Zoo, Junagarh, Gujarat (Table 1).

DNA isolation

From each animal, ~10 ml blood was drawn using vacutainers having EDTA as anti-coagulant and transported from the Zoo to the laboratory at room temperature where the same were stored at 4⁰C till further use. The total genomic DNA was isolated from each of the samples following the lysis-freeze fracture method [20] for molecular analysis.

DNA amplification and sequencing for nucleotide diversity/ DNA polymorphisms

To understand the genetic status of HW in relation to other related canids from India and other parts of the world, nucleotide diversity was ascertained for three taxonomically informative domains of the mitochondrial genome. These included, part of cytochrome-B gene, complete cytochrome-B gene for few representative samples, hypervariable d-loop control region and 16S rDNA. In each case, the target mitochondrial sequences were PCR amplified using standard protocols on PE9600 thermocycler. The PCR reactions were set using approximately 50 ng of template genomic DNA in 15-20 µl reaction volume containing: 5 pico-moles of each primer, 150 µM dNTP, 1.5 mM MgCl₂, 0.1 M KCl, 20 mM Tris-HCl, and 0.5 to 1.0 U of AmpliTaq Gold polymerase (Perkin Elmer). In general, PCR profile comprised an initial denaturation of 10 min at 95°C, followed by 35 cycles of: 94°C for 1 min., 50-55°C for 1 min., 72°C for 2 min.; and 72°C for 5 min. The PCR products were treated with exonucleaseI and shrimp alkaline phosphatase at 37°C and 80°C, 15 min each to remove the remaining primers and free nucleotides. Subsequently, the purified PCR products (~100 ng) were sequenced for both strands using the Big dye terminator ready reaction kit (Perkin Elmer). Further, to confirm sequence data, many of the samples were sequenced more than once. Sequencing PCR cycle condition were: 30 cycles of 96°C for 10 sec., 50°C for 5 sec. and 60°C for 4 min. Extended products were purified by alcohol precipitation followed by washing with 70% ethanol. The processed samples were then dissolved in loading dye and sequenced using automated DNA sequencer ABI Prism377 or ABI Prism3700. Sequences were edited using auto-assembler software package for further analysis.

Cytochrome-B gene: A ~332 bp conserved region of mitochondrial cytochrome-B gene towards the amino-terminus of the protein was amplified and directly sequenced using primers L14841 [5'-AAAAGCTTCCATCCAACATCTCAGCATGATGAAA-3'] and H15149 [5'-AAACTGCAGCCCCTCAGAATGATATTTGTCCTCA-3'] described by Kocher et al [21], for all the samples. In some cases, the PCR amplified product was gel purified using gene clean kit (BIO 101), and cloned in pMOS T vector (Amersham Pharmacia Biotech) following manufacturers details. In each case about 10 recombinant clones were randomly chosen and sequenced as above using M13 universal sequencing primers.

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 [5'-CGAAGCTTGATATGAAAAACCATCGTTG-3'], H-15915 [5'-AACTGCAGTCATCTCCGGTTTACAAGAC-3'] and internal primers L-15162 [5'-GCAAGCTTCTACCATGAGGACAAATATC-3'], L-15513 [5'-CTAGGAGACCCTGACAATA-3'] and H-15149 [see above] [22].

D-loop control region: For each of the sample, the complete d-loop region (Figure-1), 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 [9]. Initially the region was amplified using primers DL-14 [5'-TGTAACGACGGCCAGTGCTCTTGCTCCACCATCA-3'] and DH-7 [5'-CAGGAAACAGCTATGGCCCGTGCGACTCATCTTGGC-3'], which amplified a product of ~1300 bp. To sequence the amplified products, two more internal primers were used along with the above two. The two internal primers were DH-6 [5'-CAGGAAACAGCTATGGCCCCCTTGATTTTATGCG-3'] and DH-10 [5'-TGTAACGACGGCCAGTATCACTCATCTACGACCG-3'].

Partial d-loop sequences were also amplified and sequenced using the primers DLH-16340 [5'-CCTGAAGTAGGAACCAGATG-3'] and Thr-L15926 [5'-CAATTCCCCGGTCTTGTGAACC-3'] [14], which essentially gives ~first 250 bases that are obtained using the primer DL-14 (that span the t-RNA^{Pro} gene, see above) in addition to about 70 additional bases corresponding to phenylalanine t-RNA^{Phe} that is present on the left of t-RNA^{Pro}.

16S rDNA: Part of 16 sRDNA was amplified and sequenced using the primers 16Sar-L [5'-CGCCTGTTTATCAAAAACAT-3'] and 16Sbr-H [5'-CCGGTCTGAACTCAGATCACGT-3'] [23] for all the 27 samples analyzed in the study.

EMBL databse sequences for comparative analysis

The Genbank database (National Center for Biotechnology Information, USA: NCBI Home page <http://www.ncbi.nlm.nih.gov>) was searched for similar sequences using BLAST search. Based on the search, large number of reference mitochondrial DNA sequences described for wolf and other related canid species were retrieved from the database, for each of the analyzed domain for 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 Tables 2 and 3. The 90 d-loop reference sequences retrieved for comparative analysis represented the wolf

samples from all over the globe across all the continents. In comparison, there were only few wolf reference sequences specific to cyto-B gene and none for 16S ribosomal DNA. Accordingly, for these two mitochondrial domains reference sequences were from related canid genera/species.

Data analysis

The sequences obtained in the study were aligned with corresponding reference sequences from the EMBL database using the CLUSTAL program [24] to identify variable nucleotides. The sequences were further edited using GENEDOC program [25]. Diversity statistics of the compared sequences were obtained using Dna-SP ver. 3.51 [26; <http://www.bio.ub.es/~julio/DnaSP.html>]. The nucleotide diversity was used to derive genetic distance estimates and phylogenetic relationships using various analytical routines available in the software package PHYLIP 3.6 [27; <http://evolution.genetics.washington.edu/phylip.html>] and Phylo_Win [28]. The aligned sequences, manually checked for gaps and edited at the ends to avoid missing information in some of the compared taxa/entries were used to calculate genetic distance coefficients corrected for multiple hits by the Kimura two parameter [29], TN [30] and HKY model [31] which take into account unequal nucleotide frequencies in present day sequences and also unequal transitions/transversion rates. The distance estimates were used to construct phylogenetic trees using neighbor joining (NJ) and UPGMA methods [32-34]. The phylogenetic analysis of the character-state matrices (aligned sequence data) was also carried out using the maximum-parsimony and maximum-likelihood approaches and Kimura 2-parameter with gamma correction. Support for nodes found on the shortest tree was assessed by bootstrap analysis [35].

The phylogenetic trees based on mtDNA d-loop polymorphisms, were rooted using Indian jackal, which represent one of the nearest outgroup species of gray wolf in the *Canis* genus [8]. Initially, raccoon dog and fox, the two distant canid relatives of wolf (Table 2) were also used for comparison, but were not used in final analysis. Similarly, for 16S rDNA based analysis also Indian jackal was used as an outgroup taxa, as the corresponding data was not available for any other canid in the EMBL database. In comparison, no particular outgroup was defined in the cytochrome-b gene based analysis, as the genetic status of wolf is well defined in relation to the other canids and the purpose of the study was only to define the place of Indian wolves relative to the wolf from elsewhere.

The corrected Kimura distance estimate observed for HW-GW pair wise comparisons of cytochrome-B sequence data were used to derive the estimate of evolutionary time separating the two, using the conservative estimates of sequence divergence rates reported for canid mtDNA divergence [8].

Results and Discussion

Molecular genetic assessments have become necessary and important guides for description of biotic diversity and its management. This is so because the conventional taxonomy might fail to recognize phylogenetically distinct forms, misdirecting the conservation efforts for endangered species for protecting biological diversity [36]. In India there are two small populations of wolves (HW and GW) believed to be two different subspecies of *C. lupus* that are found in many other parts of the Middle-East/Central Asia. One of these populations of HW which numbers only ~350 in the wild [15] and is highly endangered, is not even recognized in SSC/IUCN list on wolves of the world probably, among other reasons, because of its purported status of a subspecies *chanco* that is known from many other countries. The present study, first on molecular genetic analysis of HW and other canid samples from India, involving DNA typing of haplotypic diversity across cytochrome-B gene, 16S rDNA and D-loop control region of mtDNA, reveals many novel insights into the evolutionary past and taxonomic relationships of the present day extant populations of wolves of India. Most importantly, the data reveal HW and also the GW from India to be the most ancient wolf lineages genetically unique from all those found world-wide suggesting a need for revision of the wolf taxonomy and present day understanding of its origin and evolution. The results have highly significant implications in the conservation priorities and efforts on wolf both at the National and International level.

mtDNA polymorphisms

We amplified and sequenced 2000-2200 bp of three domains of mtDNA, ~332 bp of protein coding cytochrome-B gene, 544 bp of 16S rRNA gene and 1200-1370 bp spanning the total d-loop control region from 23 wolf samples (18 HW, 5 GW) and 2 samples of Indian jackal. (Table 1). In case of Indian wild dog, only regions of cytochrome-B and 16S rRNA gene could be amplified and sequenced. In addition, the complete cytochrome-B gene was also amplified and sequenced for one sample each of HW, GW and jackal. Using BLASTn search with the respective sequences, most of the related sequences mainly of different wolf populations/sub-species and many of the related canid species were retrieved from the EMBL database for comparison and phylogenetic analysis (Tables 2, 3). The sequence data thus compared and analyzed for DNA polymorphisms are summarized in table 4.

Cytochrome-B gene polymorphism

The mitochondrial cytochrome-B sequencing analysis of 322 bp region, of the 18 HW and 3 GW samples revealed a total of four transitions (sequence divergence = 1.24%) distinguishing the two

types, and two additional transitions defining two haplotypes for each of them. In comparison, HW differed for 8 transitions (3 A>G, 1 G>A, 4 T>C; sequence divergence = 2.48%) and GW for 4 transitions (2 A>G, 2 T>C) from the holarctic wolf. The data thus show the three wolf types to be genetically different from each other with HW wolf to be most divergent. Further, comparisons with the other canid species reveal HW to be most basal among the three types nearest to their immediate outgroup species in the canidae family (Figure 2a).

The rate of mitochondrial DNA sequence divergence within the canids has been estimated to be 1.3-1.7% per million years [8]. Considering a conservative average estimate of approximately 1.5% per MY, the above observed sequence divergence would mean a divergence of ~0.8 MYA between HW and GW and ~1.65 MYA between HW and holarctic wolf. Further, these divergence estimates are much more than what has been reported for intraspecific variation [cf. 8], suggesting the three to be more akin to three different species.

The complete 1140 bp sequencing of the mitochondrial cytochrome-B gene also revealed similar divergence between HW and GW and HW to be more near to jackal (Figure 2b).

16S rDNA

Sequencing of ~550 bp of the mitochondrial 16S rDNA revealed relatively low variation among all the 27 samples. Only 32 nucleotides were found to be polymorphic and informative. EMBL database search didn't reveal similar sequences for the reference wolf populations, and thus detailed phylogenetic analysis was not done. Nevertheless, the data were analyzed for the samples sequenced in the study that revealed the broad generic affinities of HW and GW to the jackal /dog clad and these being very divergent from the wild dog *Cuon alpinus* (data not shown).

D-loop control region

Sequencing of complete D-loop analysis revealed a very large number of substitutions and indels specific to HW and GW compared to all other wolf haplotypes documented world-wide, all across the 5' CR-I, hypervariable repetitive domain as well as relatively conserved 3' CR-II region.

In general, the d-loop CR-I region was found to be most informative in ascertaining the genetic status of the Indian wolf types (HW and GW) because of the availability of the corresponding data for other wolf types from outside India. The sequence comparison and phylogenetic analysis of the CR-I polymorphism over its varying lengths established Indian wolves not only to be genetically very distinct from all other wolves but also forming a non-overlapping separate clade basal to the whole phylogenetic tree of world-wide wolves, more closer to the Jackal used as outgroup canid

species in the study (Figures 3-6). This genetic differentiation was overwhelmingly evident in the large number of unique sequence polymorphisms revealed in HW and GW sequences. When compared to all other reported haplotypes over 655 bp of the CR-I region, we found 2 deletions (comprising 6 bp), 1 transition (C>T) and 1 transversion (G/A>T) unique to both HW and GW. In addition, there were 12 transitions (4 G>A, 3 T>C, 5 C>T) and 2 deletions specific to the HW, and 8 transitions (5 C>T, 2 T>C, 1 G>A) specific to GW. All the 18 HW samples were represented by only one haplotype (HW-India) whereas there were two haplotypes seen for GW samples (GW-1, GW-2) differing in only one transition.

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 repeats as well as due to extensive intra-individual heteroplasmy for specific nucleotides of the repeat motifs (Figure 1). The 10-mer repeat motif of HW and GW was similar to what had been reported earlier by 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 to HW and GW, the repeat motif for Indian jackal though of same length and sequence, but had two variable nucleotides in its motif. The data could not be compared, as no corresponding data are available for other wolf populations.

Similar to CR-1 region, many variations in the 300 bp CR-II region were found unique for HW and GW when compared to the few corresponding haplotypes reported earlier studies. There was only one haplotype for all the samples of HW and similarly one in case of GW. There were three transitions (1 C>T, 1 A>G, 1 G>A) and one transversion (A>C) unique to HW haplotype and one 4-bp indel (TATA) and one transition (T>C) unique to the GW haplotype when compared to the eight other haplotypes documented in earlier studies.

Phylogenetic analysis and uniqueness of Indian wolves

Phylogenetic analysis done using different character state (ML, MP) and distance (NJ) approaches for all the three types of mitochondrial DNA sequences analyzed in the study revealed almost identical tree topologies with little variation, if any, in branch lengths and/or finer arrangements in a given cluster. Significantly, in all the analyses, the two Indian wolf types (HW and GW) turned out to be genetically unique and most ancient from all other worldwide wolf-populations, but essentially belonging to same clade of the gray wolf in the context of canidae family with jackal as the closest ancient relative (Figures 2-6). The taxonomic potential of the three different domains analyzed in the present study is in complete agreement with what has been widely demonstrated and recently

outlined by Tautz et al [37]. While the quickly evolving mitochondrial control d-loop region was found to be most informative to understand the taxonomic affiliations of Indian wolves to other related wolf populations, the relatively conserved mitochondrial cytochrome-B gene and 16S rDNA were useful in resolving their higher level phylogenetic status in the canidae family.

DNA polymorphism of the partial cytochrome-B gene has been used to reconstruct the canidae systematics [8], confirms the two Indian wolf types to be members of the jackal-cyote-wolf clade, more similar but distinctly basal to the holarctic gray wolf and dog lineage (Figure 2a). Similarly, comparison of the complete cytochrome-B gene for which only two reference sequences belonging to canidae could be found in the public domain [38], also reveal HW and GW to be descendant of the jackal lineage but basal to the dog supported with high bootstrap values (Figure 2b). Between HW and GW, the former was conspicuously more basal in the lineage. The taxonomic status of HW and GW from India to the Jackal-Wolf-Dog clade was equally evident in 16S rDNA based phylogram (not shown).

In comparison, the d-loop polymorphism for which the corresponding data are available for most of the world-wide wolf populations, was found to be most informative to ascertain the genetic status of HW and GW vis-à-vis all the previously analyzed wolf populations from outside of India. Our extensive comparative analysis involved almost all the wolf haplotypes described in the literature to date, covering different lengths of the d-loop CR-1 region [viz., 233 bp (76 sequences, 32 haplotypes), 545 bp (46 sequences= 35 haplotypes), 651 bp (29 sequences, 24 haplotypes)]. The analysis conclusively revealed both Indian wolf types (HW and GW) to be genetically most divergent and forming a completely separate clade without even a single overlap with any of the haplotypes seen in the other wolves of the world. Moreover, the HW and GW clade was found to be closest to the jackal, one of the nearest ancestral canid species and basal to the other major clade comprising all other wolf haplotypes, clearly indicating these to be the derivatives of an more ancient independent wolf radiation.

Significantly, both HW and GW were found to carry significant 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 later for 24 nucleotides over 655 bp of CR-I region and does not carry any of the latter's diagnostic polymorphism. These data conclusively show that HW found in India is genetically very different from the more abundant and widely distributed Tibetan wolf *chanco*. Similarly, the GW samples analyzed in the study, neither have the diagnostic nucleotide polymorphism (A>G) observed by Tsuda et al [9] in *pallipes* samples

originated from Afghanistan, nor show any similarity with the wolf haplotype 'w12' reported for the only other *C. lupus pallipes* sample that also originated from India [14, 18]. Moreover, a large number of diagnostic polymorphisms (see above) were seen in the GW samples in the present study that were not present in any of the above few samples of *C. lupus pallipes* reported earlier [9, 14, 18]. The 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 analyzed by Vila et al [14] having a very different haplotype (w12) can be explained by the possibility of a wolf-female dog hybridization, a phenomenon that has been observed and reported earlier. Incidentally, in the analysis carried out by Vila et al [14], 'w12' haplotype of the Indian wolf was also one of the closest wolf haplotypes to those of the dog haplotypes. Simultaneously, it is significant to mention here that the GW samples in the present study, though few in numbers (n=5) were wild animals, caught from geographically different parts of the peninsular India giving strength and credence to the results of the study.

Possible origin of HW and GW of India

All the three types of data strongly suggest HW to represent one of the oldest wolf lineage that was followed by the one that gave rise GW. These two lineages were revealed to be distinct from the third seemingly more recent lineage leading to a broad loosely structured clade comprising all other global wolf populations. It was noteworthy that in all the phylogenetic reconstructions tested in the study, both the lineages specific to Indian wolves appear together in a genetically distinct completely separate monophyletic clade with no overlap with any of the haplotypes of the above referred other broader clade for wolves from rest of the world. These results suggest that Indian wolves represent a relic ancestral lineage of wolf that evolved in isolation for a very long time from the evolutionary lineage that is represented by all other wolves from outside India. The estimates of evolutionary time based on rate of sequence divergence suggest ~1.5 mya of age for the lineage that resulted in the HW and GW clade.

In light of the above results and our present day understanding of canid systematics [8, 18], we hypothesize that the wolves of India represent a pre-wolf, post-jackal, ancestral canid radiation that got separated/arrived in India sometime in early Pleistocene about 1-2 million years ago and since then evolved independently without much interaction with the evolving wolf lineages in other parts of the world. This in turn would suggest that Indian subcontinent had been one major center of wolf/canid evolution, an inference that is in agreement with the school of taxonomists that ascribe an Asian origin to the wolf like canids [6; <http://www.naturalworlds.org/wolf/history/evolution.htm>].

Furthermore, the results of the present study raise interesting questions about the origin and domestication of dogs in India viz., a) whether there are unique breeds that might have been domesticated from the HW and GW independently; or b) dog breeds in India are from Middle-East, Central Asia and Eurasia brought in by invading Mughals, Mongols, Europeans and other visitors over many centuries. The available data from extensive studies on dog origin and domestication [18, 39] would support the second alternative, as few Indian dog samples analyzed therein align with wolf haplotypes (from the above geographical regions) which are essentially very different from the HW and GW analyzed in the present study. Nevertheless, answers to these questions are beyond the scope of the presented data, and calls for further detailed studies to address the origin and evolution of dog breeds in India.

Conservation Significance

Molecular genetic information can be of considerable value in appraisal of phylogenetic discontinuities within and among species. Taxonomies based on morphological traits alone sometime provide inadequate or misleading guides to phylogenetic distinctions but still remain central to our perceptions of biotic diversity including recognition of endangered forms/species. Many molecular genetic studies in recent years have demonstrated such prior systematic errors wherein distinct defined taxonomic groups show little evolutionary differentiation and vice versa [40], raising serious concerns about the veracity and efficacy of conservation efforts for endangered species. Such cases highlight the need for inputs from high-resolution molecular genetics approaches that can provide firmer foundation for updating the taxonomy and hence help management of biotic diversity in more realistic way [36].

The above concerns are vastly vindicated by the present study that conclusively discounts the currently believed taxonomic status of Indian wolves based on conventional systematics. Neither HW nor GW shares any haplotypic similarity with the purported sub-species of *C. lupus* i.e., *chanco* and *pallipes*, which the two are believed to represent. On the contrary, the genetic diversity analyzed in the present study demonstrates these to be genetically unique to the level of almost new species/sub-species in evolution. These novel findings greatly enhance the biological significance of these already endangered Indian wolf populations and calls for urgent reorientation and prioritization of conservation efforts for their protection and management. Simultaneously, the study reinforces the need for more exhaustive efforts to update and redefine the already complex and debated wolf taxonomy.

Finally, our results demonstrate that HW and GW from India are probably the most enigmatic wolves of the world and there is a necessity to revise their existing nomenclature, rechristening them to new species/sub-species status i.e., may be *C. himalayensis* (or *C. lupus himalayensis*) and *C. indica* (*C. lupus indica*), respectively. Infact, such a proposal for HW was made earlier by the first author at the ‘International Conference on Canid Biology and Conservation’ held at the Oxford University, UK, in 2001, alongwith the proposal to include it in the Red list of Canids’. To our understanding, these proposals were identified for recommendation to SSC/IUCN and wolf specialist group for further consideration. However, it was felt that the study should be extended with more samples to validate and strengthen the case for a separate species status, which we strongly believe is fulfilled in this present work.

Acknowledgements

We thank the Central Zoo Authority of India, New Delhi, Chief Conservator of Forests, West Bengal, Andhra Pradesh and Gujarat, Ministry of Environment and Forests, Government of India for necessary support and permissions, and Mr. M. Banerjee, Curator Nehru Zoological Park, Hyderabad, Dr. B. R. Sharma, curator Padmaja Naidu Himalayan Zoological Park, Darjeeling, Curator, Sakkarbagh Zoo, Junagarh, and their teams of veterinary doctors and animal keepers, for helping with the collection of samples; Mehar Sultana for primer synthesis, Raman Agrawal, TP Velavan and Udaykumar for helping with sequencing, Dr. R. Rajkumar for helping with data analysis and Dr. Yogendra Sharma for checking many drafts of the manuscript and helpful suggestions. The study was carried out under the CZA, DBT and CSIR supported mega project LACONES on conservation of endangered animals of India.

References

- [1] Goldman EA: **The wolves of the North America, Part-II. Classification of wolves.** In: *The wolves of North America.* eds Young SP, Goldman EA. Dover, New York/American Wildlife Institute, Washington USA, **1944**: 389-636.
- [2] Mech LD: *The wolf: the ecology and behaviour of an endangered species.* The Natural History Press, New York, USA, **1970**: 384 pp.

- [3] Ellegren H, Savolainen P, Rosen B: **The genetical history of an isolated population of the endangered grey wolf *Canis lupus*: a study of nuclear and mitochondrial polymorphisms.** *Phil Trans R Soc Lond B* **1996**, **351**: 1661-1669.
- [4] Anonymous: *The New Encyclopedia of mammals*. Ed: Macdonald D. Oxford University Press, Oxford, UK, **2001**: 42-43.
- [5] Hall ER: *The mammals of North America*. John Wiley & Sons, New York, USA, **1981**: 1181 pp.
- [6] Busch RH: *The Wolf Almanac*. The Lyons Press, **1993**: pp 241.
- [7] Roy MS, Geffen E, Smith D, Ostrander EA, Wayne RK: **Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci.** *Mol Biol Evol* **1994**, **11**: 553-570.
- [8] Wayne RK, Geffen E, Girman DJ, Koepfli KP, Lau LM, Marshall C: **Molecular systematics of the canidae.** *Syst Biol* **1997**, **46**: 622-653.
- [9] Tsuda K, Kikkawa Y, Yonekawa H, Tanabe Y: **Extensive interbreeding occurred among multiple matriarchial 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* **1997**, **72**: 229-238.
- [10] Wayne RK, Jenks SM: **Mitochondrial DNA analysis implying extensive hybridisation of the endangered red wolf *Canis rufus*.** *Nature* **1991**, **351**: 565-568.
- [11] Gottelli D, Sillero-Zubiri C, Applebaum GD, Roy MS, Girman DJ, Garcia-Moreno J, Ostrander EA, Wayne RK: **Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*.** *Mol Eco* **1994**, **3**: 301-312.
- [12] Ellegren H: **Interbreeding and relatedness in Scandinavian gray wolves.** *Hereditas* **1999**, **130**: 239-244.
- [13] Wayne RK, Lehman N, Allard NW, Honeycutt RL: **Mitochondrial DNA variability of the grey wolf: genetic consequences of population decline and habitat fragmentation.** *Conserv Biol* **1992**, **6**: 559-569.
- [14] Vila C, Amorim IR, Leonard JA, Posada D, Castroviejo J, Petrucci-Fonseca F, Crandall KA, Ellegren H, Wayne RK: **Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*.** *Mol Ecol* **1999**, **8**: 2089-2103.
- [15] Fox JL, Chundawat RS: **Wolves in the Transhimalayan region of India: the continued survival of a low density population.** In: *Ecology and conservation of wolves in a changing world*- Proceedings of the 2nd North American Symposium on wolves held in Edmonton, Alberta, Canada (25-27 Aug, 1992), **1993**: 95-104.
- [16] Shahi PS: **Status of grey wolf (*Canis lupus pallipes*) in India: a preliminary survey.** *J Bombay Nat His Soc* **1982**, **79**: 493-502.

- [17] Fox JL, Sinha SP, Chundawat RS, Das PK: **Status of the snow leopard, *Panthera uncia*, in north-west India.** *Biol Conserv* **1991**, **55**: 283-298.
- [18] Vila C, Savolainen P, Maldonado JE, Amorim IR, Rice JE, Honeycutt RL, Crandall KA, Lundeberg J, Wayne RK: **Multiple and ancient origins of the domestic dog.** *Science* **1997**, **276**: 1687-1689.
- [19] Anonymous: *Inventory of animals in Indian Zoos 2000-2001*. Central Zoo Authority, New Delhi, India, **2002**: pp 314.
- [20] Aggarwal RK, Majumdar KC, Lang JW, Singh L: **Generic affinities among crocodylians: as revealed by DNA fingerprinting using a Bkm-derived probe.** *Proc Natl Acad Sci USA* **1994**, **91**: 10601-10605.
- [21] Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC: **Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers.** *Proc Natl Acad Sci USA* **1989**, **86**: 6196-6200.
- [22] Irwin DM, Kocher TD, Wilson AC: **Evolution of the cytochrome-b gene of mammals.** *J Mol Evol* **1991**, **32**: 128-144.
- [23] Palumbi SR, Martin A, Romano S, McMillan WO, Stice L, Grabowski G: *The simple fool's guide to PCR, version 2*. Zoology Department, University of Hawaii, Honolulu, HI-96822. **1991**: pp. 46.
- [24] Higgins DG, Sharp PM: **Fast and sensitive multiple sequence alignments on a microcomputer.** *CABIOS* **1989**, **5**: 151-153.
- [25] Nicholas KB, Nicholas HBJr: **GeneDoc: a tool for editing and annotating multiple sequence alignments.** [www.psc.edu/biomed/genedoc] **1997**.
- [26] Rozas J, Rozas R: **DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis.** *Bioinformatics* **1999**, **15**: 174-175.
- [27] Felsenstein J: *Documentation of PHYLIP (Phylogeny Inference Package) version 3.5c*. Univ Washington, Seattle, USA, **1993**.
- [28] Galtier N, Gouy M, Gautier C: **SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny.** *CABIOS* **1996**, **12**: 543-548.
- [29] Kimura M: **A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences.** *J Mol Evo* **1980**, **16**: 111-120.
- [30] Tajima F, Nei M: **Estimation of evolutionary distance between nucleotide sequences.** *Mol Biol Evol* **1984**, **1**: 269-285.
- [31] Hasegawa M, Kishino H, Yano T: **Dating of the human-ape splitting by a molecular clock of mitochondrial DNA.** *J Mol Evol* **1985**, **21**: 160-174.
- [32] Saitou N, Nei M: **The neighbour-joining method: a new method for reconstructing phylogenetic trees.** *Mol Biol Evo* **1987**, **4**: 408-425.

- [33] Sokal RR, Michener CD: **A statistic method for evaluating systematic relationships.** *Univ Kans Sci Bull* **1958**, **28**:1409-1438.
- [34] Sneath PHA, Sokal RR: *Numerical taxonomy*. WH Freeman, San Francisco **1973**.
- [35] Felsenstein J: **Confidence limits on phylogenies: An approach using the bootstrap.** *Evolution* **1985**, **39**:783-791.
- [36] Avise JC: **A role for molecular genetics in the recognition and conservation of endangered species.** *TREE* **1989**, **4**: 279-281.
- [37] Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP: **A plea for molecular taxonomy.** *Trends Ecol Evol* **2003**, **18**: 70-74.
- [38] Castresana J: **Cytochrome-b phylogeny and the taxonomy of great apes and mammals.** *Mol. Biol Evol* **2001**, **18**: 465-471.
- [39] Savolainen P, Zhang Y, Luo J, Lundeberg J, Leitner T: **Genetic Evidence for an East Asian origin of domestic dogs.** *Science* **2002**, **298**: 1610-1613.
- [40] Avise JC, Nelson WS: **Molecular genetic relationships of the extinct Dusky Seaside Sparrow.** *Science* **1989**, **243**: 646-648.
- [41] Lento GM, Hickson RE, Chambers GK, Penny D: **Use of spectral analysis to test hypotheses on the origin of pinnipeds.** *Mol Biol Evol* **1995**, **2**: 28-52.
- [42] Randi E, Lucchini V, Christensen MF, Mucci N, Funk SM, Dolf G, Loeschcke V: **Mitochondrial DNA variability in Italian and East European wolves: detecting the consequences of small population size and hybridization.** *Conserv Bio*, **2000**, **14**(2): 464-473.
- [43] Okumura N, Ishiguro N, Nakano M, Matsui A, Sahara M: **Intra- and interbreed genetic variations of mitochondrial DNA major non-coding regions in Japanese native dog breeds (*Canis familiaris*).** *Anim Genet* **1996**, **27**: 397-405.
- [44] Pilgrim KP, Boyd DK, Forbes SH: **Testing for wolf-coyote hybridization in the Rocky Mountains using mitochondrial DNA.** *J. Wildl Manage* **1998**, **62**(2): 683-689.
- [45] Koop BF, Crockford S, Burbidge M, Rink U: **Polyphyletic origin of dogs [Canada – sub species crassdon].** As per EMBL entries, **1998**.
- [46] Valiere N, Fumagalli L, Gielly L, Miquel C, Lequette B, Pouille ML, Weber JM, Arlettaz R, Taberlet P: **Wolf recolonization of France and Switzerland inferred from non-invasive genetic sampling over a period of 10 years.** As per EMBL entries, **2001**.
- [47] Jedrzejewski W, Branicki W, Veit C, Medjugorac I, Bunevich A, Jedrzejewska B, Schmidt K: **Genetic diversity and population structure of wolves, *Canis lupus*, in Bialowieza Primeval Forest (Poland and Belarus).** As per EMBL entries, **2001**.
- [48] Savolainen P, Arvestad L, Lundeberg J: **Mitochondrial DNA tandem repeats in Russian wolf (only repeat region).** As per EMBL entries, **1999**.

Legends to the figures

Figure 1:

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).

Figure 2a:

ML tree based on polymorphism in 322 bp of Cyto-B gene showing relationship of HW with other canids from India and elsewhere; Distance: Gamma corrected K 2-P.

Figure 2b:

ML tree based on complete Cyto-B gene polymorphism showing relationship of HW with jackal and gray wolf from India and 2 reference canids from elsewhere; Distance: Gamma corrected K 2-P.

Figure 3:

ML tree based on 233 bp of d-loop region-I polymorphism showing relationship of HW with 75 other wolf populations from India and elsewhere and 2 dog species; Distance: Gamma corrected K 2-P. Reference taxa are as per codes in Table-3.

Figure 4:

NJ tree based on 545 bp of d-loop control region-I polymorphism showing relationship of HW with 45 other wolf populations from India and elsewhere and 2 dog entries; Indian jackal is used as outgroup; Distance: Corrected K 2-P. Reference taxa are as per codes in Table-3.

Figure 5:

ML tree based on 651bp of d-loop control region-I polymorphism showing relationship of HW with 28 other wolf populations from India and elsewhere and 2 dog entries; Indian jackal is used as outgroup; Distance: Gamma corrected K 2-P. Reference taxa are as per codes in Table-3.

Figure 6:

NJ tree based on 675bp complete d-loop control region-I polymorphism showing relationship of HW with 20 other wolf populations from India and elsewhere and 1 dog species; Indian jackal is used as outgroup; Distance: corrected K 2-P. Reference taxa are as per codes in Table-3..

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 ^a	Wild caught	Pennisular India	Nehru Zoological Park, Hyderabad (Andhra Pradesh);
Jackal (J), [<i>C. aureus</i>]	2	Wild caught	Pennisular India	Sakkarbaug Zoo, Junagarh (Gujarat)
Wild dog (WD) [<i>Cuon alpinus</i>]	2	Wild caught	Pennisular India	

DNA sequencing Analysis done:

Mitochondrial Domain	Samples analyzed (n=number)
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)

a : Samples originated from three different states of pennisular India

Table-2:
Reference mitochondrial Cytochrome-b and 16S RNA gene sequences used in the phylogenetic analysis

	Species	Common name	Continent / Country	EMBL Acc.No.	Reference
<i>Partial Cytochrome-b gene reference sequences</i>					
1	<i>C. lupus lupus</i>	Gray wolf	Holarctic	AF028141	[8]
2	<i>C. simensis</i>	Ethiopian wolf	Africa	AF028144	[8]
3	<i>C. familiaris</i>	Dog		U12829	[41]
4	<i>Cuon alpinus</i>	Dhole	Asia	AF028137	[8]
5	<i>Arelocynus microtis</i>	Small eared dog	Asia (Japan)	AF028135	[8]
6	<i>C. aureus</i>	Golden jackal	Old world	AF028138	[8]
7	<i>C. mesomelas elongae</i>	Black-backed jackal	Eastern Africa	AF028142	[8]
8	<i>C. adustus</i>	Side-striped jackal	Eastern Africa	AF028136	[8]
9	<i>C. latrans</i>	Coyote	North America	AF028140	[8]
10	<i>Pseudalopex sechurae</i>	Desert fox	NW S. America	AF028154	[8]
<i>Complete Cytochrome-b gene reference sequences</i>					
1	<i>C. familiaris</i>	Dog	?		[38]
2	<i>Vulpes vulpes</i>	Fox	?		[38]
<i>16S rDNA reference sequence</i>					
1	<i>C. familiaris</i>		Asia (Japan)	AB048590	

Table 3:

Reference mitochondrial control-region sequences obtained from EMBL database and used in the phylogenetic analysis

Sample Code	Species	Subspecies	Country/Continent		EMBL	D-loop	Haplotype ^a	Reference	D-loop region sequenced
					AccNo	(bp)			
1-lu1	<i>C. lupus</i>	lupus	Yugoslavia	Europe	AB007372	672		[9]	Complete 5'-CR-I
2-lu2	<i>C. lupus</i>	lupus	Yugoslavia	Europe	AB007373	672		[9]	Complete 5'-CR-I
3-p1	<i>C. lupus</i>	pallipes	Afganisathian	Asia	AB007374	672		[9]	Complete 5'-CR-I
4-ch1	<i>C. lupus</i>	chanco	Mongol, China	Asia	AB007375	673		[9]	Complete 5'-CR-I
5-ch2	<i>C. lupus</i>	chanco	Mongol, China	Asia	AB007376	673		[9]	Complete 5'-CR-I
6-ch3	<i>C. lupus</i>	chanco	Mongol, China	Asia	AB007377	673		[9]	Complete 5'-CR-I
7-ch4	<i>C. lupus</i>	chanco	Mongol, China	Asia	AB007378	673		[9]	Complete 5'-CR-I
8-ch5	<i>C. lupus</i>	chanco	Mongol, China	Asia	AB007379	673		[9]	Complete 5'-CR-I
9-Wo11	<i>C. lupus</i>	crassodon	Canada	N. America	AF098125	707	Wo-11	[45]*	Complete 5'-CR-I
10-Wo10	<i>C. lupus</i>	crassodon	Canada	N. America	AF098124	707	Wo-10	[45]*	Complete 5'-CR-I
11-Wo9	<i>C. lupus</i>	crassodon	Canada	N. America	AF098123	707	Wo-9	[45]*	Complete 5'-CR-I
12-Wo4	<i>C. lupus</i>	columbianus	Canada	N. America	AF098118	707	Wo-4	[45]*	Complete 5'-CR-I
13-Wo5	<i>C. lupus</i>	columbianus	Canada	N. America	AF098119	707	Wo-5	[45]*	Complete 5'-CR-I
14-Wo6	<i>C. lupus</i>	columbianus	Canada	N. America	AF098120	707	Wo-6	[45]*	Complete 5'-CR-I
15-Wo7	<i>C. lupus</i>	columbianus	Canada	N. America	AF098121	707	Wo-7	[45]*	Complete 5'-CR-I
16-Wo8	<i>C. lupus</i>	columbianus	Canada	N. America	AF098122	707	Wo-8	[45]*	Complete 5'-CR-I
17-Wo2	<i>C. lupus</i>	lupus	Canada	N. America	AF098116	707	Wo-2	[45]*	Complete 5'-CR-I
18-Wo3	<i>C. lupus</i>	lupus	Canada	N. America	AF098117	707	Wo-3	[45]*	Complete 5'-CR-I
19-Wo1	<i>C. lupus</i>	tundrarum	Canada	N. America	AF098115	707	Wo-1	[45]*	Complete 5'-CR-I
20-W27	<i>C. lupus</i>		China	Asia	AF008135	713	W-27	[18]	Complete 5'-CR-I
21-W 6	<i>C. lupus</i>		Romania	Europe	AF008136	712	W- 6	[18]	Complete 5'-CR-I
22-W 1	<i>C. lupus</i>		Spain	Europe	AF008137	712	W- 1	[18]	Complete 5'-CR-I
23-W 4	<i>C. lupus</i>		France/Italy	Europe	AF008138	712	W- 4	[18]	Complete 5'-CR-I
24-W 5	<i>C. lupus</i>		Greece/Romania	Europe	AF008139	712	W- 5	[18]	Complete 5'-CR-I
25-W12	<i>C. lupus</i>		India	Asia	AF008140	712	W-12	[18]	Complete 5'-CR-I
26-W15	<i>C. lupus</i>		SaudiArabia	Asia	AF008141	713	W-15	[18]	Complete 5'-CR-I
27-W14	<i>C. lupus</i>		SaudiArabia/China	Asia	AF008142	712	W-14	[18]	Complete 5'-CR-I
28-w 1	<i>C. lupus</i>		Italy-East European	Europe	AF115687	542	w- 1	[42]	Most of 5'-CR-I
29-w 2	<i>C. lupus</i>		Italy-East European	Europe	AF115688	542	w- 2	[42]	Most of 5'-CR-I
30-w 3	<i>C. lupus</i>		Italy-East European	Europe	AF115689	542	w- 3	[42]	Most of 5'-CR-I
31-w 4	<i>C. lupus</i>		Italy-East European	Europe	AF115690	542	w- 4	[42]	Most of 5'-CR-I
32-w 5	<i>C. lupus</i>		Italy-East European	Europe	AF115691	542	w- 5	[42]	Most of 5'-CR-I
33-w 7	<i>C. lupus</i>		Italy-East European	Europe	AF115692	542	w- 7	[42]	Most of 5'-CR-I
34-w 8	<i>C. lupus</i>		Italy-East European	Europe	AF115693	542	w- 8	[42]	Most of 5'-CR-I
35-w 9	<i>C. lupus</i>		Italy-East European	Europe	AF115694	542	w- 9	[42]	Most of 5'-CR-I
36-w 10	<i>C. lupus</i>		Italy-East European	Europe	AF115695	542	w- 10	[42]	Most of 5'-CR-I
37-w 11	<i>C. lupus</i>		Italy-East European	Europe	AF115696	542	w- 11	[42]	Most of 5'-CR-I
38-w 12	<i>C. lupus</i>		Italy-East European	Europe	AF115697	542	w- 12	[42]	Most of 5'-CR-I
39-w 13	<i>C. lupus</i>		Italy-East European	Europe	AF115698	542	w- 13	[42]	Most of 5'-CR-I
40-w 14	<i>C. lupus</i>		Italy-East European	Europe	AF115699	542	w- 14	[42]	Most of 5'-CR-I
41-w 15	<i>C. lupus</i>		Italy-East European	Europe	AF115700	542	w- 15	[42]	Most of 5'-CR-I
42-w 16	<i>C. lupus</i>		Italy-East European	Europe	AF115701	542	w- 16	[42]	Most of 5'-CR-I
43-w 19	<i>C. lupus</i>		Italy-East European	Europe	AF115702	542	w- 19	[42]	Most of 5'-CR-I
44-w 20	<i>C. lupus</i>		Italy-East European	Europe	AF115703	542	w- 20	[42]	Most of 5'-CR-I
45-Do9	<i>C. familiaris</i>		Dog			707	Do-9		Complete 5'-CR-I
46-DogABO	<i>C. familiaris</i>		Dog		AB048590	737			Complete 5'-CR-I
47	<i>Vulpus vulpus</i>		Red fox	Japan	D-83639	666		[43]	Complete 5'-CR-I
48	<i>Procyonoides</i>		Raccoon dog	Japan	D-83614	667		[43]	Complete 5'-CR-I
49-W1	<i>C. lupus</i>		France/Italy/Spain	Europe	AF338803	342	W-1	[46]*	Part of 5'-CR-I
50-W2	<i>C. lupus</i>		France/Italy/Spain	Europe	AF338804	342	W-2	[46]*	Part of 5'-CR-I
51-W3	<i>C. lupus</i>		France/Italy/Spain	Europe	AF338805	342	W-3	[46]*	Part of 5'-CR-I
52-W4	<i>C. lupus</i>		Croatia	Europe	AF338806	342	W-4	[46]*	Part of 5'-CR-I
53-W5	<i>C. lupus</i>		Croatia	Europe	AF338807	342	W-5	[46]*	Part of 5'-CR-I
54-W6	<i>C. lupus</i>		Portugal	Europe	AF338808	342	W-6	[46]*	Part of 5'-CR-I
55-W7	<i>C. lupus</i>		Spain	Europe	AF338809	342	W-7	[46]*	Part of 5'-CR-I
56-W8	<i>C. lupus</i>		Romania	Europe	AF338810	342	W-8	[46]*	Part of 5'-CR-I
57-W9	<i>C. lupus</i>		Poland	Europe	AF338811	342	W-9	[46]*	Part of 5'-CR-I
58-W10	<i>C. lupus</i>		Spain	Europe	AF338812	342	W-10	[46]*	Part of 5'-CR-I
59-W2	<i>C. lupus</i>		Polland-Belarus	Europe	AF344300	257	W-2	[47]*	Part of 5'-CR-I
60-W3	<i>C. lupus</i>		Polland-Belarus	Europe	AF344301	257	W-3	[47]*	Part of 5'-CR-I
61-W4	<i>C. lupus</i>		Polland-Belarus	Europe	AF344302	257	W-4	[47]*	Part of 5'-CR-I
62-p3	<i>C. lupus</i>		USA	N. America	AF020699	402		[44]	Part of 5'-CR-I
63-W27	<i>C. lupus</i>		China	Asia	AF008160	305	W-27	[18]	3'-CR-II
64-W6a	<i>C. lupus</i>		Romania/Russia	Eurasia	AF008161	304	W-6a	[18]	3'-CR-II
65-W 1	<i>C. lupus</i>		Portugal/Spain	Europe	AF008162	304	W- 1	[18]	3'-CR-II
66-W 4	<i>C. lupus</i>		France/Italy	Europe	AF008163	300	W- 4	[18]	3'-CR-II
67-W 5	<i>C. lupus</i>		Greece/Romania	Europe	AF008164	272	W- 5	[18]	3'-CR-II
68-W12	<i>C. lupus</i>		India	Asia	AF008165	304	W-12	[18]	3'-CR-II
69-W15	<i>C. lupus</i>		SausiArabia	Asia	AF008166	304	W-15	[18]	3'-CR-II
70-W14	<i>C. lupus</i>		SausiArabia/China	Asia	AF008167	304	W-14	[18]	3'-CR-II
71-w2	<i>C. lupus</i>		Croatia/Portugal/Sweder	Europe	AF005296	257	w-2	[18]	Part of 5'-CR-I
72-w3	<i>C. lupus</i>		Poland/Spain	Europe	AF005297	257	w-3	[18]	Part of 5'-CR-I
73-w7	<i>C. lupus</i>		Bulgaria/SaudiArabia	Eurasia	AF005298	257	w-7	[18]	Part of 5'-CR-I
74-w8	<i>C. lupus</i>		Greece	Europe	AF005299	257	w-8	[18]	Part of 5'-CR-I
75-w9	<i>C. lupus</i>		Greece	Europe	AF005300	257	w-9	[18]	Part of 5'-CR-I
76-w10	<i>C. lupus</i>		Estonia/Finland/Russia/Spain	Europe	AF005301	257	w-10	[18]	Part of 5'-CR-I
77-w11	<i>C. lupus</i>		Israel	Asia	AF005302	257	w-11	[18]	Part of 5'-CR-I
78-w16	<i>C. lupus</i>		Iran	Asia	AF005303	257	w-16	[18]	Part of 5'-CR-I
79-w17	<i>C. lupus</i>		Iran	Asia	AF005304	257	w-17	[18]	Part of 5'-CR-I
80-w13	<i>C. lupus</i>		SaudiArabia	Asia	AF005305	257	w-13	[18]	Part of 5'-CR-I
81-w18	<i>C. lupus</i>		Afganisathian	Asia	AF005306	257	w-18	[18]	Part of 5'-CR-I
82-w19	<i>C. lupus</i>		China	Asia	AF005307	257	w-19	[18]	Part of 5'-CR-I
83-w20	<i>C. lupus</i>		Alaska	N. America	AF005308	257	w-20	[18]	Part of 5'-CR-I
84-w22	<i>C. lupus</i>		Alaska/Alberta/Montana/N-W territories	N-W territories	AF005309	257	w-22	[18]	Part of 5'-CR-I
85-w21	<i>C. lupus</i>		Yukon	Europe	AF005310	258	w-21	[18]	Part of 5'-CR-I
86-w23	<i>C. lupus</i>		Yukon	Europe	AF005311	258	w-23	[18]	Part of 5'-CR-I
87-w24	<i>C. lupus</i>		Yukon	Europe	AF005312	258	w-24	[18]	Part of 5'-CR-I
88-w25	<i>C. lupus</i>		Mexico	N. America	AF005313	258	w-25	[18]	Part of 5'-CR-I
89-w26	<i>C. lupus</i>		Russia	Eurasia	AF005314	257	w-26	[18]	Part of 5'-CR-I
90-lu3	<i>C. lupus</i>		Russia	Eurasia	AF202894	360		[48]*	Repetitive region
a :	Haplotype code given in the reference by the respective authors								
* :	Unpublished reference; details taken from corresponding EMBL entries								

Table 4 : Summary statistics of DNA polymorphisms analyzed in the study

Mitochondrial domain used for taxonomic comparisons									
Sequence Parameters		Complete Cyto-B gene	Partial Cyto-B gene	16S rRNA gene	D-loop (CR-I) 233bp	D-loop (CR-I) 545bp	D-loop (CR-I) 651bp	D-loop (CR-I) 675bp	D-loop Region-II
Sequences compared	5	35	35	27	80	50	33	24	11
Sequence size compared (bp)	1140	332	332	554	233	545	651	675	294
Sequence size without gaps/'N' (bp)	1132	280	280	554	210	505	608	656	252
Polymorphic nucleotides	232	77	77	32	35	72	73	75	22
Singleton polymorphic nucleotides	178	34	34	.	14	33	39	44	16
Informative nucleotides	54	43	43	32	21	39	34	31	6
Base composition (average)									
A :	28.9%	28.2%	28.2%	26.7%	30.5%	26.9%	27.3%	27.2%	33.6%
C :	27.6%	23.0%	23.0%	20.3%	24.0%	27.1%	26.5%	25.6%	21.4%
G :	14.2%	16.6%	16.6%	21.7%	12.3%	14.9%	15.3%	16.3%	11.3%
T :	29.4%	32.2%	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	8.70	11.65	19.57	9.89	7.60	6.96	14.20
Minimum	2.80	0.00	0.00	5.70	0.00	0.00	0.00	0.00	0.00
Maximum	18.75	26.00	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. of haplotypes	5	18	18	6	36	39	28	10	10
Number of Average substitutions	106.0±66.0	12.35±9.57	12.35±9.57	7.21±0.93	4.97±0.15	9.65±0.83	11.15±1.7	5.53±1.5	2.19±0.03
Nucleotide diversity (%)	9.39±4.87	4.37±1.15	4.37±1.15	1.30±0.50	2.37±0.04	1.91±0.04	1.83±0.03	1.83±0.03	1.83±0.03
Ks (%)	57.2±48.6	19.2±16.5	19.2±16.5						
Ka (%)	1.49±0.79	0.77±0.80	0.77±0.80						
HW vs. GW and wolf/dog popuⁿ from elsewhere									
<i>Number of Average substitutions</i>									
<i>Minimum-maximum substitutions</i>					14.1±1.6	20.7±1.9	21.3±1.9	7.1±1.6	
<i>Median (Mode) substitutions</i>					11--16	17--24	18--25	5--9	
					14 (16)	20 (19)	21 (21)	7 (7)	
GW vs. HW and wolf/dog popuⁿ from elsewhere									
<i>Number of Average substitutions</i>									
<i>Minimum-maximum substitutions</i>					13.4±1.3	17.3±1.6	18.0±2.0	5.1±1.6	
<i>Median (Mode) substitutions</i>					10--17	14--21	14--21	4--9	
					13 (13)	17 (17)	18 (19)	5 (4)	
Between different wolf/dog popuⁿ excluding HW/GW from India									
<i>Number of Average substitutions</i>									
<i>Minimum-maximum substitutions</i>					4.5±2.23	7.4±3.2	7.6±3.4	3.1±1.3	
<i>Median (Mode) substitutions</i>					0--10	0--15	0--15	0--5	
					4 (4)	8 (7)	8 (7)	3 (3)	

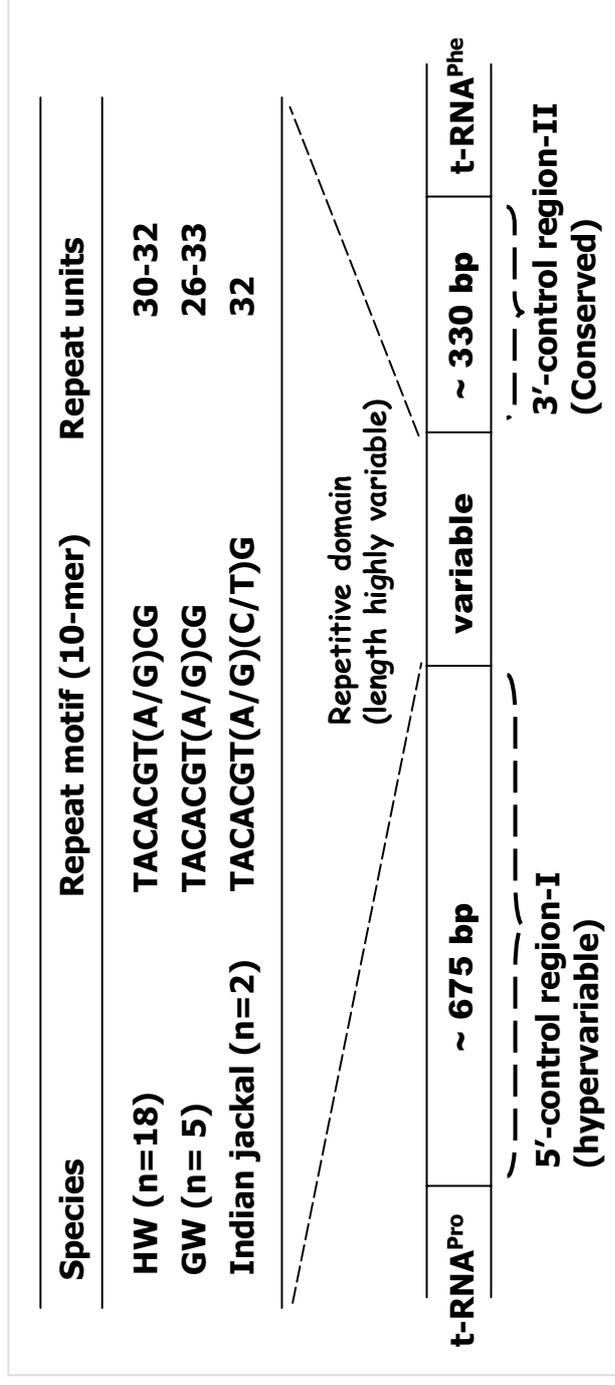


Figure 1: Mitochondrial d-loop control region organization. Upper panel shows the sequence polymorphism seen in the canids from India (present study).

Figure 2a:

ML tree based on Cyto-B gene polymorphism showing relationship of HW with other canids from India and elsewhere; Distance: Gamma corrected K 2-P.

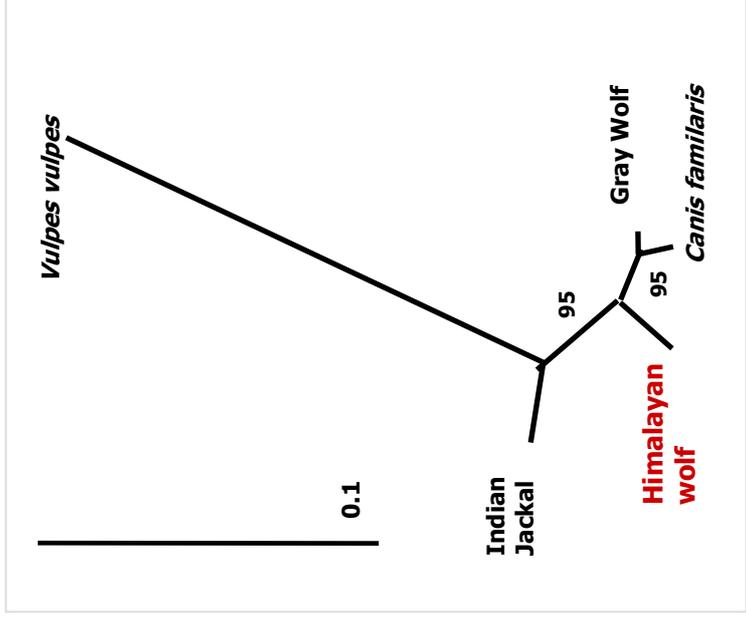
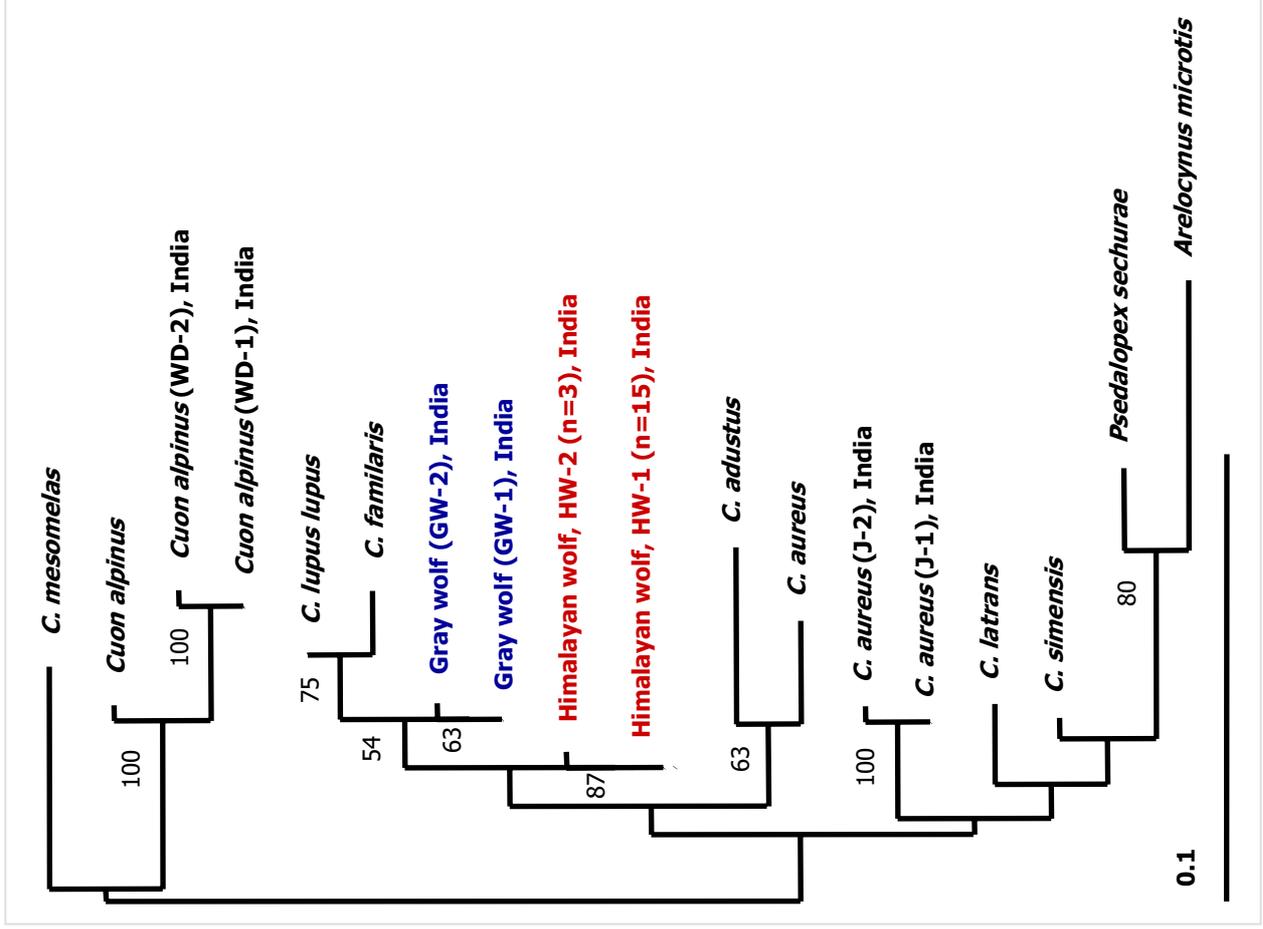


Figure 2b:

ML tree based on complete Cyto-B gene polymorphism showing relationship of HW with jackal and grey wolf from India and 2 reference canids from elsewhere; Distance: Gamma corrected K 2-P.

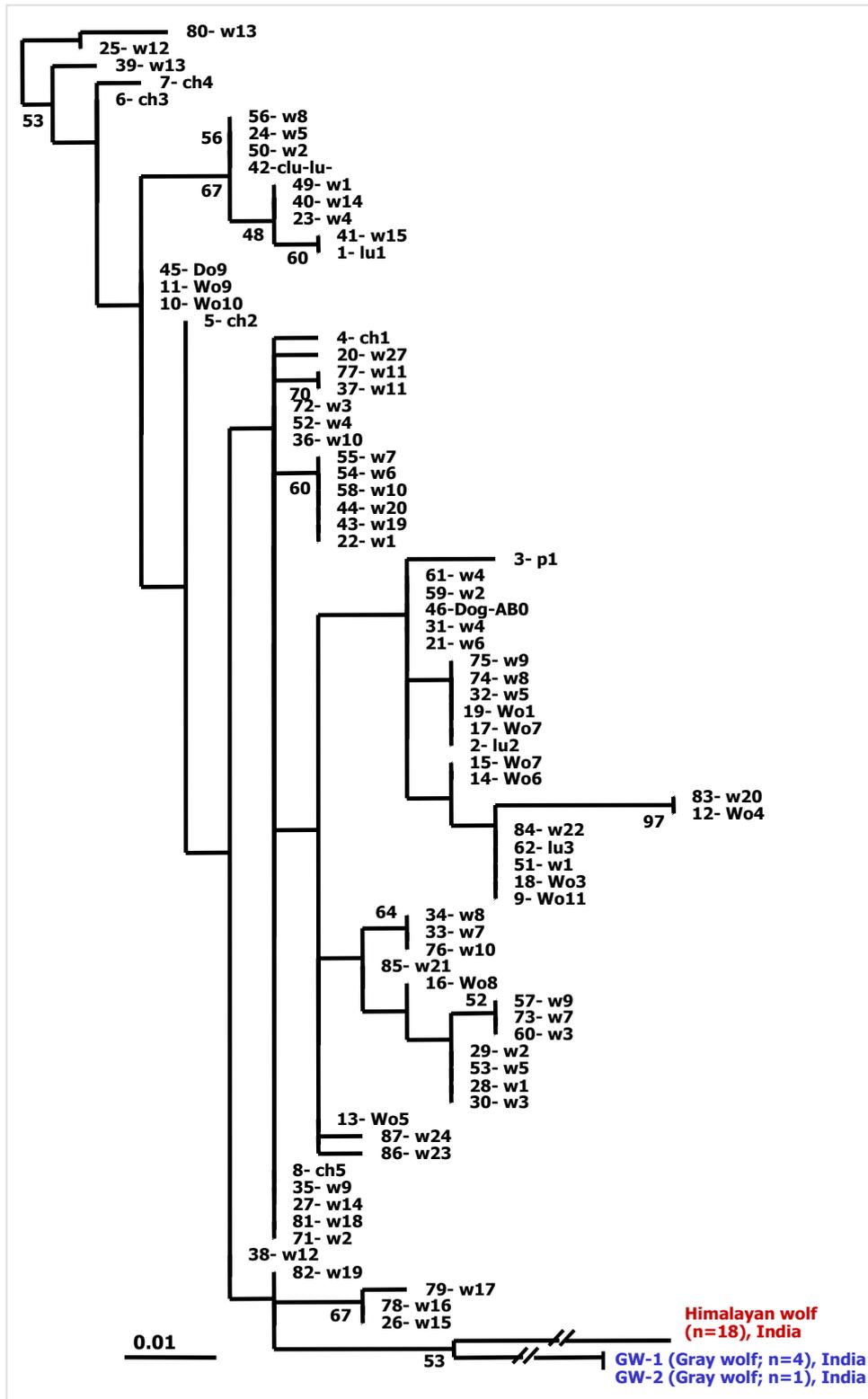


Figure 3: ML tree based on 233 bp of d-loop region-I polymorphism showing relationship of HW with 75 other wolf populations from India and elsewhere and 2 dog species; Distance: Gamma corrected K 2-P. Reference taxa are as per codes in Table-3.



Figure 4:

NJ tree based on ~540 bp of d-loop control region-I polymorphism showing relationship of HW with 45 other wolf populations from India and elsewhere and 2 dog entries; Indian jackal is used as outgroup; Distance: Corrected K 2-P. Reference taxa are as per codes in Table- 3.

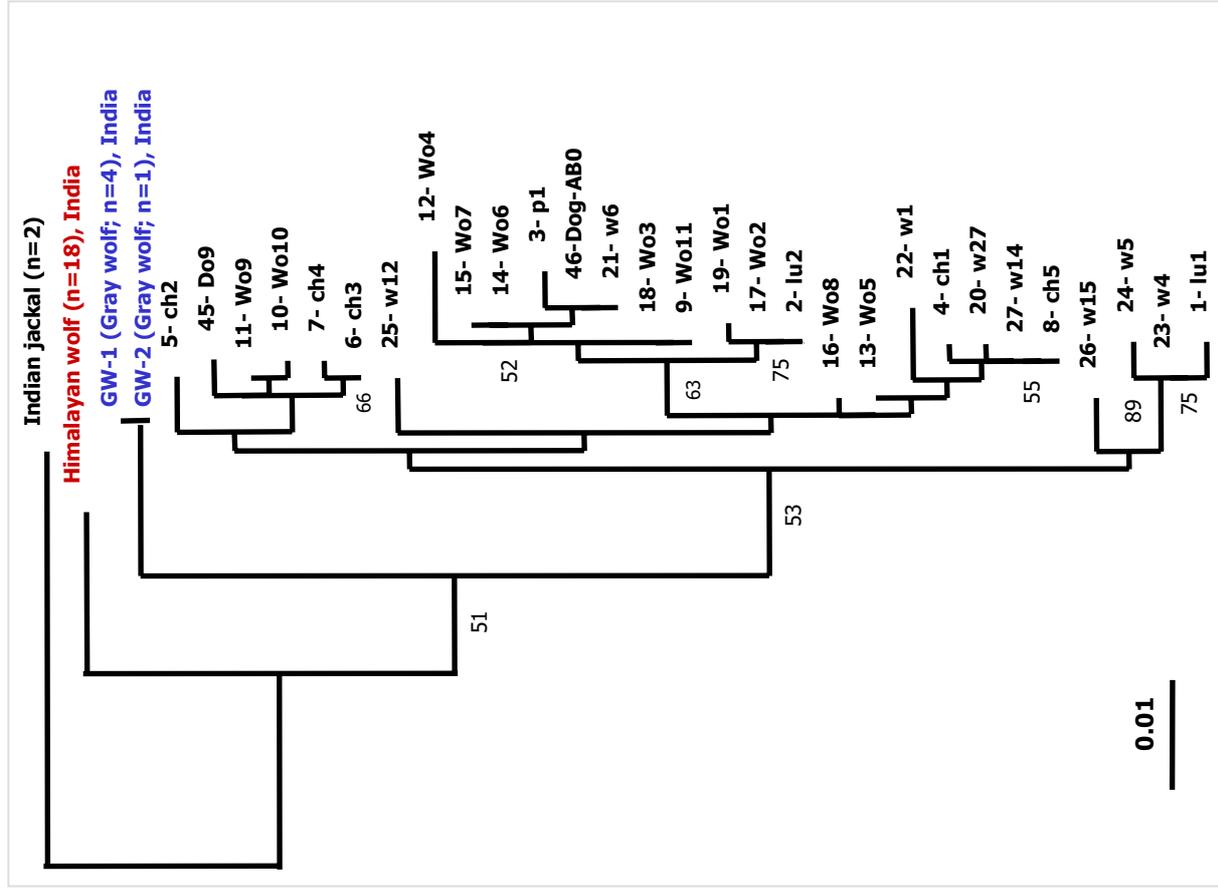


Figure 5:

ML tree based on ~655 bp of d-loop control region-I polymorphism showing relationship of HW with 28 other wolf populations from India and elsewhere and 2 dog entries; Indian jackal is used as outgroup; Distance: Gamma corrected K 2-P. Reference taxa are as per codes in Table-3.

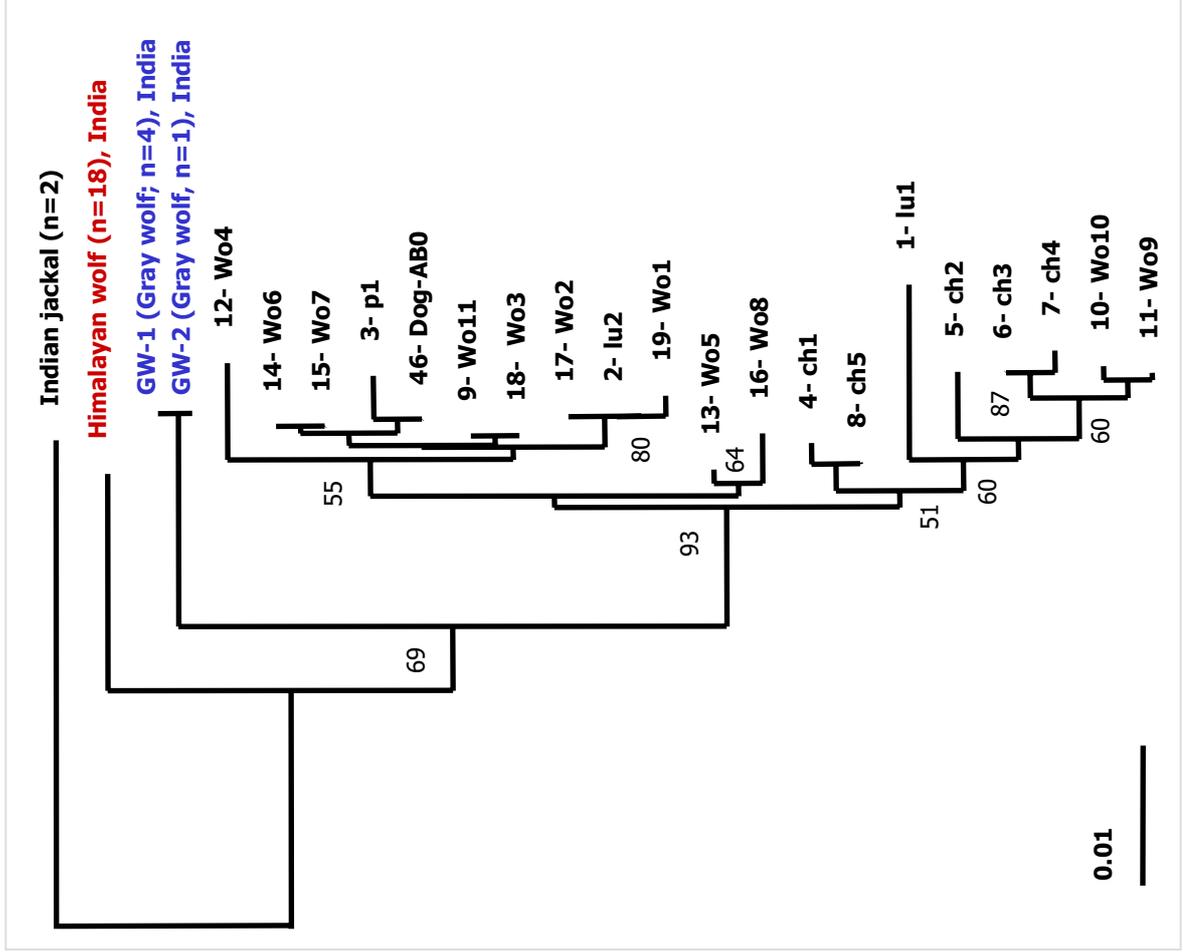


Figure 6:

NJ tree based on 675bp complete d-loop control region-I polymorphism showing relationships of HW with 20 other wolf populations from India and elsewhere and 1 dog species; Indian jackal is used as outgroup; Distance: corrected K 2-P. Reference taxa are as per codes in Table-3.