Phylogeography, genetic structure, and diversity in the dhole (*Cuon alpinus*)

A. IYENGAR,*† V. N. BABU,‡ S. HEDGES,§ A. B. VENKATARAMAN,¶ N. MACLEAN† and
P. A. MORIN* **

Laboratory for Conservation Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, †School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, United Kingdom, ‡Centre for Ecological Sciences, Indian Institute of Science, Bangalore, India, §Wildlife Conservation Society — International Programs, 2300 Southern Boulevard, Bronx, New York, NY 10460, USA, ¶Asian Elephant Research and Conservation Centre, Centre for Ecological Sciences, Indian Institute of Science, Bangalore, India

Abstract

The Asiatic wild dog or dhole was once very widely distributed across Asia but now has a very fragmented range. In this first genetic study of this little-known species, we obtained information on genetic diversity, phylogeography, and social structure using both mitochondrial control region sequencing and microsatellite genotyping of noninvasive faecal samples from wild populations, as well as from museum and captive samples. A pattern largely consistent with isolation by distance across the Asian mainland was observed, with no clear subspecies distinctions. However, two major phylogeographical groupings were found across the mainland, one extending from South, Central, and North India (south of the Ganges) into Myanmar, and the other extending from India north of the Ganges into northeastern India, Myanmar, Thailand and the Malaysian Peninsula. We propose a scenario involving glaciation events that could explain this pattern. The origin of the dhole populations in Sumatra and Java is enigmatic and requires further study. Very low levels of genetic diversity were observed among wild dholes from Baluran National Park in Java, Indonesia, but in contrast, high levels were observed in Mudumalai Wildlife Sanctuary in South India.

Keywords: Asia, Asiatic wild dog, Canidae, conservation genetics, microsatellite, mitochondrial DNA

Received 1 December 2004; revision accepted 14 March 2005

Introduction

The dhole (*Cuon alpinus*), also called the Asiatic wild dog, red dog, red wolf, or whistling dog, was widespread across North America, Europe and Asia during the Pleistocene. The species' range, however, became restricted to Asia after the late Pleistocene mass extinctions *c.* 12 000–18 000 BP, when it became extinct across North America and Europe, along with several other large species such as mammoths and dire wolves (Thenius 1954; Dundas 1999). Dholes are adapted to life in very different environments ranging from mountainous alpine regions in Russia to open steppes in

Correspondence: Arati Iyengar, Fax: + 44 2380594459; E-mail: a.iyengar@soton.ac.uk.
**Present address: Protected Resources Division, Southwest Fisheries Science Center (SWFSC), 8604 La Jolla Shores Drive, La Jolla CA 92037, USA.
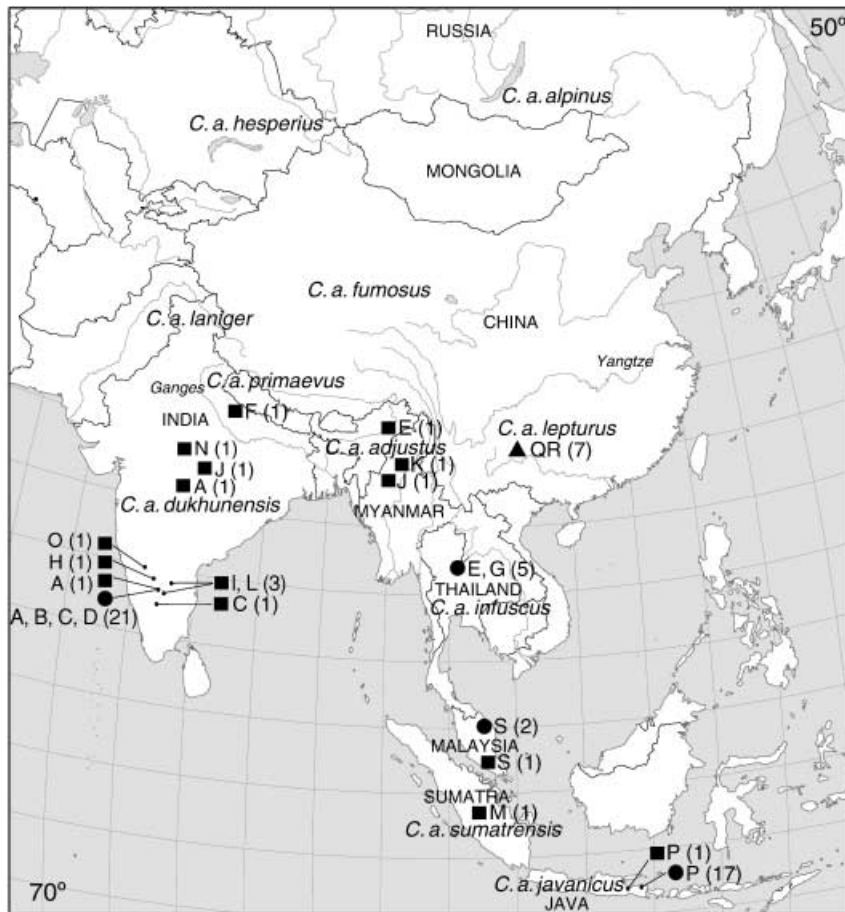
Tibet to scrubland and dense forests in South Asia. Unfortunately in recent times, due to habitat fragmentation, collapse of prey populations, and increased disturbance and persecution by humans, the range of the dhole has been much reduced. The dhole's range was known to have encompassed most of South, East, and Southeast Asia (Fig. 1). The present range of the species is imperfectly known, but dholes are thought to occur in Bhutan, Cambodia, China (including Tibet), India, Indonesia (Sumatra and Java), Lao PDR, the Malaysian Peninsula, Myanmar, Thailand, and Vietnam. Dholes are also thought to likely still remain in Bangladesh and Nepal, and there may be relic populations in the Central Asian states of the former USSR, Korea, Mongolia, Pakistan, and Russia (Durbin *et al*. 2004b). The species, given protected status in most countries only in the 1970s, is at high risk of extinction in many parts of its range and is listed as endangered in the

Fig. 1 Map of Asia, showing approximate distribution ranges of the 11 putative subspecies of dholes, with sampling locations, sample numbers, and observed haplotypes. ●—samples from extant wild populations, ■—museum samples, ▲—captive samples. Original template map obtained from the BYU Geography Department (http://www.geog.byu.edu/maps/outline/outline.html)

2004 International Union for Conservation of Nature and Natural Resources (IUCN) Red List of Threatened Species and the IUCN Canid Action Plan (Durbin *et al.* 2004a, b).

Molecular phylogenetic analyses have placed the dhole as an early divergent within the wolf-like canid group consisting of the grey wolf (*Canis lupus*), the coyote (*Canis latrans*), the Ethiopian wolf (*Canis simensis*), some jackals (*Canis aureus*, *Canis mesomelas*) and the dhole (Wayne *et al.* 1997). The dhole is the sole species in the genus *Cuon*. Eleven subspecies are sometimes recognized (see approximate distribution ranges in Fig. 1). Differences in pelage length and colour have been noted among these putative subspecies, with a general increase in pelage length with latitude and a much darker and brighter pelage in the high-rainfall areas of Southeast Asia (Pocock 1936; Cohen 1978; Durbin *et al.* 2004b). Although dholes have been well studied in terms of their behavioural ecology within India (e.g. Cohen *et al.* 1978; Johnsingh 1982, 1992; Fox 1984; Karanth & Sunquist 1995, 2000; Venkataraman *et al.* 1995; Venkataraman 1998), aside from a few brief surveys (Johnsingh 1985; Stewart 1993, 1994), little is known about the status, distribution, and biology of this species in the rest of its extensive range.

Studies in India have revealed that similar to other social canids such as grey wolves and painted hunting dogs (*Lycaon pictus*), dholes are highly social animals with a rigid structure of fixed dominance hierarchies. They also exhibit a high level of cooperative behaviour and live in packs of 3–20 individuals. Circumstantial evidence suggests that packs have a sex ratio biased towards males as a consequence of higher female dispersal (Johnsingh 1982; Venkataraman 1998). In contrast to grey wolves from North America and Europe, and painted hunting dogs from Africa, for which species extensive genetic studies have been conducted (e.g. Girman *et al.* 1997; Vilà *et al.* 1999; Girman *et al.* 2001; Lucchini *et al.* 2002, 2004), this is the first genetic study of the dhole. Using both microsatellite markers and mitochondrial DNA sequencing, we examine the extent and phylogenetic patterns of DNA variation across a number of regions in Asia. Since dholes are highly elusive carnivores and, moreover, a protected species, we relied on noninvasive faecal samples from extant wild populations. Faecal samples are widely used in the field of conservation genetics for these reasons (e.g. Adams *et al.* 2003). We also obtained samples from dhole skulls collected over the period 1846–1939 and held in the Natural

History Museum in London. Museum samples have proved to be very useful in studies involving endangered species with fragmented and restricted distributions, since they allow a reconstruction of recent evolutionary history across the historical range of the species (e.g. Girman *et al*. 2001; Godoy *et al*. 2004).

Materials and methods

Samples

The two main collections of faecal samples were made in protected areas in East Java (Indonesia) and South India (Fig. 1). We are confident that these faecal samples originated from dholes and not domestic dogs based on the very low chance of encountering domestic dog scats in these areas, the appearance of the scats, and the congruence in results obtained with DNA sequencing. Fifty-one fresh (< 12 h old) faecal samples were collected in 1998 from Baluran National Park (BNP), an area 250 km² situated in East Java, by placing a small amount into 2-mL tubes containing 20% dimethyl sulfoxide (DMSO), 0.25 M ethylenediamene tetra-acetic acid (EDTA), 100 mM Tris, pH 7.5, saturated with NaCl DET buffer — Frantzen *et al*. 1998. A much larger set of 143 samples was collected during 2001 from Mudumalai Wildlife Sanctuary (MWS) in South India. Samples were placed into 50-mL tubes containing *c*. 30 g silica gel (type III indicating, Sigma) with a small piece of filter paper separating the faecal material from the silica gel (Wasser *et al*. 1997). The tubes were then held at ambient temperature for several weeks prior to being stored at 4 °C. MWS is an area of 326 km² within the much larger Nilgiri Biosphere Reserve where the home ranges of five packs of dholes have been estimated by authors VNB and ABV. Minimum convex polygons (MCPs) were constructed with the Animal Movement extension for ARCVIEW 3.2® using GPS locations of pack sightings and GPS locations of fresh scats attributed to a pack based on the number of scats found and similarity to the number of individuals within the pack. Between 30 and 60 location points were used for the estimation of each home range (except in the case of the Gamehut pack where only 15 such points were available). Site fidelity tests and mitigation of outlier effects were both carried out within ARCVIEW 3.2®. Although individual dholes can only be identified with difficulty, pack identification was possible because overlap between home ranges was not always extensive and pack sizes were different from one another. Faecal sample collection was carried out mainly from communal latrines (Johnsingh 1982) within estimated home ranges (Fig. 2). Collection was frequently closely preceded by or followed by actual visual sighting of the pack, allowing us to allocate samples to packs with a good level of certainty, even when collections were made in regions of overlap. Repeat collections were carried out within home ranges in order to attempt to sample as many individuals as possible (see Fig. 2). We also obtained a few faecal samples from Phu Khieo Wildlife Sanctuary (PKWS) in northeastern Thailand and Taman Negara National Park (TNNP) in Peninsular Malaysia. All sampling sites are shown in Fig. 1. In addition, we obtained blood samples from captive dholes from three zoos within Germany and scrapings of internal tissue from several dhole skulls held at the Natural History Museum in London (see Table 3 for details).

DNA extractions

DNA was extracted from the faecal samples in a dedicated area using the QIAamp® DNA stool minikit (QIAGEN) according to the manufacturers instructions but with the following modifications: approximately 100 mg dried faecal material was separated from the bolus on a sterile Petri dish, placed in an Eppendorf tube with 1.8 mL ASL buffer and allowed to incubate at 37 °C for 12–24 h, and the final postextraction elution step was carried out for 30 min. A maximum of 15 samples were processed at one time with 1–2 negative controls. DNA from blood samples was extracted using a QIAamp® DNA blood mini kit following manufacturer's instructions. DNA from museum samples was extracted in a dedicated 'ancient DNA' lab, by placing the entire material (no more than *c*. 200 mg) into 5 mL of extraction buffer (0.45 M EDTA pH 8, 1% sarcosyl, 0.4 mg/mL proteinase K), followed by the procedure described in Vigilant *et al*. (2001).

Quantification of faecal DNA samples using qPCR

Quantitative polymerase chain reaction (qPCR) was used to quantify the amount of amplifiable genomic DNA in all faecal DNA samples, and to determine the number of replications required when genotyping in order to obtain high levels of accuracy. PCRs were carried out in dedicated areas with appropriate negative controls. Heterozygous genotypes were accepted once confirmed in two separate amplifications, while in the case of homozygous genotypes, repeats were carried out between four and seven times (for template amounts of 26–200 pg/reaction, as described in Morin *et al*. 2001).

Microsatellite analyses

A set of 13 microsatellite loci was used, consisting of both dinucleotide (CXX436, CXX466, AHT130, CXX608, CXX250, CXX279, CXX253, CXX434, CXX374) and tetranucleotide repeats (2140, 2010, 2001, 2096), all of which have been isolated in dogs (Ostrander *et al*. 1993, 1995; Holmes *et al*. 1995; Francisco *et al*. 1996). PCRs were carried out in a 15-µL volume containing 2 µL of DNA extract from faeces,
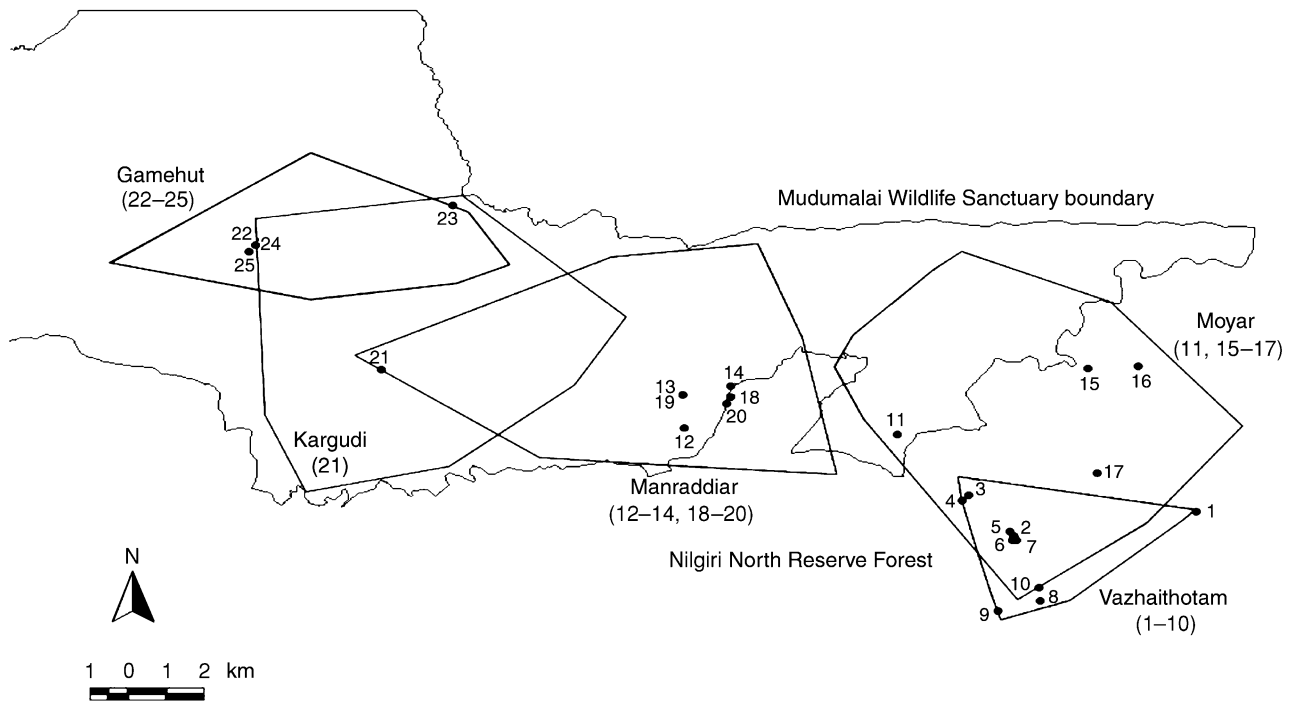
Fig. 2 Estimated home ranges of five packs of dholes within Mudumalai Wildlife Sanctuary in Tamilnadu, South India, with locations of faecal sample collections. South of the MWS boundary lies the Nilgiri North Reserve Forest. Numbers refer to individual collections. Repeat collections were carried out within most packs (10 within Vazhaithotam, 6 within Manraddiar, 4 within Moyar, 4 within Gamehut).

1× PCR buffer [ABgene, 75 mM Tris-HCl, pH 8.8, 20 mM $(NH_4)_2SO_4$, 0.01% (v/v) Tween 20®], 0.8–3.0 mM $MgCl_2$, 12 μg bovine serum albumin (BSA) (Roche), 200 μM each dNTP, 200 nM each primer and 0.4 U *Taq* DNA polymerase (ABgene). Amplification conditions consisted of initial denaturation for 4 min followed by 45 cycles at 94 °C for 30 s, 50–63 °C annealing temperature for 30 s, and 72 °C for 30 s followed by a final extension at 72 °C for 10 min. The 5′ end of the forward primer was fluorescently labelled and the products were separated using capillary electrophoresis (ABI Prism 310). Alleles were sized relative to an internal standard (HD400 with Rox label) and scored using GENESCAN 3.0 and GENOTYPER software (Applied Biosystems).

Mitochondrial DNA sequencing

A 650-bp fragment of the 5′ region of the dhole control region was initially amplified on DNA from fresh blood using a set of primers used in wolf (Randi *et al.* 2000). The product was then sequenced and dhole-specific primers designed to amplify 310 bp of the dhole 5′ control region sequence (DHDLOOPFOR 5′-CTACCATCAACCCCCAAAGC and DHDLOOPR310 5′-GCAAGGATTGATGGTTTCTCG). PCR amplification of the 310-bp product was carried out in a 30-μL volume containing 2 μL template DNA, 1× PCR buffer (ABgene), 2.5 mM $MgCl_2$, 24 μg BSA, 200 μM each dNTP,

200 nM each primer and 0.5 U *Taq* DNA polymerase (ABgene). Amplification conditions were as follows: initial denaturation for 4 min, followed by 35–40 cycles of 94 °C for 30 s, 57 °C for 30 s and 72 °C for 30 s followed by a final extension at 72 °C for 10 min. PCR products were electrophoresed on a 1% agarose gel and the 310-bp PCR product cut out and purified using QIAQuick gel purification columns (QIAGEN). Sequencing was carried out using an ABI3700 or an ABI377 (Applied Biosystems). A grey wolf faecal sample obtained from Leipzig Zoo was also used for PCR amplification of the same 310-bp region.

A 321-bp fragment of the mitochondrial cytochrome *b* region was also amplified using universal primers CB1 and CB2 described in Palumbi (1996) using amplification conditions as described above, except with an annealing temperature of 50 °C.

Since the occurrence of several nuclear copies of mitochondrial sequences including the control region has been reported in the canine genome (Ishiguro *et al.* 2002), and this appears to be a common problem across many taxa (e.g. Thalmann *et al.* 2004), a number of procedures were followed in order to exclude all sequences of nuclear origin from our analyses. First, dhole-specific primers were designed from a larger mitochondrial product from DNA obtained from fresh blood. Second, every sequence was put through BLAST in GenBank to ensure high homology

to existing canid mitochondrial D-loop sequences and no homology to existing nuclear copies of mitochondrial sequences. Finally, every variable haplotype was confirmed by cloning the PCR product into pDrive vector (PCR cloning kit, QIAGEN) and sequencing five to six clones in each case (12 clones each for haplotypes B and C) and comparing the sequences obtained to the one obtained with direct PCR sequencing. Although a number of single substitutions (singletons) were seen in some sequences and attributed to polymerase error, every substitution seen upon direct sequencing of the PCR product was also confirmed across every one of the clones. Only one museum sample, which was not satisfactorily confirmed, was removed from the analyses.

Sexing of dholes

Individuals were initially sexed using primers for the Y-chromosome-based SRY gene (*SRYfor* 5′-CTCGCGATC-AAAGGCGCAAGAT and *SRYrev* 5′-TTCGGCTTCTGT-AAGCATTTTC, Meyers-Wallen *et al.* 1995) generating a product of 104 bp, and as an internal PCR control, the *cmyc* proto-oncogene (*CMYC_E3_F1U1* 5′-GCCAGAGG-AGGAACGAGCT and *CMYC_E3_R1U1* 5′-GGGCCTTTT-CATTGTTTTCCA, Morin *et al.* 2001) generating a product of 81 bp. Since these gene sequences are highly conserved across mammals, in order to eliminate the possibility of obtaining incorrect results from male human handler contamination, we also designed primers within the *Canis* SRY sequence in regions of dissimilarity to the human SRY sequence so as to generate a product of 199 bp (*CanisSRYfor* 5′-ATGGCTCTAGAGAATCCCCA-3′ and *CanisSRYrev* 5′-GCAATTTGTGACTTTTCTGTGC-3′) and as the control, to the dhole *cmyc* gene (GenBank Accession no. AF519448) generating a product of 127 bp (*Dholecmycfor* 5′-GAGGAGGAACGAGCTG-3′ and *Dholecmycrev* 5′-TTGGACGGACAGGATGTAC). Both sets of primers were successfully used for sexing dholes in a volume of 15 μl containing 2 μl DNA extract, 1× PCR buffer (ABgene), 2.5 mM MgCl$_2$, 12 μg BSA, 200 μM each dNTP, 200 nM each of the four primers and 0.4 U *Taq* DNA polymerase (ABgene). Amplification conditions were as follows: initial denaturation for 4 min, followed by 35–40 cycles of 94 °C for 30 s, 57 °C for 30 s and 72 °C for 30 s and a final extension at 72 °C for 10 min. Every PCR was repeated three times in the case of apparent females to ensure accuracy, and the final results of the two assays were 100% concordant (data not shown).

Statistical analyses

Microsatellites. Polymorphism within populations, measured as the total number of alleles, mean number of alleles per locus (*A*), mean observed heterozygosity (H_O) and mean expected heterozygosity (H_E) was calculated using

GENEPOP (Raymond & Rousset 1995). Departures from Hardy–Weinberg equilibrium for each locus and population were assessed using the Markov chain method (Guo & Thompson 1992) implemented in GENEPOP and a sequential Bonferroni test was applied to the results. Deviation from linkage equilibrium for all pairwise locus combinations was tested using the exact probability test in GENEPOP.

The probability of identity P_{ID} (the probability that two individuals drawn at random from a population will have the same genotype at multiple loci) was calculated using the program API-CALC version 1.0 (Ayres & Overall 2004). This program estimates P_{ID} (or PI_{ave}, as denoted by these authors) while incorporating the effects of population substructure and inbreeding and also allowing for different proportions of close relatives to be tested.

Evidence for the occurrence of genetic bottlenecks was investigated using BOTTLENECK (Cornuet & Luikart 1996).

Pairwise estimates of the coefficient of relatedness (*r*) for all individuals were calculated using both Queller and Goodnight (r_{xyQG}) (Queller & Goodnight 1989) and Lynch and Ritland (r_{xyLR}) (Lynch & Ritland 1999) measures within the program IDENTIX using empirical allele frequencies (Belkhir *et al.* 2002) and used in comparisons for within and between pack relatedness levels. The problem of pseudo-replication exists in using every pairwise *r* value in such a manner with an associated artificial increase in the number of degrees of freedom (Prugnolle & de Meeus 2002). Consequently, in a comparison of a regression of *r* values with distance (*d*) values between packs, we used ranked *r* values as described in Knight *et al.* (1999), with *d* values allocated as follows: 0 within a pack, 1 when neighbouring packs, 2 when separated by one intervening pack, 3 when separated by two intervening packs and a maximum of 4 when separated by three packs (e.g. Vazhaithotam and Gamehut). To further evaluate the patterns of relatedness, a Monte Carlo resampling procedure implemented in IDENTIX was used with 1000 permutations in order to compare the observed distribution of *r* to that expected in non-kin-structured populations.

Mitochondrial DNA phylogenetic analyses. All sequences were checked by eye, edited, and then aligned using BIOEDIT version 5.0.9 (Hall 1999). The model of DNA substitution that best fitted the data was selected using MODELTEST, version 3.06 (Posada & Crandall 1998). The model HKY + G (ti/tv ratio = 28.6459, alpha = 0.0874) was selected by both the hierarchical likelihood ratio test and the Akaike information criterion (AIC).

Phylogenetic relationships were analysed using maximum-likelihood (ML) and maximum-parsimony (MP) approaches in PAUP, version 4.0b10 (Swofford 2002), using a heuristic search with the tree-bisection–reconnection (TBR) swapping algorithm. A homologous sequence obtained from a

Location	N	#loci	#A	A	R	#UA	H_O (± SE)	H_E (± SE)
MWS	21	13	47	3.81	4.0	37	0.54 (± 0.07)	0.52 (± 0.06)
BNP	17	13	18	1.38	0.54	8	0.14 (± 0.06)	0.15 (± 0.06)

Table 1 Genetic diversity in dholes from southern India and Java

MWS, Mudumalai Wildlife Sanctuary, South India; BNP, Baluran National Park, Java; *N*, number of individuals with unique multilocus genotypes; #loci, number of microsatellite loci typed; #*A*, total number of alleles across all loci; *A*, mean allele number per locus; *R*, range of allele size expansion in repeat motif number; #UA, number of unique alleles found; H_O, mean observed heterozygosity; H_E, mean expected heterozygosity (SE, standard error).

grey wolf faecal sample (found to be identical to GenBank sequence AF098117) was used as the outgroup for the analyses. Node support was assessed using 1000 bootstrap replicates. Mean HKY + G and uncorrected *P* distances between groups of haplotypes were measured in PAUP and MEGA version 2.1 (Kumar *et al.* 2001), respectively.

An analysis of molecular variance (AMOVA) was employed to assess the significance of genetic differentiation between various possible groupings using the program ARLEQUIN version 2.001 (Schneider *et al.* 2000). In an AMOVA, where correlations among haplotype distances at various hierarchical levels are used as *F* statistic analogues designated Φ statistics, Φ_{CT} measures the proportion of genetic variation among groups of populations (Excoffier *et al.* 1992). Since ARLEQUIN does not contain the HKY model, we used the Tamura–Nei model + G (alpha = 0.0874) in addition to total pairwise differences between haplotypes to estimate Φ_{CT} values.

A mismatch distribution of pairwise substitutional differences among haplotypes and a range of neutrality statistics which are capable of detecting the genetic traces of population growth, decline or stability were examined using DNASP (version 4.0, Rozas *et al.* 2003). Values for Fu's *F* statistic (F_S, which specifically tests for population growth and detects excesses of low-frequency alleles in an expanding population) and Fu and Li's *F** and *D** statistics were obtained and compared. Median-joining networks were estimated using the median-joining method described by Bandelt *et al.* (1999) and the software NETWORK 3.1.1.1 (http://www.fluxus-engineering.com), assigning equal weights to all variable sites and with default values for the epsilon parameter (epsilon = 0).

Results

Microsatellite analyses

Rates of amplification from faecal DNA samples. Only 35 of the 143 silica-dried faecal samples (*c.* 25%) from MWS were found to contain DNA in adequate quantities to allow accurate genotyping based on a qPCR assay. Of these 35 samples, nine multilocus genotypes were found to be from the same two individuals (six samples from one and three from another), necessitating the removal of repeat genotypes from the analyses. A further seven samples had to be eliminated from the final analyses due to inconsistent amplification across all loci. Thus, 21 individual genotypes were used for all subsequent analyses. In the case of the BNP samples, 30 of the 51 (*c.* 60%) DET buffer-preserved samples were found to contain adequate amounts of DNA by qPCR. Eleven of these were found to be repeat sampling of individuals, and two were found to show inconsistent amplification with resulting missing information. Upon deleting these individuals, 17 individuals with unique multilocus genotypes remained, which were used for all subsequent analyses.

Genetic diversity. The dholes from BNP were found to have fixed alleles at 8 of the 13 loci (all except CXX466, 2140, 2010, CXX608, CXX253), with extremely low overall allelic diversity and mean observed and expected heterozygosity values. The samples from MWS, in contrast, had high allelic diversity and high observed and expected heterozygosity levels (Table 1). One locus (CXX608) was fixed in the MWS samples, so only 12 loci were used in subsequent analyses. One locus (CXX250) was found to deviate significantly from Hardy–Weinberg equilibrium in this population after Bonferroni correction for multiple comparisons (*P* = 0.018). Since this locus demonstrated a heterozygote deficit which may be attributed to the existence of null alleles, we eliminated this locus from subsequent pack structure analyses. Tests for linkage disequilibrium after Bonferroni correction did not reveal any that were significant.

Dholes have a history of human persecution, and in India, dholes were widely shot or poisoned and bounties paid for carcasses right up to 1972, when they were given legal protection (Durbin *et al.* 2004b). Hence the possible existence of a genetic bottleneck in the MWS population was investigated. No such evidence was detected using a two-phase mutation (TPM) model (90% stepwise-mutation model) and the Wilcoxon's test (*P* = 0.38, normal L-shaped distribution).

The average probability of identity PI_{ave} was calculated in the MWS population by incorporating the F_{ST} value

Table 2 Estimates of pack size and results of molecular sexing

	Pack size estimates		
Sample	From field observations	# unique multilocus genotypes	Sex
BNP, Java		17	
MWS, South India		21	
Vazhaitotham	3	5	2M, 3F
Kargudi	4	1	1M
Gamehut	9	8	3M, 5F
Manraddiar	12	3	1M, 2F
Moyar	10	4	1M, 3F

(estimated at 0.0744) and allowing for a range of sib proportions within the population (Ayres & Overall 2004). Values of 3.6×10^{-8}, 8.5×10^{-6}, and 3.3×10^{-4} were obtained for sib proportions of 0, 0.5 and 1, respectively. Thus, the probability of two individuals within the MWS population having the same multilocus genotype is extremely small, since even the very conservative estimate of PI_{ave} obtained upon assuming that every individual was a full sibling (i.e. the $P_{(ID)sib}$), was very low, at 3.3×10^{-4}. So we can be confident that only unique individuals were used in subsequent analyses.

Pack size and structure in MWS. The number of unique multilocus genotypes detected among faecal samples from individual home ranges and comparison with pack number estimates from field observations (Table 2), reveals that despite repeat sampling within home ranges between 2 to 10 times (except Kargudi, with only one sampling), it was not possible to sample every individual within a pack. There was disagreement between the field-based estimate of pack size and the number of unique multilocus genotypes found in Vazhaitotam, with five unique multilocus genotypes, but only three dholes observed. In the Gamehut pack with a large number of individuals (nine) most of which (eight) have been sampled, a male bias was not evident even if it was assumed that the unsampled individual was a male.

Pairwise relatedness (r) values for both Queller and Goodnight (r_{xyQG}) and Lynch and Ritland (r_{xyLR}) estimators were calculated among individuals within MWS to look for differences in values within and between packs. Within pack values were significantly higher than between pack values in males [mean r_{xyQG} within packs = 0.43 (range = 0.12–0.64), mean r_{xyQG} between packs = −0.09 (range = −0.58–0.51), t-test, $P = 0.006$, d.f. = 5; mean r_{xyLR} within packs = 0.30 (range = 0.11–0.46), mean r_{xyLR} between packs = −0.14 (range = −0.49–0.27), t-test, $P = 0.004$]. A similar trend was also seen in females [mean r_{xyQG} within packs = 0.22 (range = −0.51–0.77), mean r_{xyQG} between packs = −0.06 (range = −0.60–0.72), t-test, $P = 0.007$, d.f. = 24; mean r_{xyLR} within packs = 0.12 (range = −0.42–0.74), mean r_{xyLR} between packs = −0.09 (range = −0.53–0.51), t-test, $P = 0.01$]. A comparison of the regressions of ranked r_{xyQG} and r_{xyLR} values and corresponding distance values (in pack units) for both sexes revealed a clear negative trend of mean r values with distance in both sexes (r_{xyLR} regression shown in Fig. 3a). In addition, the slopes of the two regression lines were found to be different ($P < 0.05$). The slope in the case of females is steeper than in the case of males, which suggests that females may be more likely to disperse close to their natal packs and conversely, males may be dispersing longer distances. A permutation test using r_{xyLR} values clearly supported the presence of kin structure within the population, with 99% of the permuted populations showing mean r scores higher than the mean observed value (Fig. 3b). The use of r_{xyQG} values however, did not show significant results with the permutation test (data not shown). The r_{xyLR} measure has been found to outperform the r_{xyQG} measure in a recent study (Russello & Amato 2004).

Mitochondrial DNA diversity and phylogeography. Approximately 50% and 80% of faecal DNA samples from MWS and BNP, respectively, was found to amplify the 310-bp control region product, of which 246 bp were successfully used for all phylogenetic analyses. Despite this higher success rate with mitochondrial DNA amplification, sequences from the 21 and 17 individuals with unique multilocus genotypes in MWS and BNP, respectively, were used to avoid repeated sampling. A total of 19 haplotypes were found among all samples (faecal, museum, and captive) with a total of 37 polymorphic sites (Table 3). Only one haplotype was detected among all the BNP samples (haplotype P), while four haplotypes were detected among the MWS samples (haplotype A in 18 out of 21 samples and haplotypes B, C and D in one individual each). These haplotypes were also found among museum samples from these regions: haplotype P was detected in a 1935 sample from East Java; haplotype A was detected in a 1937 sample from the MWS region, as well as in a 1911 sample from Central India; and haplotype C was detected in an 1888 museum sample from the Anamalai Hills region in South India (Fig. 1, Table 3).

Haplotypes were found to be shared between samples across vast distances on the Asian mainland, e.g. haplotype E, found in a 1922 museum sample from Assam in northeastern India was also found in two individuals from an extant population within PKWS in Thailand and haplotype J was found in museum samples from Central India, North India (south of the Ganges river) and North Myanmar (Fig. 1, Table 3).

Both maximum-likelihood (ML) and maximum-parsimony (MP) analyses produced identical trees. One tree was obtained with ML and a strict consensus of six most parsimonious trees was constructed. The bootstrap consensus
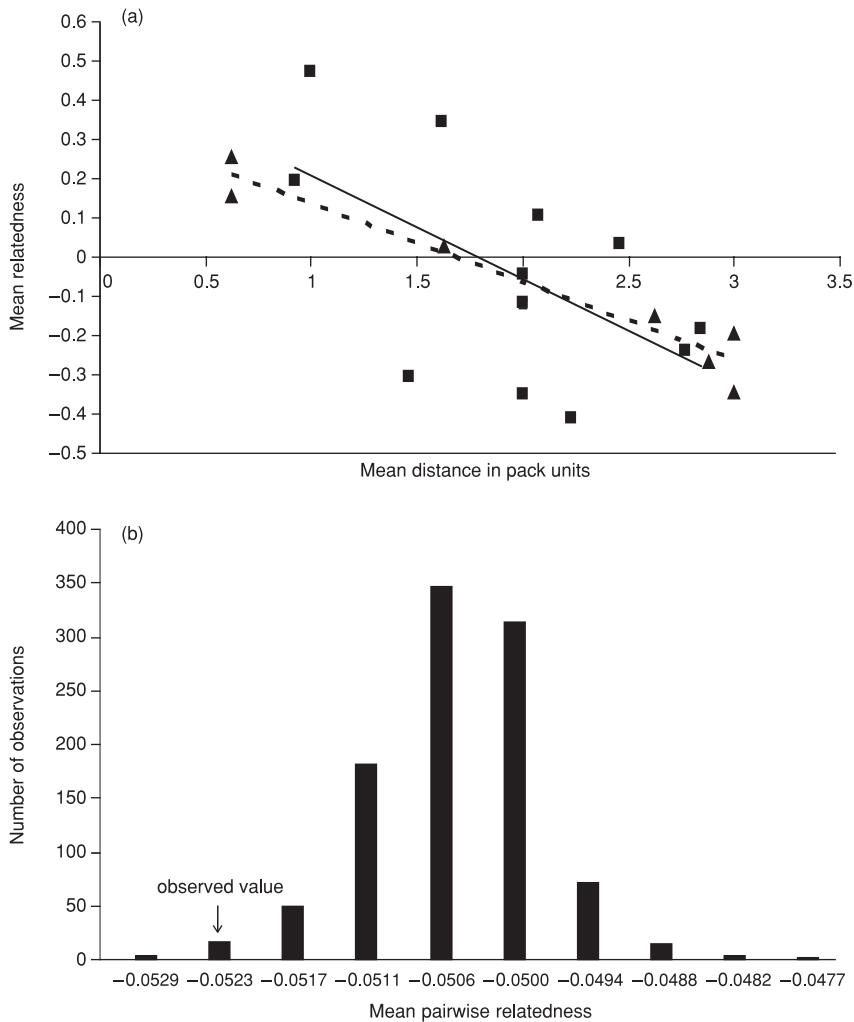
Fig. 3 (a) Regression of mean ranked r_{xyLR} values and corresponding mean distance values in males and females from MWS. ■ and solid line, females ($r^2 = 0.33$ $P = 0.05$) ▲ and dotted line, males ($r^2 = 0.94$, $P = 0.0003$). (b) Results from the permutation test comparing the observed mean pairwise r_{xyLR} value to those expected in randomized unstructured populations.

tree (1000 replicates) shows two sister clades with moderate bootstrap support (Fig. 4). Mean uncorrected *P* distance was 3.4% and 1.9% within clades I and II, respectively, and 6.7% between the two clades. All haplotypes from South, Central and North India (south of the Ganges River) were found within clade I. A strongly supported grouping of haplotypes M, P, and R from Sumatra, Java and captive zoo animals (presumed to belong to the *lepturus* subspecies), was placed within a separate clade with weak/moderate support as part of the larger clade I. Another strong grouping consisted of haplotypes C, L, O, all from South India. Thus, high divergence of haplotypes in samples from South India was evident. Clade II consisted solely of haplotypes from North India (north of the Ganges), northeastern India, Myanmar, Thailand and Malaysia.

Pairwise mean uncorrected *p* and HKY + G distances between haplotypes and haplotype groups are shown in Table 4. Haplotype F, seen in a museum sample from north of the Ganges in North India, was found to be very distinct from the other North Indian and Central Indian haplo-

types. Haplotype K (grouping in clade II) was found to be very distinct from the other haplotype detected in Myanmar (haplotype J which grouped within clade I). The two clades are shown pictorially in Fig. 5 and a median-joining network is shown in Fig. 6. Results from an AMOVA show that significant genetic differentiation of populations is seen in several cases (Table 5). However, rather than in the case of a straightforward grouping by geographical location (groups 3, 5 and 6), high Φ_{CT} values were obtained when haplotypes from south of the Ganges in India were grouped separately from haplotypes from north of the Ganges and from other clade II haplotypes, and also when haplotypes from Sumatra and Java were grouped separately from haplotypes from Malaysia and other clade II haplotypes (groups 7, 8, 9 and 10). The highest Φ_{CT} values were obtained when all clade I haplotypes were grouped separately from clade II haplotypes, with haplotypes from Sumatra, Java and captive samples either contained within the clade I group, or as a separate third group (groups 12 and 11, respectively).

Table 3 Variable sites within the various mtDNA control region haplotypes. Extant samples are faecal samples from current wild populations in MWS (South India), PKWS (northeastern Thailand), BNP (East Java) and TNNP (Malaysia). Numbers of samples are listed in parentheses. Museum samples are listed with museum accession numbers, date of collection and location details. Captive samples from zoos within Germany are listed with information on origin

Hap	11111 2299901223 6913477053	1111111111 4455555666 2902349046	1111111111 6667777777 7890134567	1111222 7899123 8617613	GenBank accession numbers	Sample number, date and location information
A	GCGGGAGCTC	GTTACTTTTA	TTATCGGCTA	GAACGAT	AY682699	Extant (MWS, S. India) (18), ZD1937.1.10.24 (1937, Masinagudi, S. India), ZD1934.10.4.3a (1911, Sipna valley, Berar, C. India)
B	.T........	..CG......	......ATC.	.G..A..	AY682700	Extant (MWS, S. India) (1)
C	..A.....C.	..CGT....G	...C..AT..	.G..A..	AY682701	Extant (MWS, S. India) (1), ZD1888.2.5.22 (1888, Anamalai Hills, S. India)
D	....	....	....	.G.....	AY682702	Extant (MWS, S. India) (1)
E	........T	T.C.....C.	C.CC.AATCG	AG...G.	AY682703	Extant (PKWS, Thailand) (2), ZD1922.12.22.4 (1922, Naga Hills, Assam, E. India)
F	........T	T.C....CC.	C.CC.AA.CG	AG...G.	AY682704	ZD1929.10.19.5 (1909, Oudh, U.P., N. India)
G	........T	T.C.....CG	C.CC.AATC.	AG...G.	AY682705	Extant (PKWS, Thailand) (3)
H	.......C.	..CG......	......ATC.	.G..A..	AY682706	ZD1919.6.2.23 (1919, Kotagiri, Nilgiri Hills, S. India),
I	....	..CG.....G	......ATC.	.G..A..	AY682707	ZD1935.1.1.3 (1935, B.R. Hills, Coimbatore, S. India), ZD1930.5.24.125 (1929, Salem, E. Ghats, S. India)
J	....	......C...	....	....	AY682708	ZE1952.4.7.17 (1923, Nagpur, C. India), ZD1937.12.3.33 (1937, Upper Burma), ZD1907.10.18.3 (1907, Mirzapur, U.P., N. India)
K	....AG.T.T	TCC.....C.	C.CC.AATC.	AG...GC	AY682709	ZD1937.12.3.34 (1937, Katha Dist., Upper Burma)
L	..A...A.C.	..CGT....G	...C..AT..	.G..A..	AY682710	ZD1934.8.3.2 (1934, B.R. Hills, Coimbatore, S. India)
M	...AA.....	..C..CC...	C...T.....	...TA..	AY682711	ZD1845.3.19.5 (1839, Sumatra)
N	....	..CG......	......ATC.	.G..A..	AY682712	ZD1934.10.4.2 (1912, Hoshangabad, C. India)
O	..A.....C.	..C.T....G	...C.AAT..	.G..A..	AY682713	ZD1911.5.31.4 (1910, Kanara, S. India)
P	....	..C..CC...	C...T.....	...TA..	AY682714	Extant (BNP, E. Java) (17), ZD1935.2.26.1 (1935, Modjokerto E. Java)
Q	A.........	..C.......	....	..G.A..	AY682715	Captive (Dresden zoo, Germany, origin Winnipeg zoo, Canada, wild origin presumed to be China) (3)
R	..A.......	..C..CC...	.C..T.....	...TA..	AY682716	Captive (Allwetter & Hodenhagen zoos, Germany origin Novosibirsk zoo, Russia & European zoos, wild origin presumed to be China) (4)
S	.......CT	T.C.....C.	C.CCTAATC.	AG...G.	AY682717	Extant (TNNP, Malaysia) (2), ZD 1846.5.13.2 (1846, Malacca)

To attempt further resolution of these results, a 321-bp fragment of the mitochondrial cytochrome *b* region was also amplified from a number of samples. MWS and BNP faecal samples (*n* = 5 and 6, respectively), museum samples from East Java (1935), South India (1937, 1888, 1919, 1935, 1929), Myanmar (1937 — control region haplotype J), and captive samples from Allwetter Zoo, Germany, were found to have identical sequences (haplotype 1). However, a single substitution (C-T transition, resulting in an amino acid change from arginine to cysteine; haplotype 2) was seen in the PKWS and TNNP faecal samples (*n* = 5 and 2, respectively), and in the museum sample from north of the Ganges (control region haplotype F) (see Fig. 5 for distribution pattern). This result supports the finding that samples grouping in clade II are very distinct from samples in clade

I. The existence of this substitution within the sample from Myanmar and northeastern India (control region haplotypes K and E, respectively), which also grouped in clade II, could not be confirmed since these samples, along with several other museum samples, were found to amplify contaminating human cytochrome *b* sequences.

A mismatch distribution of pairwise differences among haplotypes did not show a unimodal pattern (generally interpreted as a signature of an ancient population explosion; Rogers & Harpending 1992). Instead, it showed a multimodal, erratic (ragged) pattern even when samples from each clade were analysed independently (data not shown). However, Fu's F_S was found to be significant when either all haplotypes [F_S = −10.3 (95% CI −3.35, 4.56)] or just clade I haplotypes [F_S = −7.56 (95% CI −3.22, 5.81)] were analysed,
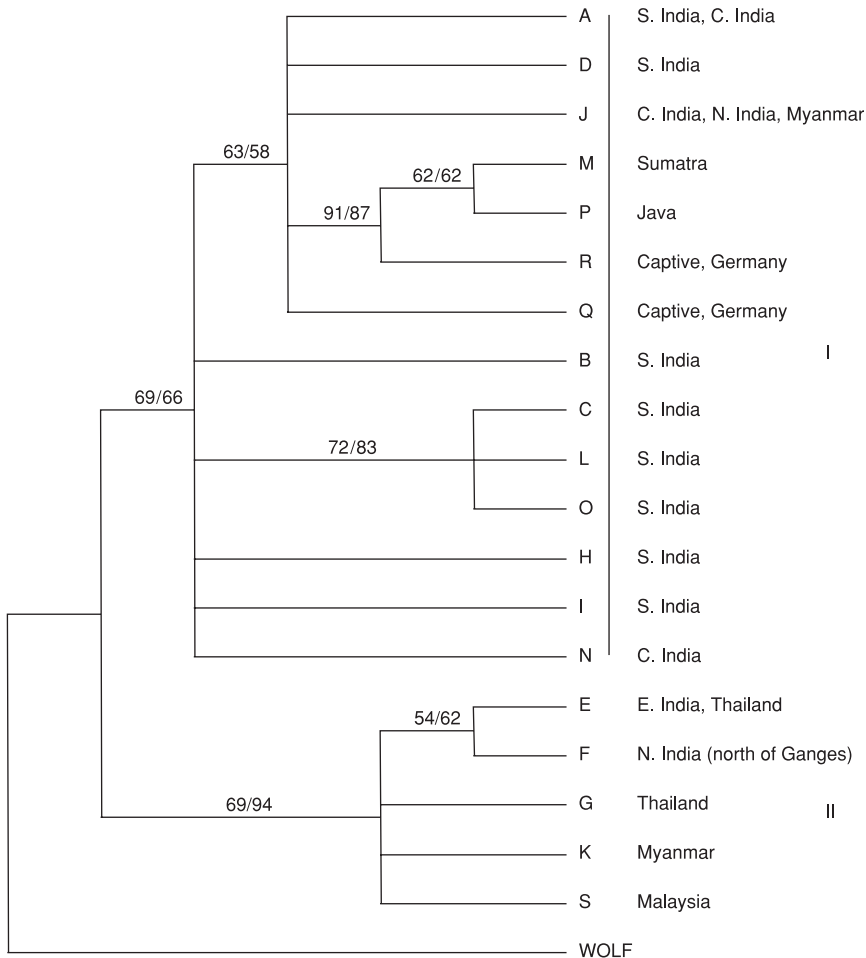
Fig. 4 Bootstrap consensus tree depicting haplotype relationships based on maximum-likelihood (HKY + G) and maximum-parsimony analysis. Numbers above nodes are bootstrap support for the node in 1000 replicates under likelihood on the left and parsimony on the right.

but not when just clade II haplotypes were analysed [F_S = −1.48 (95% CI −3.64, 4.85)]. A significantly large negative value of Fu's F_S rejects population stasis/neutrality, indicating an excess of recent mutations and thus population expansion and/or selection. Range expansion can be distinguished from the effects of selection by the patterns of significance of F_S, F^* and D^* (e.g. Peck & Congdon 2004). Thus a range expansion is indicated when F_S is significant and F^* and D^* are not, while the reverse suggests selection. Nonsignificant results were obtained for Fu and Li's F^* and D^* in all cases.

Discussion

Genetic diversity

The single mitochondrial haplotype and extremely low microsatellite diversity seen in the BNP samples is suggestive of a highly genetically impoverished group of dholes in this part of Java. Such low levels of microsatellite diversity were reported in the endangered Florida puma

(*Puma concolor*), which showed a sevenfold depletion of microsatellite variation compared to other pumas (Driscoll *et al*. 2002) and displayed stark phenotypic signs of inbreeding (Roelke *et al*. 1993). Upon isolation of the islands of Sumatra and Java from the mainland of Asia *c*. 10 000 BP (Meijaard 2003), various species most likely underwent a genetic bottleneck in comparison to their Asian mainland counterparts. Island species are known to retain less molecular diversity than mainland species (Frankham 1998). Massive human population growth and large-scale habitat destruction and fragmentation over the past century or so in Indonesia in general, and in Java in particular, may have led to this highly genetically impoverished group of dholes. BNP, which is situated in the northeastern corner of East Java is surrounded by areas of scrubby terrain with teak plantations, agricultural land and human settlements, which are unlikely to support dholes for any length of time. It is unclear whether two other national parks on the southern coast of East Java (Alas Purwo and Meru Betiri) which have been known to contain dholes in the recent past, retain some interconnectivity with BNP for

Table 4 ×100 pairwise uncorrected *p* (below diagonal) and HKY + G distances (above diagonal) between haplotypes and haplotype groups

		S. India	C.N. India	N. India N. Ganges	Myanmar J	Myanmar K	E. India	Thailand	Malaysia	Captive, Germany Q	Captive, Germany R	Sumatra	Java
S. India (8,2.7)			5.5	16.7	6.7	23.6	13.4	12.5	13.4	6.1	10.2	13.9	10.1
C.N. India (3, 2.2)		2.8		14.5	1.9	23.6	13.4	13.4	15.5	3.0	6.2	7.4	5.1
N. India (N. Ganges)		6.4	6.1		16.6	5.3	0.9	1.5	2.7	19.1	33.2	29.1	21.9
Myanmar (2, 8.1)	J	3.4	1.2	6.5			16.6	16.6	19.1	2.7	4.3	5.3	3.5
	K	7.5	7.4	3.2			3.5	3.5	4.3	33.2	57.4	38.5	38.0
E. India		5.8	5.8	0.8	6.5	2.4		1.4	1.4	19.1	33.2	29.1	22.0
Thailand (2, 0.8)		5.6	5.8	1.2	6.5	2.4	0.4		1.4	19.1	33.2	29.1	22.0
Malaysia (1, 0)		5.8	6.2	2.0	6.9	2.8	1.2	1.2		22.1	29.1	25.6	19.2
Captive	Q	3.4	2.2	6.9	2.0	8.6	6.9	6.9	7.3			6.4	4.3
	R	4.7	3.5	8.5	2.8	10.2	8.5	8.5	8.1			2.7	1.4
Sumatra		5.4	3.9	8.1	3.2	8.9	8.1	8.1	7.7	3.7	2.0		0.9
Java (1, 0)		4.6	3.1	7.3	2.4	8.9	7.3	7.3	6.9	2.8	1.2	0.8	

Haplotypes from Myanmar and captive animals are shown separately. Values in parentheses in column 1 denote number of haplotypes and within-group *p* distance. S. India, South India; C.N. India, Central and North India; N. India, North India; N. Ganges, North of Ganges River; E. India, East India. Only one sample was available from N. India (North of Ganges), East India and Sumatra.

dholes through such terrain. What is also not clear at this stage is whether there are fitness costs associated with this genetic impoverishment.

In contrast to the BNP samples, in the MWS samples from South India, we observed high levels of microsatellite diversity and four mitochondrial control region haplotypes with a 3% mean uncorrected *p* genetic distance. One haplotype (A) was found to be the prevailing haplotype across the majority of individuals in the MWS packs. It has been suggested that dholes undergo periodic population 'boom and busts', the reasons for which are unclear (e.g. Davidar 1975). It is possible that such 'busts' result in a temporary loss/skew of genetic variability in dhole populations, but that this is stabilized over time with immigration and genetic exchange. The overall high genetic variability detected within this small area suggests that any 'busts' have not resulted in a big loss of genetic diversity in this region, perhaps because large tracts of interconnected habitat suitable for dholes still remain, facilitating the movement of animals. In addition, despite a history of human persecution for centuries in India, no evidence for a recent genetic bottleneck was detected in the MWS samples using microsatellite data.

Pack size and structure in MWS

The 25% microsatellite amplification from silica-dried faecal samples obtained in this study is much lower than what has been reported in other carnivores, e.g. 65% and 45% from alcohol-preserved winter and summer scat samples from wolves in Italy (Lucchini *et al.* 2002), and 36.5% from silica-dried faecal samples from brown bears

(Murphy *et al.* 2002). The DET buffer preservation method resulted in higher success rates (60%). However, for reasons of safety in the field and during transportation, it remains desirable to evaluate alternative methods for dhole faecal preservation (e.g. Nsubuga *et al.* 2004; Roeder *et al.* 2004). Low amplification rates and repeat identification of individuals in MWS precluded the genotyping of every individual within a pack. However, our results allow us to draw a number of conclusions about dhole pack structure. First, within-pack relatedness is significantly greater than between-pack relatedness in both males and females. This supports delayed dispersal in both sexes as previously proposed by Venkataraman (1998). These results are not likely to be affected by the sampling of predispersal-age juveniles because samples were pooled from across a number of separate packs and high numbers of such individuals across every pack is not likely. The clear trend in both sexes for a decline in mean relatedness values with an increase in distance values suggests that both dhole males and females are more closely related to individuals from the same or from neighbouring packs than more distant packs. Clear evidence for the existence of kin structure was also obtained using a permutation method. Such a kin-based social structure has been observed in other pack-living canids practising cooperative breeding and hunting, e.g. painted hunting dogs (Girman *et al.* 1997). Our results also suggest that females may be more likely to disperse close to their natal packs while males may disperse longer distances, along similar lines to what has been observed in painted hunting dogs (Girman *et al.* 2001). A male-biased pack structure previously reported from field observations (Johnsingh
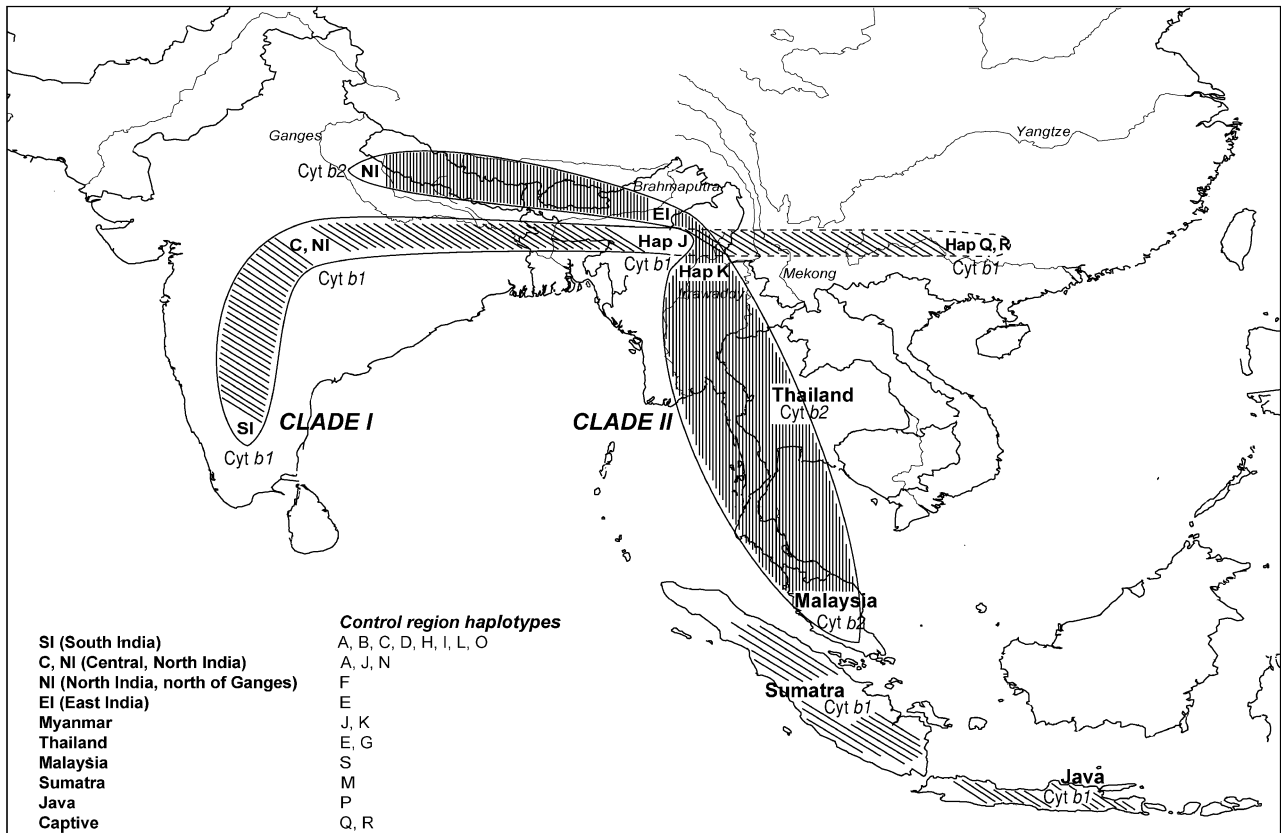
Fig. 5 A diagrammatic representation of the two major phylogeographical groupings seen with mitochondrial control region haplotypes. Haplotypes J and K from Myanmar, and Q and R from captive animals presumed to originate from China are shown separately. Dotted lines represent possible extension of clade I. Mitochondrial cytochrome *b* haplotypes identified from the various locations are also shown.

1982; Venkataraman 1998) was not supported by the results obtained for the Gamehut pack using molecular sexing. However, inadequate overall sampling precludes a definitive conclusion on this issue. The discrepancy seen in the estimates of individuals within the Vazhaithotam pack based on field observations and multilocus genotyping may be due to temporary influx of individuals into the Vazhaithotam pack or single individuals or 'floaters' passing through the home range.

Phylogeography

The phylogenetic analyses reveal two major geographically abutting clades. The deeper phylogenetic structure and longer branches within clade I when compared to clade II suggest older coalescence times and/or higher historical effective population sizes for these populations. This is also reflected by the fact that several of the clade I haplotypes are positioned internally within the network (Posada & Crandall 2001). In the case of clade II, shorter branch lengths and more terminal nodes within the network suggest a more recent coalescence. While we are reluctant to estimate times of divergence using a short

control region sequence, we propose a scenario that could explain the observed patterns. There is evidence for considerable effects of glaciation events on the climate and biogeography within South Asia. During glacial periods, the climate is thought to have been drier, colder, and more seasonal, with lowland areas suffering from increased desertification, and reduction of moist tropical forests into savannah and patchy deciduous forests (e.g. Brandon-Jones 1996; Meijaard 2003, 2004). Much of northern and western India is thought to have been desert at glacial maxima (Fleischer *et al.* 2001). Also, recurrent patterns of disjunct distribution of closely related, forest dependent taxa are observed across Asia that are often best explained in light of the effects of glaciation events where they are thought to have become isolated within ice age refugia and represent relics of a former continuous population (Brandon-Jones 1996; Karanth 2003; Meijaard 2003). For example, a number of taxa found in the wet evergreen forests of southwest India and Sri Lanka are absent from the rest of the Indian subcontinent and then found again in the wet evergreen forests of northeast India and Southeast Asia (Karanth 2003). High levels of diversity were found in dholes from South India, parts of which (the Western
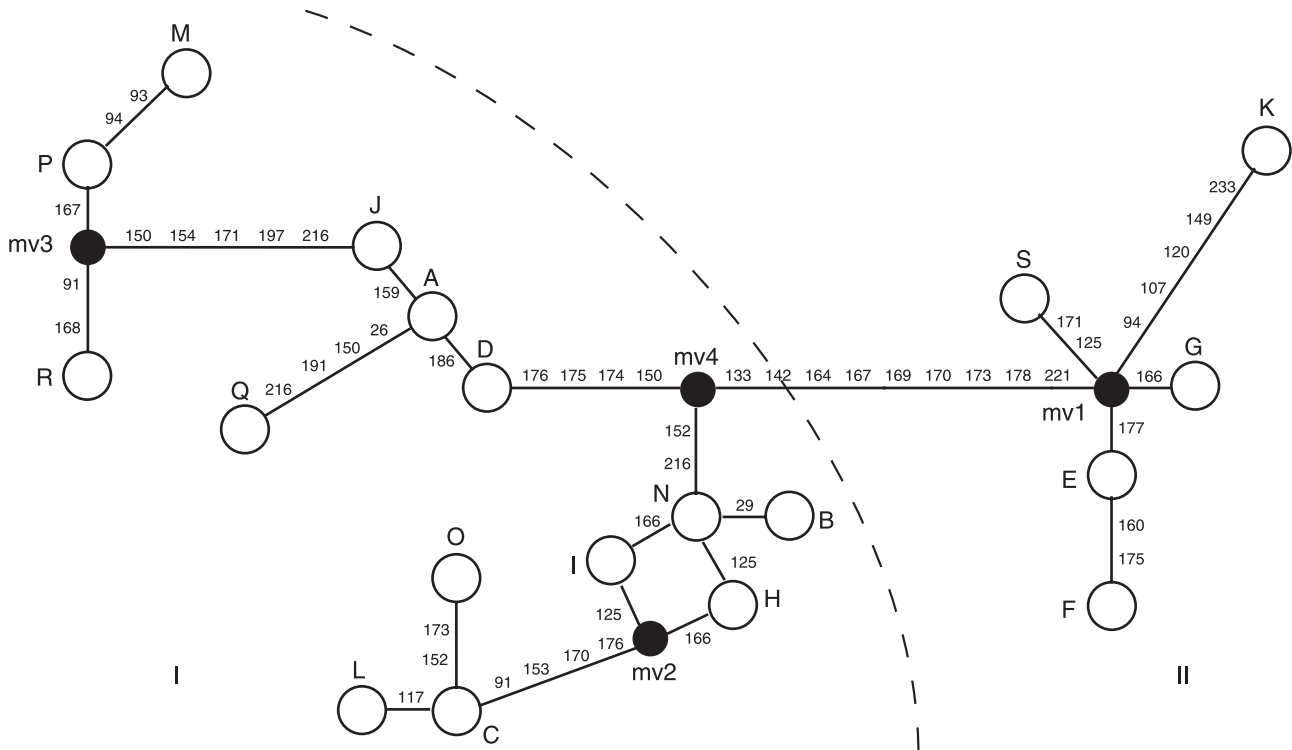
Fig. 6 Median-joining network of haplotypes. The curved dividing line separates the two clades seen in the bootstrapped consensus tree. Open circles represent haplotypes obtained in this study, while closed circles represent median vectors, i.e. presumed unsampled or missing intermediates. Numbers on branches represent substitution positions.

Ghats) are considered to be among the most biologically diverse regions within South Asia. Thus, since South India represents a region that is likely to have acted as a refugium during glaciation, we propose that ancestral clade I haplotypes from such a refugium expanded northwards and eastwards, with a barrier to gene flow in the north in the form of the Ganges River system. Haplotype F from north of the Ganges in India was very distinct from the other North and Central Indian haplotypes from south of the Ganges. Results from cytochrome *b* sequencing and AMOVAs also support this finding. The distinct group of clade II haplotypes are likely to have originated more recently from one or more refugia within this region, separately from clade I haplotypes. A number of studies have suggested the existence of glacial refugia in northern Indochina and the Malay Peninsula (Brandon-Jones 1996; Gorog *et al.* 2004). Our sampling is not sufficiently extensive to confirm or reject the apparent monophyly of clades I and II, and to determine whether Myanmar, where haplotypes from both clades were found, may represent a zone of more recent secondary contact.

Significantly large negative values for Fu's F_S were obtained both when all haplotypes were analysed simultaneously and when clade I haplotypes were analysed separately. In addition, nonsignificant values for Fu and Li's *F** and *D** were seen in both these cases, a pattern that is highly suggestive of an ancient population expansion as expected from a refugial expansion scenario. Fu's F_S is a powerful test for population expansion and has been found to be considerably more sensitive than other such statistics (Fu 1997). However, a mismatch distribution did not show a unimodal pattern characteristic of such an ancient population explosion. Interpretation of mismatch distribution results have on occasion been considered tentative since a number of factors have been found to affect the results, such as time of population expansion, population size before expansion, and subdivision of populations (Marjoram & Donnelly 1994).

Haplotypes M and P from Sumatra and Java, which grouped within clade I, showed high genetic distance values to the haplotype found in Peninsular Malaysia, and were very distinct from the group of haplotypes within clade II. Cytochrome *b* haplotypes from these samples were also identical to those found in samples from clade I and different to that found in Peninsular Malaysia and Thailand. Sumatra and Java are known to have been connected to the Asian mainland via the Malaysian Peninsula as recently as 10 000 years BP when the Sunda Shelf was exposed as a consequence of low sea levels during glaciation (Meijaard 2003). The existence of this land bridge has led some researchers to conclude widespread dispersal of animals across the Sunda Shelf (e.g. Heaney 1986). However, others have suggested that dispersal of terrestrial species

Table 5 Results from AMOVA on various mtDNA haplotype groupings

No.	Pairwise difference		Tamura Nei + G		Grouping
	Φ_{CT}	P	Φ_{CT}	P	
1	0.245	0.057	0.046	0.211	[S. India, C. India, N. India] [N. India N. Ganges, E. India, Myanmar (J and K), captive, Thailand, Malaysia] [Sumatra, Java]
2	0.382	0.007	0.446	0.007	[S. India, C. India, N. India, Myanmar J] [N. India N. Ganges, E. India] [Myanmar K, Thailand] [Malaysia, Sumatra, Java] [captive]
3	0.507	0	0.602	0.002	[S. India, C. India, N. India, N. India N. Ganges] [E. India, Myanmar (J and K)] [Thailand, Malaysia] [Sumatra, Java] [captive]
4	0.510	0	0.661	0	[S. India, C. India, N. India, Myanmar J, Sumatra, Java] [N. India N. Ganges, E. India, Thailand, Myanmar K, Malaysia] [captive]
5	0.534	0	0.522	0.003	[S. India, C. India, N. India] [N. India N. Ganges, E. India], [Myanmar (J and K), captive] [Thailand, Malaysia] [Sumatra, Java]
6	0.563	0	0.672	0	[S. India, C. India, N. India] [N. India N. Ganges, E. India, Myanmar (J and K)] [Thailand, Malaysia] [Sumatra, Java] [captive]
7	0.590	0	0.704	0	[S. India, C. India, N. India, Myanmar J, captive] [N. India N. Ganges, E. India, Myanmar K, Thailand, Malaysia] [Sumatra, Java]
8	0.608	0	0.716	0	[S. India, C. India, N. India, Myanmar J] [N. India N. Ganges, E India] [Myanmar K, Thailand, Malaysia] [Sumatra, Java] [captive]
9	0.609	0	0.734	0	[S. India, C. India, N. India, Myanmar J] [N. India N. Ganges, E. India] [Myanmar K, Thailand, Malaysia] [Sumatra, Java, captive]
10	0.613	0	0.722	0	[S. India, C. India, N. India, Myanmar J] [N. India N. Ganges, E. India, Myanmar K, Thailand, Malaysia] [Sumatra, Java] [captive]
11	0.615	0	0.740	0	[S. India, C. India, N. India, Myanmar J] [N. India N. Ganges, E. India, Myanmar K, Thailand, Malaysia] [Sumatra, Java, captive]
12	0.627	0	0.760	0	[S. India, C. India, N. India, Myanmar J, Sumatra, Java, captive] [N. India N. Ganges, E India, Thailand, Myanmar K, Malaysia]

between the land masses of Sundaland was limited due to the presence of a number of ecological and physical barriers (Brandon-Jones 1996; Meijaard 2003, 2004). In this study, dhole samples from Sumatra and Java were found to be very distinct from the samples from Malaysia, supporting the existence of such barriers. Lower genetic distances were seen between these haplotypes and other haplotypes within clade I, i.e. South, Central and North India (south of the Ganges), Myanmar (haplotype J) and possibly China (haplotypes Q and R). Further studies are required to clarify these results but in the absence of alternative explanations, these results may be suggestive of human translocation of dholes from one of these regions into Sumatra and/or Java. But since there is no documented evidence for such translocation(s), and given that dholes have been long considered vermin (and not hunted for sport, cf. red fox introductions to Australasia), this hypothesis must remain highly speculative. Nonetheless, the Sumatran and Javan populations of dholes are of high conservation importance given the relatively few populations known from the Southeast Asian mainland (Durbin *et al.* 2004b), making further study into the origins of these dholes a high priority for conservation. The low genetic distances displayed by both haplotypes (Q and R) from captive animals thought to have their origins in China and belonging to the *lepturus*

subspecies (H. Maisch, personal communication) to haplotype J within clade I and high distances to haplotypes within clade II, suggest that they may be part of the eastwards expansion of clade I.

Haplotypes were found to be shared between individuals across vast distances, reflecting the highly mobile nature of these carnivores. No clear distinction between subspecies was noted, with haplotypes being shared between *Cuon alpinus dukhunensis* & *Cuon alpinus adjustus* and *C. alpinus adjustus* & *Cuon alpinus infuscus*, and very low genetic distances between several subspecies. A similar lack of distinctness between subspecies of widely distributed Asian mammals has been reported for tigers (*Panthera tigris*) by Cracraft *et al.* (1998). However, the presence of two distinct phylogeographical groups and the grouping of Sumatran and Javan haplotypes with those from India (south of the Ganges) and Myanmar, as opposed to those from Malaysia and Thailand, should be noted.

Conservation implications

Our results do not support the recognition of 11 subspecies of dholes and instead, show admixture over vast areas. However, conservationists should recognize the existence of the two distinct phylogeographical groupings that we have

observed. The unexpected results, which indicate that dholes may have been introduced to Sumatra and Java, need further investigation, given the likely conservation importance of dhole populations on these two large islands. Dholes from the Baluran National Park in East Java were found to be highly genetically impoverished, stressing the requirement for additional genetic studies within this park, and other protected areas in Java in order to better assess the status of these populations and the possible need for genetic management. *Cuon alpinus primaevus*, reported as 'very rare' from a survey in the early 1980s (Johnsingh 1985), and found to be very distinct from *C. alpinus dukhunensis* in this study, must be accorded particularly high priority for conservation action. Within South India, the continued maintenance of habitat interconnectivity is important to retain maximal amounts of the high genetic diversity seen.

Acknowledgements

References

Adams JR, Kelly BT, Waits LP (2003) Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and Coyotes (*Canis latrans*). *Molecular Ecology*, **12**, 2175–2186.

Ayres KL, Overall ADJ (2004) API-CALC 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives. *Molecular Ecology Notes*, **4**, 315–318.

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Belkhir K, Castric V, Bonhomme F (2002) IDENTIX, a software to test for relatedness in a population using permutation methods. *Molecular Ecology Notes*, **2**, 611–614.

Brandon-Jones D (1996) The Asian Colobinae (Mammalia: Cercopithecidae) as indicators of Quaternary climatic change. *Biological Journal of the Linnean Society*, **59**, 327–350.

Cohen JA (1978) *Cuon alpinus. Mammalian Species*, **100**, 1–3.

Cohen JA, Fox MW, Johnsingh AJT, Barnett BD (1978) Food habits of the dhole in south India. *Journal of Wildlife Management*, **42**, 933–936.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Cracraft J, Feinstein J, Vaughn J, Helm-Bychowski K (1998) Sorting out tigers (*Panthera tigris*): mitochondrial sequences, nuclear inserts, systematics, and conservation genetics. *Animal Conservation*, **1**, 139–150.

Davidar ERC (1975) Ecology and behavior of the dhole or Indian wild dog *Cuon alpinus* (Pallas). In: *The Wild Canids: Their Systematics, Behavioral Ecology and Evolution* (ed. Fox MW), pp. 109–119. Van Nostrand Reinhold Company, New York.

Driscoll CA, Menotti-Raymond M, Nelson G, Goldstein D, O'Brien SJ (2002) Genomic microsatellites as evolutionary chronometers: a test in wild cats. *Genome Research*, **12**, 414–423.

Dundas RG (1999) Quaternary records of the dire wolf, *Canis dirus*, in North and South America. *Boreas*, **28**, 375–385.

Durbin LS, Hedges S, Duckworth JW, Tyson M, Iyengar A, Venkataraman A (2004a) *Cuon alpinus*. In: *2004 IUCN Red List of Threatened Species*. http://www.redlist.org.

Durbin L, Venkataraman A, Hedges S, Duckworth JW (2004b) Dhole (*Cuon alpinus*). In: *Status Survery and Conservation Action Plan. Canids: Foxes, Wolves, Jackals and Dogs* (eds Sillero-Zubiri C, Hoffman M, Macdonald DW), pp. 210–219. IUCN, Gland, Switzerland and Cambridge, UK.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Fleischer RC, Perry EA, Muralidharan K, Stevens EE, Wemmer CM (2001) Phylogeography of the Asian elephant (*Elephas maximus*) based on mitochondrial DNA. *Evolution*, **55**, 1882–1892.

Fox MW (1984) *The Whistling Hunters. Field Studies of the Asiatic Wild Dog (*Cuon alpinus*)*. State University of New York Press, New York.

Francisco LV, Langston AA, Mellersh CS, Neal CL, Ostrander EA (1996) A class of highly polymorphic tetranucleotide repeats for canine genetic mapping. *Mammalian Genome*, **7**, 359–362.

Frankham R (1998) Inbreeding and extinction: island populations. *Conservation Biology*, **12**, 665–675.

Frantzen MA, Silk JB, Ferguson JW, Wayne RK, Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology*, **7**, 1423–1428.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Girman DJ, Mills MGL, Geffen E, Wayne RK (1997) A molecular genetic analysis of social structure, dispersal, and interpack relationships of the African wild dog (*Lycaon pictus*) *Behavioral Ecology & Sociobiology*, **40**, 187–198.

Girman DJ, Vilà C, Geffen E *et al.* (2001) Patterns of population subdivision, gene flow and genetic variability in the African wild dog (*Lycaon pictus*) *Molecular Ecology*, **10**, 1703–1723.

Godoy JA, Negro JJ, Hiraldo F, Donázar JA (2004) Phylogeography, genetic structure and diversity in the endangered bearded vulture (*Gypaetus barbatus* L.) as revealed by mitochondrial DNA. *Molecular Ecology*, **13**, 371–390.

Gorog AJ, Sinaga MH, Engstrom MD (2004) Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*). *Biological Journal of the Linnean Society*, **81**, 91–109.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics*, **48**, 361–372.

Hall TA (1999) BIOEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Heaney LR (1986) Biogeography of mammals in SE Asia: estimates of rates of colonization, extinction and speciation. *Biological Journal of the Linnean Society*, **28**, 127–165.

Holmes NG, Dickens HF, Parker HL, Binns MM, Mellersh CS, Sampson J (1995) Eighteen canine microsatellites. *Animal Genetics*, **25**, 132–133.

Ishiguro N, Nakajima A, Horiucji M, Shinagawa M (2002) Multiple nuclear pseudogenes of mitochondrial DNA exist in the canine genome. *Mammalian Genome*, **13**, 365–372.

Johnsingh AJT (1982) Reproductive and social behaviour of the dhole *Cuon alpinus* (Canidae). *Journal of Zoology, London*, **198**, 443–463.

Johnsingh AJT (1985) Distribution and status of dhole *Cuon alpinus* Pallas, 1811 in South Asia. *Mammalia*, **49**, 203–208.

Johnsingh AJT (1992) Prey selection in three large sympatric carnivores in Bandipur. *Mammalia*, **56**, 517–526.

Karanth PK (2003) Evolution of disjunct distributions among wet–zone species of the Indian subcontinent: testing various hypotheses using a phylogenetic approach. *Current Science*, **85**, 1276–1283.

Karanth KU, Sunquist ME (1995) Prey selection by tiger, leopard and dhole in tropical forests. *Journal of Animal Ecology*, **64**, 439–450.

Karanth KU, Sunquist ME (2000) Behavioural correlates of predation by tiger (*Panthera tigris*), leopard (*Panthera pardus*) and dhole (*Cuon alpinus*) in Nagarahole, India. *Journal of Zoology, London*, **250**, 255–265.

Knight ME, Van Oppen MJH, Smith HL, Rico C, Hewitt GM, Turner GF (1999) Evidence for male-biased dispersal in Lake Malawi cichlids from microsatellites. *Molecular Ecology*, **8**, 1521–1527.

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular evolutionary genetics analysis software. *Bioinformatics*, **17**, 1244–1245.

Lucchini V, Fabbri E, Marucco F, Ricci S, Boitani L, Randi E (2002) Noninvasive molecular tracking of colonising wolf (*Canis lupus*) packs in the western Italian Alps. *Molecular Ecology*, **11**, 857–868.

Lucchini V, Galov A, Randi E (2004) Evidence of genetic distinction and long-term population decline in wolves (*Canis lupus*) in the Italian Apennines *Molecular Ecology*, **13**, 523–536.

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. *Genetics*, **152**, 1753–1766.

Marjoram P, Donnelly P (1994) Pairwise comparisons of mitochondrial DNA sequences in subdivided populations and implications for early human evolution. *Genetics*, **136**, 673–683.

Meijaard E (2003) Mammals of South-east Asian islands and their late Pleistocene environments. *Journal of Biogeography*, **30**, 1245–1257.

Meijaard E (2004) Biogeographic history of the Javan leopard *Panthera pardus* based on a craniometric analysis. *Journal of Mammalogy*, **85**, 302–310.

Meyers-Wallen VN, Palmer VL, Acland GM, Hershfiled B (1995) Sry-negative XX sex reversal in the American cocker spaniel dog. *Molecular Reproduction and Development*, **41**, 300–305.

Morin PA, Chambers KE, Boesch C, Vigilant L (2001) Quantitative PCR analysis of DNA from non-invasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Molecular Ecology*, **10**, 1835–1844.

Murphy MA, Waits LP, Kendall KC, Wasser SK, Higbee JA, Bogden R (2002) An evaluation of long-term preservation methods for brown bear (*Ursus arctos*) faecal DNA samples. *Conservation Genetics*, **3**, 435–440.

Nsubuga AM, Robbins MM, Roeder AD, Morin PA, Boesch C, Vigilant L (2004) Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. *Molecular Ecology*, **13**, 2089–2094.

Ostrander EA, Sprague GF, Rine J (1993) Identification and characterisation of dinucleotide repeat (CA) n markers for genetic mapping in dog. *Genomics*, **16**, 207–213.

Ostrander EA, Mapa FA, Yee M, Rine J (1995) 101 new simple sequence repeat-based markers for the canine genome. *Mammalian Genome*, **6**, 192–195.

Palumbi SR (1996) Nucleic acids II: the polymerase chain reaction. In: *Molecular Systematics* (eds Hillis DM, Moritz C, Mable BK), pp. 205–247. Sinauer Associates, Sunderland, Massachusetts.

Peck DR, Congdon BC (2004) Reconciling historical processes and population structure in the sooty tern *Sterna fuscata*. *Journal of Avian Biology*, **35**, 327–335.

Pocock RI (1936) The Asiatic wild dog or dhole (*Cuon javanicus*). *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 33–55.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution*, **16**, 37–45.

Prugnolle F, de Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity*, **88**, 161–165.

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution*, **43**, 258–275.

Randi E, Lucchini V, Christensen MF *et al.* (2000) Mitochondrial DNA variability in Italian and East European wolves: detecting the consequences of small population size and hybridization. *Conservation Biology*, **14**, 464–473.

Raymond M, Rousset R (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Roeder AD, Archer FI, Poinar HN, Morin PA (2004) A novel method for collection and preservation of faeces for genetic studies. *Molecular Ecology Notes*, **4**, 761–764.

Roelke ME, Martenson JS, O'Brien SJ (1993) The consequences of demographic reduction and genetic depletion in the endangered Florida panther. *Current Biology*, **3**, 340–350.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Rozas J, Sánchez-DeI, Barrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Russello MA, Amato G (2004) *Ex situ* population management in the absence of pedigree information. *Molecular Ecology*, **13**, 2829–2840.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, *ver 2.001. A software for population genetics data analysis*. URL: http://lgb.unige.ch/arlequin/.

Stewart P (1993) Mapping the dhole. *Canid News*, **1**, 18–21.

Stewart P (1994) Mapping the dhole update. *Canid News*, **2**, 35–36.

Swofford DL (2002) PAUP*: *Phylogenetic Analysis Using Parsimony (and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.

Thalmann O, Hebler J, Poinar HN, Pääbo S, Vigilant L (2004) Unreliable mtDNA data due to nuclear insertions: a cautionary tale from analysis of humans and other great apes. *Molecular Ecology*, **13**, 321–335.

Thenius E (1954) On the origins of the dhole. *Osterr Zoological Zeitsch*, **5**, 377–388.

Venkataraman A (1998) Male-biased adult sex ratios and their significance for cooperative breeding in dhole, *Cuon alpinus* packs. *Ethology*, **104**, 671–684.

Venkataraman A, Arumugam R, Sukumar R (1995) The foraging ecology of dhole (*Cuon alpinus*) in Mudumalai Sanctuary, Southern India. *Journal of Zoology, London*, **237**, 543–561.

Vigilant L, Hofreiter M, Siedel H, Boesch C (2001) Paternity and relatedness in wild chimpanzee communities. *Proceedings of the National Academy of Sciences, USA*, **98**, 12890–12895.

Vilà C, Amorim IR, Leonard JA *et al.* (1999) Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. *Molecular Ecology*, **8**, 2089–2103.

Wasser SK, Houston CS, Koehiler GM, Cadd GG, Fain SR (1997) Techniques for the application of faecal DNA methods to field studies of Ursids. *Molecular Ecology*, **6**, 1091–1097.

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

A. Iyengar is interested in the application of molecular techniques to address issues important to the conservation of endangered species. V. N. Babu is interested in canid ecology and behaviour, and carnivore community ecology. S. Hedges currently works on Asian mammal conservation projects focusing on elephants, wild cattle, and dholes. A. B. Venkataraman is interested in applying knowledge of behavioural ecology in the design of landscapes for conservation. N. Maclean is professor of genetics at the University of Southampton, with interests in genetic manipulation of fish and molecular ecology. P. A. Morin currently directs the Molecular Ecology Laboratory at the Southwest Fisheries Science Centre, conducting basic and applied genetics research for marine mammal conservation.