npg

Population differentiation within and among Asian elephant (*Elephas maximus*) populations in southern India

TNC Vidya[1], P Fernando[2], DJ Melnick[2,3,4] and R Sukumar[1,4]

[1]*Centre for Ecological Sciences, Indian Institute of Science, Bangalore 560012, India;* [2]*Centre for Environmental Research and Conservation, Columbia University, 1200 Amsterdam Avenue, New York, NY 10027, USA;* [3]*Department of Ecology, Evolution and Environmental Biology, Columbia University, 1200 Amsterdam Avenue, New York, NY 10027, USA*

Southern India, one of the last strongholds of the endangered Asian elephant (*Elephas maximus*), harbours about one-fifth of the global population. We present here the first population genetic study of free-ranging Asian elephants, examining within- and among-population differentiation by analysing mitochondrial DNA (mtDNA) and nuclear microsatellite DNA differentiation across the Nilgiris-Eastern Ghats, Anamalai, and Periyar elephant reserves of southern India. Low mtDNA diversity and 'normal' microsatellite diversity were observed. Surprisingly, the Nilgiri population, which is the world's single largest Asian elephant population, had only one mtDNA haplotype and lower microsatellite diversity than the two other smaller populations examined. There was almost no mtDNA or microsatellite differentiation among localities within the Nilgiris, an area of about 15 000 km². This suggests extensive gene flow in the past, which is compatible with the home ranges of several hundred square kilometres of elephants in southern India. Conversely, the Nilgiri population is genetically distinct at both mitochondrial and microsatellite markers from the two more southerly populations, Anamalai and Periyar, which in turn are not genetically differentiated from each other. The more southerly populations are separated from the Nilgiris by only a 40-km-wide stretch across a gap in the Western Ghats mountain range. These results variably indicate the importance of population bottlenecks, social organization, and biogeographic barriers in shaping the distribution of genetic variation among Asian elephant populations in southern India.
Heredity (2005) **94**, 71–80. doi:10.1038/sj.hdy.6800568
Published online 29 September 2004

Introduction

The distribution of the endangered Asian elephant has contracted over the past two to three millennia to approximately 6% of its historical range (Sukumar, 2003, p 45). Currently, India holds an estimated 24 000–28 500 elephants, distributed across the north-western and north-eastern Himalayan foothills, central, and southern India, with population sizes of 1000–1500, 9000–10 000, 1500–2500, and 12 500–14 500, respectively (Sukumar and Santiapillai, 1996; Asian Elephant Research and Conservation Centre (AERCC), 1998; Bist, 2002; Sukumar, 2003). As elsewhere, elephant populations in India are affected by habitat fragmentation and loss, human–elephant conflict, and poaching for ivory. Elephants in southern India are distributed along the Western and Eastern Ghats mountain ranges, considered an important core area for the long-term conservation of the species (AERCC, 1998).

Although over 50% of free-ranging Asian elephants are found in India, with southern India accounting for 22%,

there has been no comprehensive population genetic study of elephants in India. Previous genetic studies of Asian elephants based on allozymes (Nozawa and Shotake, 1990; Hartl *et al*, 1995) and mtDNA sequence variation (Hartl *et al*, 1996; Fernando *et al*, 2000; Fleischer *et al*, 2001) have been limited in their sampling of Indian elephants and have relied exclusively on captive animals, whose provenance is often obscure. Here, we describe the first population genetic study of free-ranging elephants in India.

The elephant being a large mammal with home ranges of over 500 km² (and linear range of about 50 km) in parts of southern India (Baskaran *et al*, 1995), one normally expects little structuring at the nuclear genome over short distances, and some isolation by distance at much greater distances. On the other hand, as elephants live in matriarchal societies, substructuring at mitochondrial DNA (mtDNA) may be expected even in small populations (Fernando and Lande, 2000). We examine the scale at which structuring in populations was seen and the processes responsible for the observed patterns.

The study includes all three major populations, defined by geographic discontinuity, in southern India: the Nilgiris, Anamalai, and Periyar populations. mtDNA and nuclear microsatellite DNA markers were used to: (1) examine the extent and patterns of distribution of genetic diversity and (2) make inferences about gene flow between geographic locations in southern India.
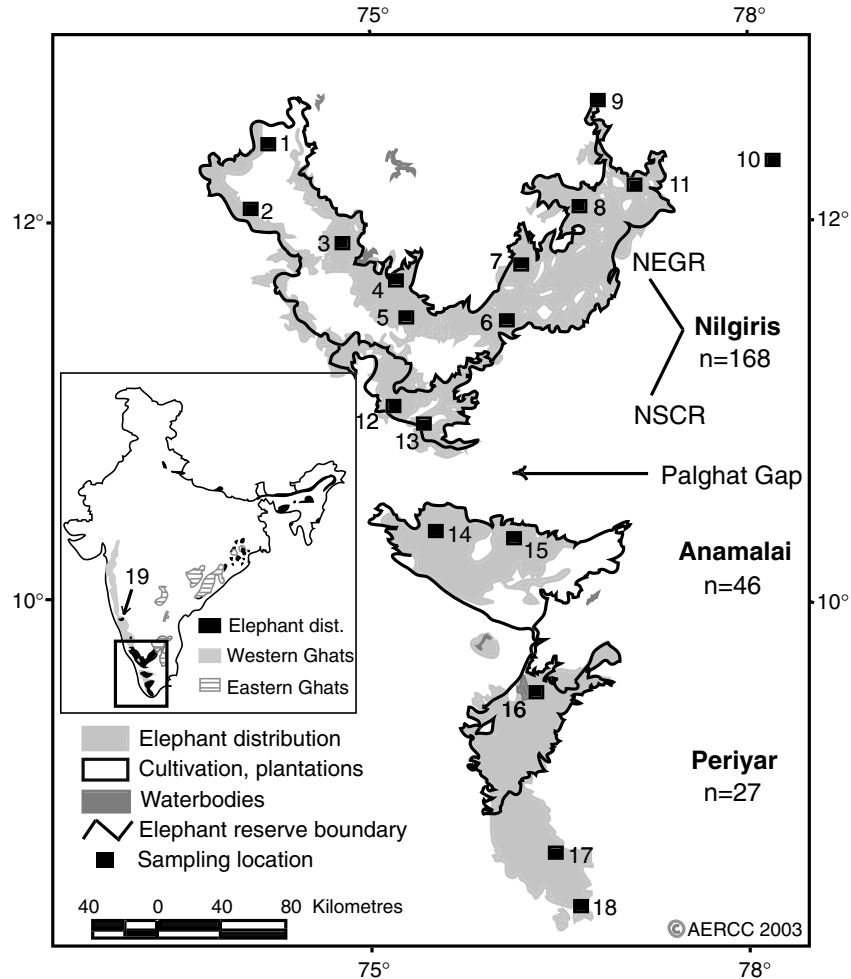
Figure 1 Elephant distribution in southern India and sampling locations. For the purpose of management, large areas of the Nilgiri, Anamalai, and Periyar populations are managed as Project Elepant Reserves: Project Elephant Reserve No. 7 or the Nilgiris-Eastern Ghats Reserve (NEGR) and Reserve No. 8 or the Nilambur – Silent Valley – Coimbatore Reserve (NSCR) in the Nilgiris, Reserve No. 9 or the Anamalai – Parambikulam Reserve in Anamalai, and Reserve No. 10 or the Periyar Reserve in Periyar. The boundaries of these reserves are shown. Dots denote the general area sampled, which in most cases, consisted of several actual sampling sites. The names of the areas sampled are: 1. Madikeri Forest Division ($n = 19$), 2. Virajpet Forest Division ($n = 1$), 3. Rajiv Gandhi National Park ($n = 18$), 4. Bandipur National Park ($n = 10$), 5. Mudumalai Wildlife Sanctuary ($n = 70$), 6. Satyamangalam Forest Division ($n = 10$), 7. BRT Wildlife Sanctuary ($n = 18$), 8. Cauvery Wildlife Sanctuary ($n = 2$), 9. Bannerghatta National Park ($n = 3$), 10. Thirupattur Division ($n = 1$), 11. Hosur Forest Division ($n = 9$), 12. Silent Valley National Park ($n = 4$), 13. Mannarkkad Forest Division ($n = 3$), 14. Parambikulam Wildlife Sanctuary ($n = 17$), 15. Indira Gandhi Wildlife Sanctuary ($n = 29$), 16. Periyar Tiger Reserve ($n = 21$), 17. Kalakkad – Mundanthurai Tiger Reserve ($n = 5$), 18. Kanyakumari Division ($n = 1$), and 19. Dandeli Wildlife Sanctuary ($n = 2$). Inset: Map of India showing the elephant distribution in India, the Western and Eastern Ghats mountain ranges, and a sampling location (19) in the North Kanara population, which is disjunct from the Nilgiris region further south.

These objectives allow us to describe the 'genetic landscape' of elephants in southern India and detect the presence of 'phylogeographic partitioning' (Avise *et al*, 1987). In addition, the distribution patterns of the two divergent Asian elephant mtDNA clades (Hartl *et al*, 1996; Fernando *et al*, 2000; Fleischer *et al*, 2001) are important in tracing the evolutionary history and phylogeography of the species.

Materials and methods

Study area and sampling
The elephant distribution in southern India covers 39 500 km^2 (AERCC, 1998; Bist, 2002), north and south of the Palghat Gap, along the Western Ghats and the Eastern Ghats mountain ranges (Figure 1). Additional

discontinuities exist, North Kanara and Nilgiris and Anamalai and Periyar, forming distinct populations to the north and south of the gap, respectively (Figure 1). The northern end of the elephant distribution in southern India is restricted to pockets of approximately 260 and 50 elephants, in Bhadra Wildlife Sanctuary and Dandeli Wildlife Sanctuary (location 19, Figure 1), respectively (unpublished census data for 2002), which together are referred to here as the *North Kanara* population. The *Nilgiris-Eastern Ghats* (Nilgiris) population, which extends over part of the Western and the Eastern Ghats (Figure 1), is the world's single largest contiguous population of Asian elephants (AERCC, 1998; unpublished census data for 2002), with an estimated 9400 individuals. The *Anamalai* population lies to the south of the Nilgiris, separated from it by the Palghat Gap. The gap is situated south of the junction of the Eastern and

the Western Ghats ranges (at 10°31′N latitude and longitude 76°45′E, 180 m above msl) and is the only major discontinuity in the Western Ghats chain of mountains (Figure 1). It is about 40 km wide and fairly steep, the Nilgiri and Anamalai blocks rising to over 2000 m above msl on either side of the gap. The Anamalai region holds approximately 1600 (AERCC, 1998) elephants. South of Anamalai, and contiguous with it till the early 20th century, lies the *Periyar* population of approximately 1700 elephants (AERCC, 1998).

Obtaining traditional sources of DNA such as tissue or blood from free-ranging, endangered, large mammals is logistically difficult, and thus will often not allow adequate sampling of such species. We therefore used dung samples as the source of DNA. In total, 224 samples were obtained from free-ranging elephants from locations shown in Figure 1. In addition, 19 samples were obtained from captive elephants in state forest department elephant camps, for which reliable details of capture were available. Samples were collected from the outer most layer of dung, which would have the least degraded DNA, usually into 95% ethanol, although 14 samples were collected in storage buffer (see Fernando *et al*, 2003a). Of the 243 samples, 91% were collected within a few hours of defaecation, 6% were several hours to less than a day old, and 3% were about two days old.

Genetic analyses

Extraction followed Fernando *et al* (2003a) and involved digestion of 0.5 g of dung with Proteinase K, followed by extraction with phenol/chloroform/isoamyl alcohol and purification using QIAGEN gel purification columns. Extractions were carried out in an area reserved for low-copy number samples. Polymerase chain reactions using the primers MDL3 and MDL5 (Fernando and Lande, 2000) were carried out to amplify a 600-bp segment of mtDNA, containing the C-terminal of cytochrome *b* and the hypervariable left domain of the noncoding control region. PCR products were sequenced from both directions using the internal primers MDLseq-1 (5′-CCTACAYCATTATYGGCCAAA-3′) and MDLseq-2 (5′-AGAAGAGGGACACGAAGATGG-3′), in 9 μl BigDye (Applied Biosystems Inc.) terminated cycle sequencing reactions. Cycle sequencing products were purified and electrophoresed on 5.6% polyacrylamide gels in an ABI Prism 377 DNA Sequencer. New haplotypes were confirmed by repeating the PCR.

Six microsatellite loci, the tri- and tetranucleotide loci EMX-1, EMX-2, EMX-3, and EMX-4, developed from an Asian elephant (Fernando *et al*, 2001a), and the dinucleotide loci LafMS02 and LafMS03, developed from African elephants (Nyakaana and Arctander, 1998), were amplified and scored following the guidelines in Fernando *et al* (2003a) and using the ABI Gene Scan analysis software v.3.1.2 (Applied Biosystems Inc.). The amplification of microsatellites from noninvasive sources of DNA may be fraught with difficulties due to inhibition (Taberlet *et al*, 1999), nonamplification of one allele (Kohn and Wayne, 1997; Frantzen *et al*, 1998; Taberlet *et al*, 1999), and/or electrophoresis artefacts (Fernando *et al*, 2001b). Errors in genotyping can be eliminated by carrying out multiple extractions and PCRs, but would constitute a huge addition in time and expense, and therefore we worked instead on minimizing the errors. Separate areas and dedicated instruments were used for low-copy number samples and amplified product, in addition to using filter tips for pipetting all samples. A comparison of 480 PCR reactions each from dung- and blood-extracted DNA of 20 identified captive elephants (two extracts from each blood/dung sample and two PCRs per extract), using the same six loci as in this study, revealed a low genotyping error of 1.88%, which included only 0.42% error due to allelic dropout (Fernando *et al*, 2003a). This level of error, while is not acceptable in studies requiring individual identification, does not make a difference to the results of a population genetic study. In addition, 10–15 random samples from those used in the present study were genotyped a second time at each of the six loci, and none of these showed any error. A different set of 10 samples (of total DNA concentration 9–190 ng/μl) were also used in a dilution experiment (in addition to the dilution experiment in Fernando *et al*, 2003a), and no allelic dropout was detected at a lower dilution than $\frac{1}{20}$ to $\frac{1}{50}$, suggesting that sufficient target DNA was present in dung-extracted DNA extracts.

MtDNA sequence data analyses

Sequences were aligned and edited using SEQUENCHER v.3. 1.1 (Gene Codes Corporation, 1999). A sequence that varied by one or more nucleotides was considered a different haplotype. Haplotype diversity ($\hat{H}$) (Nei, 1987, p 180) and nucleotide diversity (Nei, 1987, p 257) were calculated using C programs (written by TNCV, available on request). Sequence divergence between haplotypes was computed using the Jukes–Cantor distance (Jukes and Cantor, 1969). A minimum-spanning network was created using Arlequin v.2.000 (Schneider *et al*, 2000), combining haplotypes from this study along with previously published haplotypes of Fernando *et al* (2000). F_{ST} between populations was also calculated using Arlequin v.2.000.

Microsatellite data analyses

Allele frequencies, heterozygosity, and allelic richness were calculated using C programs (written by TNCV, available on request). Allelic richness (mean number of alleles per locus) was computed by multiple random subsampling using a sample size smaller than that of the smallest population (see Leberg, 2002), and differences across populations tested using Mann–Whitney *U*-tests in STATISTICA v.6 (StatSoft, Inc., 2001). Overall allelic richness across loci was calculated similarly by sampling multiple sets of individuals from each population, but counting the total number of distinct alleles each set of individuals had across all loci, instead of using different random sets of individuals for each locus. Linkage disequilibrium tests between pairs of loci, with the null hypothesis of random association of genotypes at these loci, and the Hardy–Weinberg equilibrium test, with the null hypothesis of random union of gametes at each locus, were performed for each population using Genepop v.3.1 (Raymond and Rousset, 1995). Type I errors were corrected for by applying the sequential Bonferroni test *a posteriori* (see Rice, 1989).

Pairwise population differentiation among the Nilgiri, Anamalai, and Periyar populations, and between pairs of locations within the Nilgiris, were estimated using the

statistics F_{ST} (based on Weir and Cockerham, 1984) and R_{ST} (Slatkin, 1995), which are calculated based on variance in allele frequencies and allele lengths, respectively. Arlequin v.2.000 (Schneider *et al*, 2000) and RSTCALC v.2.2 (Goodman, 1997) were used to compute F_{ST} and R_{ST}, respectively. We examined hierarchical genetic structuring based on an analysis of molecular variance (AMOVA) (Excoffier *et al*, 1992, Michalakis and Excoffier, 1996) and locus-by-locus AMOVA, as executed in Arlequin. AMOVA measures the partitioning of variance at different levels of population subdivision, giving rise to an analogue of F-statistics called Φ-statistics.

Isolation by distance within the Nilgiri population was examined in a Mantel test (Mantel, 1967), which examines the correlation between matrices of pairwise geographic distances and F_{ST} values, using a C program (written by TNCV, available on request). To examine breaks in gene flow, a NW-NE, NW-S, and NE-S, transects connecting sampling locations were created, and pairwise $F_{ST}/(1-F_{ST})$ values corrected for geographic distance were plotted against mid-pair distances, for adjacent pairs of locations. This procedure negates the effect of geographic distance, if any, on F_{ST} and identifies barriers to gene flow.

Evidence of a recent population bottleneck was assessed using a test for heterozygosity excess (Cornuet and Luikart, 1996) and a graphical test to detect mode shifts in allele frequency distributions (Luikart *et al*, 1998). For the former, the two-phase model of mutation and 1000 randomizations in the program BOTTLENECK v.1.2.02 (Piry *et al*, 1997) were used.

Results

MtDNA sequence data

MtDNA d-loop sequences were obtained from 226 samples. Sequence alignment and analyses revealed five unique haplotypes. A single haplotype BN was observed in all 159 amplified samples from the Nilgiris and the sample from North Kanara that amplified. Three haplotypes were observed in Anamalai: BF in 37 (86%), BA in four (9.3%), and BL in two (4.7%), individuals. BF predominated in Periyar also ($n = 23$; 95.8%), with a second haplotype BB restricted to a single individual. Haplotypes BN, BF, and BL were found in both males and females and haplotypes BA and BB in females. Haplotypes BN and BL correspond to N and L of Fernando *et al* (2000), while haplotypes BF, BA, and BB (GenBank accession numbers AY245823, AY365432, AY365433) were undescribed previously.

All haplotypes from southern India were closely related to each other and formed part of the β clade of Fernando *et al* (2000) (Figure 2), corresponding to the 'A' clade of Hartl *et al* (1996) and Fleischer *et al* (2001). Mutational changes between these haplotypes included only transitions. Assuming a substitution rate of 3.5% per million years reported for 675 bp (including the 600-bp segment that we used) of mitochondrial d-loop (Fleischer *et al*, 2001) and a generation time of 17.5 years (Sukumar, 1989, p 179), a single base pair change would correspond to about 42 325 years or 2419 generations. The mean (± 1.96 SE) Jukes–Cantor distance between BN and the other haplotypes was 0.008 ± 0.0014, corresponding to
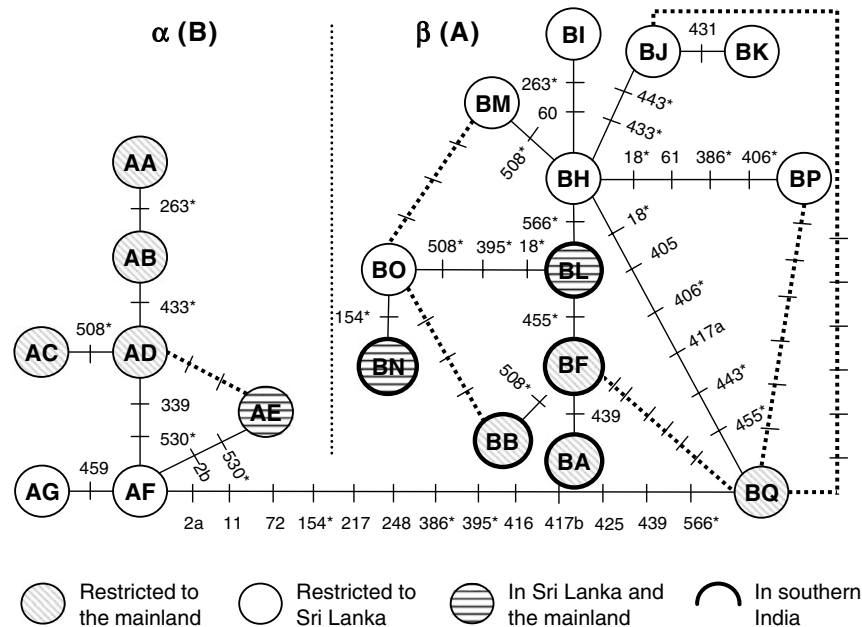
Figure 2 Minimum-spanning network of haplotypes from this study combined with those from Fernando *et al* (2000). The numbers against crosshatches indicate polymorphic positions between haplotypes. Positions with both transitional and transversional changes are denoted by 'a' and 'b', respectively. The dotted lines show alternate connections. Distances between haplotypes are proportionate to the pairwise differences between them, with the exception of BF–BQ and BJ–BQ. Locations sampled in Fernando *et al*'s (2000) study in the mainland included Bhutan ($n = 13$), northeastern India ($n = 6$), Laos ($n = 14$), and Vietnam ($n = 4$), and samples from Sri Lanka included 18 samples from northern Sri Lanka (haplotypes BH, BI, BJ, BK, BL), 25 from mid-latitude Sri Lanka (haplotypes BH, BL, BM, BN, BO, BP), and 38 samples from southern Sri Lanka (haplotypes AE, AF, AG, BL, BP). BQ was the only β clade haplotype on the mainland, found at a frequency of 0.071 in Laos.

228 555 (188 558–268 552) years or 13 060 (10 774–15 346) generations.

Haplotype diversity in Nilgiris, Anamalai, and Periyar was 0.0, 0.255 ± 0.0830, and 0.083 ± 0.0749, respectively. Taken as a single population, the overall haplotype and nucleotide diversities in southern India were 0.436 ± 0.0289 and 0.0036 ± 0.0022, respectively. F_{ST} between Nilgiris and Anamalai and Nilgiris and Periyar were 0.989 and 0.998, respectively ($P<0.001$), reflecting the absence of shared haplotypes. Anamalai and Periyar, however, were not significantly differentiated from each other, with an F_{ST} of 0.029 ($P>0.05$).
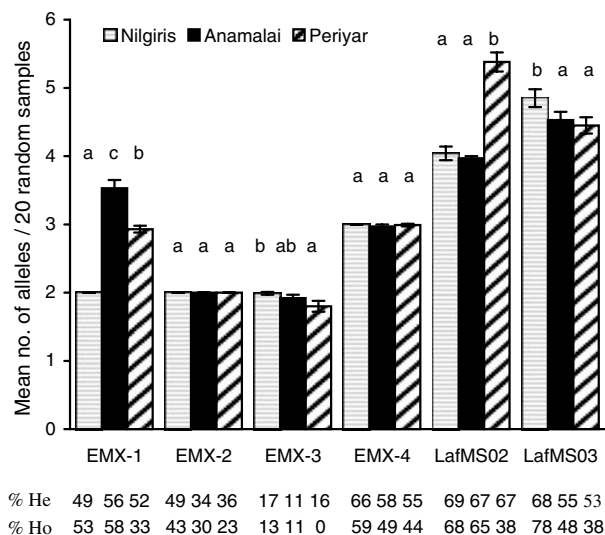
Microsatellite data

Microsatellites could be amplified from 86 to 97% of the samples, depending on the locus. Polymorphism was low at the tri- and tetranucleotide loci, with 4, 2, 2, and 3 alleles at EMX-1, EMX-2, EMX-3, and EMX-4, respectively, and moderately high at LafMS02 and LafMS03 with six alleles each. Observed heterozygosity was lower in Periyar compared to the other two populations (Wilcoxon's matched pairs test: Nilgiris and Periyar, $P=0.028$; Anamalai and Periyar, $P=0.028$; Nilgiris and Anamalai, $P=0.116$), but expected heterozygosity did not differ across populations (Wilcoxon's matched pairs test: Nilgiris and Periyar, $P=0.116$; Anamalai and Periyar, $P=0.917$; Nilgiris and Anamalai, $P=0.075$) (Figure 3). Each population had the highest number of alleles at one locus at least (Figure 3), but the overall allelic richness in Nilgiris (mean 2.78 alleles per locus) was significantly lower than those in Anamalai (2.88) and Periyar (2.86) (Mann–Whitney U-tests: Nilgiris and Periyar, $P=0.031$; Nilgiris and Anamalai, $P<0.001$; Anamalai and Periyar, $P=0.271$). Allele frequencies in the three populations are listed in Appendix A.

Microsatellite data analyses were carried out on 140, 45, and 27 samples, from Nilgiris, Anamalai, and Periyar, respectively, after excluding 19 samples from location 5, which was over-represented. The loci LafMS02 and EMX-3 deviated from Hardy–Weinberg equilibrium in the Periyar population ($P=0.0006$, 0.0014, respectively; first significant P-value for the sequential Bonferroni correction $=0.003$), possibly as a consequence of nonrandom mating, rather than selection, as indicated by significant inbreeding coefficients (F_{IS}) of 0.44 and 1.00 ($P=0.001$, <0.001, respectively) at these loci, respectively. Ivory poaching of males is known to have been prevalent in Periyar in the 1970s and 1980s giving rise to a present adult sex ratio of about one male:100 females

(Ramakrishnan *et al*, 1998), and the individuals sampled were on average 25–30 years old, based on visual estimates (see Sukumar, 1989, pp 224–226), thus giving rise to the possibility of a few males fathering offspring in the region. All loci were in Hardy–Weinberg equilibrium in the other two populations. No significant linkage disequilibrium was observed between any pair of loci in any population ($P>0.01$, first significant P-value for the sequential Bonferroni correction in each population $=0.003$).

Genetic structure

An AMOVA of the three populations indicated that a large percentage of total genetic variation (94.8%) was within populations, but an appreciable amount (5.2%; $P<0.001$) was among populations (Table 1). The locus-by-locus AMOVA revealed concordant patterns among loci ($\Phi_{ST}=0.03-0.09$, $P<0.001$), except for the tetranucleotide locus EMX-3 (the locus with the lowest heterozygosity, Figure 3), which did not show significant variation across populations ($\Phi_{ST}=-0.003$, $P=0.576$).

| % He | 49 56 52 | 49 34 36 | 17 11 16 | 66 58 55 | 69 67 67 | 68 55 53 |
| % Ho | 53 58 33 | 43 30 23 | 13 11 0 | 59 49 44 | 68 65 38 | 78 48 38 |

Figure 3 Mean ±1.96 SE number of alleles from 100 sets of 20 randomly drawn samples from the three populations. Results from pairwise Mann–Whitney U-tests between populations are denoted above the bars: 'a' is significantly smaller than 'b', which is significantly smaller than 'c' (significance at 0.05). Below each bar are the expected and observed percentages of heterozygotes at the corresponding population and locus.

Table 1 Hierarchical AMOVA based on nuclear microsatellite DNA: percentage of variation explained by different levels of population structure, variance components, and Φ-statistics

Level of analysis	df	Variance components	Percent variation	Φ-statistics	P-value
Among populations (Nilgiris, Anamalai, Periyar)	2	0.0733 Va	5.23	$\Phi_{ST}=0.052$	<0.001
Within populations	421	1.3276 Vb	94.77		
Among locations within Nilgiris	10	0.0320 Va	2.23	$\Phi_{ST}=0.022$	0.003
Within locations	265	1.4058 Vb	97.77		
Within NEGR	7	0.0088 Va	0.62	$\Phi_{ST}=0.006$	0.177
Within locations	250	1.4304 Vb	99.38		

The three populations considered are Nilgiris, Anamalai, and Periyar; the Nilgiris further consists of two zones: NEGR and NSCR, the latter being represented by only two sampling locations.

Pairwise F_{ST} and R_{ST} values differentiated Nilgiris and Anamalai and Nilgiris and Periyar, but not Anamalai and Periyar (Table 2), indicating gene flow between the latter two populations. Locus-by-locus AMOVAs of pairs of populations were consistent with this pattern, the differentiation between Nilgiris and Anamalai being supported by all loci except EMX-3, and Anamalai and Periyar being identified as a single population at all loci.

The AMOVA on locations within the Nilgiris revealed only 2.23% of the variance between locations ($\Phi_{ST} = 0.022$, $P = 0.003$). Pairwise F_{ST} and R_{ST} values between these locations showed that only two of the 45 pairwise F_{ST}'s and five of the R_{ST}'s were significant, and involved locations 12 and 13 (Table 3). However, these significant values may be due to the small sample sizes of these two locations compared with much larger sample sizes and/or the absence of data at locus EMX-4 in location 12. An AMOVA excluding locations 12 and 13 (ie using only the NEGR region of Nilgiris, Figure 1) identified 99.4% of the total variation as being within locations, and no significant substructuring was discernible ($\Phi_{ST} = 0.006$, $P = 0.177$). In the locus-by-locus AMOVA, only locus LafMS03 showed slight significant structuring within NEGR ($\Phi_{ST} = 0.029$, $P = 0.021$), which was due to location 11 as shown by its exclusion.

Isolation by distance and breaks in gene flow

Isolation by distance was not detected within the Nilgiri population (Mantel test $P > 0.05$), consistent with most pairwise F_{ST}'s being nonsignificant. The E–W transect within the Nilgiris did not show any discontinuity in gene flow. However, the N–S transects showed a distinct break in gene flow across the Palghat Gap, between Nilgiris and Anamalai (Figure 4). This was especially striking in view of the geographic distances between

locations 13 and 14 (on either side of the gap), and, for instance, between 1 and 11, or 1 and 13 (Figures 1 and 4a).
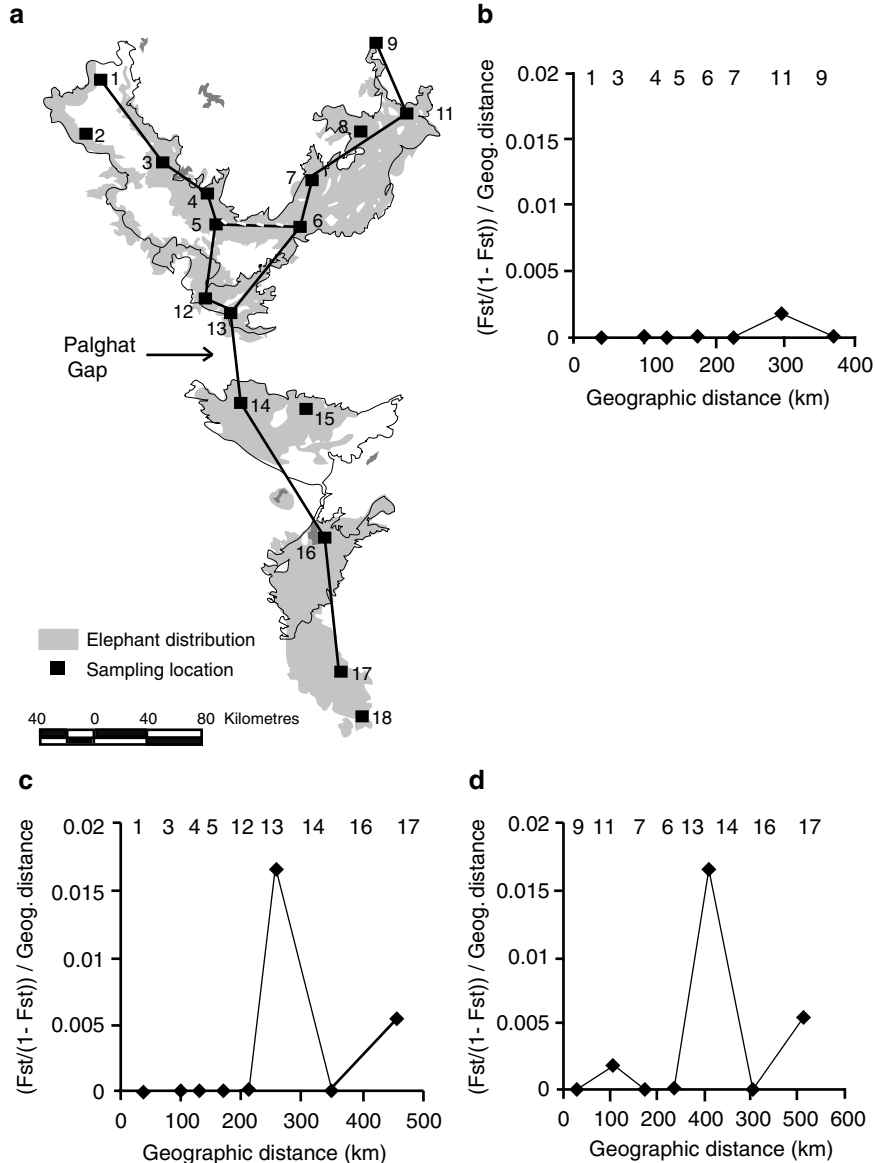
Test for recent population bottleneck

The test for heterozygosity excess was significant at two loci in the Nilgiri population (Wilcoxon's test, one-tailed P-value: EMX-4 = 0.009, EMX-2 = 0.044, all loci combined = 0.016) and was nonsignificant at all loci in Anamalai and Periyar (one-tailed $P > 0.05$). However, allele frequency distributions were normal at all loci and no mode shift was observed in any population.

Discussion

Regional patterns of mtDNA and microsatellite DNA diversity

Low mitochondrial haplotype diversity was observed in southern India, with a total of five haplotypes, and overall haplotype diversity of 0.436 ± 0.029. More strikingly, we did not find any haplotype from the α(B) clade in southern India, contrary to previous studies with few samples from India, which showed a predominance of α(B) clade haplotypes on the mainland and β(A) haplotypes in Sri Lanka and the Sunda region of Southeast Asia (Fernando *et al*, 2000; Fleischer *et al*, 2001). The presence of only β(A) haplotypes in southern India and the predominance of α(B) haplotypes in the rest of the mainland and in southern Sri Lanka is intriguing. Trade in elephants as a possible explanation for the presence of β(A) haplotypes in southern India (Fleischer *et al*, 2001) is unlikely because all the observed haplotypes are from the β(A) clade, closely related, phylogenetically distal, and in three instances unique (Figure 2), suggesting that at least some of the haplotypes arose within the region, as opposed to the introduction of basal haplotypes that may have been widespread. It is possible that the β(A) clade was displaced by the α(B) clade colonizing the continent and persisted only in the more southern reaches of the range. The three α(B) clade haplotypes in Sri Lanka may then be introduced haplotypes. Alternatively, both clades could have coexisted in the subcontinent and Sri Lanka, and southern India as a whole may have suffered an ancient severe bottleneck after a/the α(B) haplotype(s) had moved into Sri Lanka, and subsequently lost the α(B) clade. A

Table 2 F_{ST} (below diagonal) and R_{ST} (above diagonal) values between populations based on microsatellite DNA

Pop	Nilgiris	Anamalai	Periyar
Nilgiris		0.046**	0.098***
Anamalai	0.071***		NS
Periyar	0.038***	NS	

Symbols ** and *** denote significant difference from zero at $P < 0.01$ and 0.001, respectively, and 'NS' refers to nonsignificance, based on permutation tests.

Table 3 F_{ST} (below diagonal) and R_{ST} (above diagonal) values between locations within Nilgiris, based on microsatellite DNA

	1	3	4	5	6	7	9	11	12	13
1		NS	NS	NS	NS	NS	NS	NS	0.666*	NS
3	NS		NS	NS	NS	NS	NS	NS	0.560*	NS
4	NS	NS		NS	NS	NS	NS	NS	NS	NS
5	NS	NS	NS		NS	NS	NS	NS	0.563*	NS
6	NS	NS	NS	NS		NS	NS	NS	NS	NS
7	NS	NS	NS	NS	NS		NS	NS	0.654*	NS
9	NS	NS	NS	NS	NS	NS		NS	NS	NS
11	NS	NS	NS	NS	NS	NS	NS		0.585*	NS
12	NS	NS	NS	NS	NS	NS	NS	0.236*		NS
13	NS	NS	NS	NS	NS	NS	NS	0.250*	NS	

Numbers correspond to the number of the location in Figure 1. Locations with fewer than three samples were excluded from the analyses. Comparisons with location 12 are based on five loci, since there was no data at locus EMX-4 from location 12. Asterisks indicate significance after Bonferroni corrections and 'NS' refers to values not significantly different from zero.

Figure 4 (a) Map of the E–W (1-5-6-9), NW-S (1-13-17), and NE-S (9-13-17) transects used to detect breaks in gene flow. Locations with fewer than three samples were not included in these transects. Location 15 was not used as it digressed from the line, but there was no difference in the results when it was included. (b–d) $(F_{ST}/(1-F_{ST}))$/geographic distance as calculated for pairs of adjacent locations, plotted against the cumulative mid-pair distances, along the E–W, NW-S, and NE-S transects, respectively. The positions of the locations are indicated by their numbers at the top of each figure, in the order in which the transect connected the locations. The abscissa of the black points indicate the position, in terms of geographic distance, of the mid-points between the pairs of locations shown above them. The lines joining these points are only for visualization and do not imply any statistical trend. Only statistically significant F_{ST}'s were used for the figures; a value of zero was used for F_{ST}'s that were not significantly different from zero.

sampling of central, northern, and northeastern India might help resolve this puzzle.

Hartl *et al* (1996) had reported an α(B) clade haplotype in two out of nine captive elephants from the state of Kerala, under which most of Periyar and part of Anamalai fall. This may be due to the difficulty in tracing the origins of captive animals, especially in the context of regular, sometimes illegal, trade, in elephants from central and northeastern India to Kerala (data from the forest department). Considering the geographic extent and size of our sampling, the α(B) clade, if it exists in southern India, must occur at very low frequency.

Nuclear diversity was low at tri- and tetranucleotide loci and moderate at dinucleotide loci. While there is no previous extensive population-level study of microsatellite variation in Asian elephants, the number of alleles per locus over five loci (EMX-1-4, LafMS03), using six to 20 individuals from nine range countries other than India, was 1.4 (Borneo)–3.6 (Sri Lanka) (average = 2.7) (based on Fernando *et al*, 2003b). The number of alleles per locus in southern India is 3.4. In addition, the number of alleles per dinucleotide locus in southern India (4.5) was comparable to those in African savannah elephant populations, of 2.6–5.6 (mean 4.4) alleles per locus across 16 dinucleotide loci (Comstock *et al*, 2002)

and 2.1–4.1 (mean 3.1) across four dinucleotide loci (Eggert *et al*, 2002). Thus, despite mitochondrial diversity being low in southern India, nuclear diversity appears to be at a 'normal' level.

Absence of differentiation and lower diversity within the Nilgiris

It was surprising that the Nilgiri population was fixed for a single mitochondrial haplotype and, although nuclear diversity was only a little lower than those of the other two southern Indian populations, did not show any substructuring at the nuclear loci either. The absence of nuclear differentiation is probably an outcome of large home ranges in this region (Baskaran *et al*, 1995) and significant movement of animals on an evolutionary time scale. On the other hand, fixation of an mtDNA haplotype may be a result of lineage sorting, social structure, or a population bottleneck. Lineage sorting is the persistence of a single lineage due to stochastic loss of female offspring in the other lineages, and is thus a function of the effective number of females (N_f) in a population (Avise *et al*, 1984). However, lineage sorting seems an unlikely explanation in the largest population of over 9000 Asian elephants, while Sri Lanka, with a population of 2000–5000 elephants, shows 12 haplotypes (Fernando *et al*, 2000), with five haplotypes in a single location of a few square kilometres (P Schikler, personal communication).

Social structure may influence the number of haplotypes in a population due to the matrilineal inheritance of mitochondria, and the matriarchal social system of elephants. There is some evidence for differences in social organization and ranging of elephants in southeastern Sri Lanka and the Nilgiri population in southern India, with elephants in these populations displaying a 'family' level of social organization with small home ranges (~55–120 km², Fernando and Lande, 2000), and a higher 'clan' level of organization (Sukumar, 1989, p 62; Baskaran *et al*, 1995) with large home ranges (>500 km², Baskaran *et al*, 1995), respectively. A clan consists of several 'bond groups' sharing a dry season range, each bond group in turn consisting of a few 'family groups' (adult females and their offspring) that show significant social association (Douglas Hamilton, 1972; Moss and Poole, 1983). While clan formation probably occurs over decades or centuries, changes in mitochondrial haplotypes arise over tens of thousands of years, resulting in the mitochondrial haplotypes of members of a clan being identical by descent. A dominance hierarchy among clans may lead to the marginalization of subordinate clans and hence the extinction of mitochondrial haplotypes through a 'selective sweep' (Maynard Smith and Haigh, 1974), selection acting on a linked behaviour rather than a linked gene. In contrast, independent family groups occupying small home ranges may support more mtDNA haplotypes. Comprehensive studies of elephant social organization covering the major populations in southern India and Sri Lanka would be useful to assess the role of social organization in creating the observed patterns of mitochondrial diversity.

A population bottleneck is another plausible explanation for the presence of a single mitochondrial haplotype in the Nilgiris. We were unable to detect evidence of a recent bottleneck based on microsatellite data. It is possible, however, that an ancient bottleneck might have occurred, which affected, or whose effects are presently seen on, only the mitochondrial genome, as it has only one-fourth the effective population size of the nuclear genome. Since the accumulation of a single mutation in the mtDNA sequence examined takes about 42 300 years on average, the upper bound on the period of such a bottleneck would be of this order. Palaeoecological studies using stable carbon isotope ratios in montane peat deposits have ascertained periods of intense aridity in the Nilgiris, at about 45 000 years before present (BP) and 18 000 years BP (the Last Glacial Maximum), with wetter conditions resuming only in the early Holocene, about 9000–8000 years BP (Sukumar *et al*, 1993). This aridity provides a possible cause for a population reduction in the Nilgiris, while a southerly range such as Anamalai or Periyar may have served as a wetter tropical Pleistocene refugium.

A phylogeographic barrier?

Although there was no substructuring within the Nilgiris, which spans almost 15 000 km², the Nilgiri population was differentiated from the more southerly populations, separated by just the 40-km-wide Palghat Gap, at both mitochondrial and microsatellite DNA. Forests in the Palghat Gap were probably first cleared by humans when farming was established in southern India ca. 2000BC (Allchin and Allchin, 1997, p 104), and it is only in the past few centuries that high-density human settlements would have excluded elephants from the gap. The mitochondrial differentiation corresponding to about 228 550 years and the absence of a recent bottleneck in the nuclear genome suggest that the differentiation across the gap must predate this human influence and, possibly, result from prehistoric vicariant events relating to the gap itself. The gap is thought to be a Precambrian shear zone with geologically recent faulting events (Rao *et al*, 2002). However, the role of a palaeo river system and/or marine incursion in shaping the gap have also been hypothesized (see Subramanian and Muraleedharan, 1985). The Palghat Gap is considered a phytogeographic barrier, the Nilgiri and Anamalai regions having dissimilar floristic compositions (Subramanyam and Nayar, 1974). In addition, six species of passerine birds have alternate subspecies distributions on either side of the gap (based on Ali and Ripley, 1987). There is, however, no previous study reporting the gap as having served as a barrier for large mammals. It is conceivable that the Palghat Gap was a biogeographic barrier in the past, as has been shown with ancient ridges of subbasins of the Amazon. Patton *et al* (1994) have proposed that it was those ridges, rather than the present-day riverine barrier, that have shaped the genetic structure of arboreal spiny rats. If the Palghat gap was indeed a barrier to gene flow, it is likely to have moulded the genetic structure of other independently evolving species (Avise *et al*, 1987), thus exhibiting a concordance across taxa. A study of genetic distribution patterns in other large mammals, as well as other classes of vertebrates and invertebrates, could provide greater insight into the phylogeographic history of this region.

Acknowledgements

References

Ali S, Ripley SD (1987). *Compact Handbook of the Birds of India and Pakistan: Together with those of Bangladesh, Nepal, Bhutan and Sri Lanka*, 2nd edn. Oxford University Press: New Delhi.

Allchin B, Allchin R (1997). *Origins of a Civilization: The Prehistory and Early Archaeology of South Asia*. Viking: New Delhi.

Asian Elephant Research and Conservation Centre (1998). *The Asian Elephant in Southern India: A GIS Database for Conservation of Project Elephant Reserves*. Asian Elephant Research and Conservation Centre: Bangalore.

Avise J, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE *et al* (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between genetics and systematics. *Annu Rev Ecol Syst* **18**: 489–522.

Avise JC, Neigel JE, Arnold J (1984). Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J Mol Evol* **20**: 99–105.

Baskaran N, Balasubramanian S, Swaminathan S, Desai AA (1995). Home range of elephants in the Nilgiri Biosphere Reserve, south India. In: Daniel JC, Datye HS (eds) *A Week with Elephants, Bombay Natural History Society*. Oxford University Press: Bombay, pp 296–313.

Bist SS (2002). An overview of elephant conservation in India. *Indian Forester* **128**: 121–136.

Comstock KE, Georgiadis N, Pecon-Slattery J, Roca AL, Ostrander EA, O'Brien SJ *et al* (2002). Patterns of molecular genetic variation among African elephant populations. *Mol Ecol* **11**: 2489–2498.

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

Douglas-Hamilton I (1972). On the ecology and behaviour of the African elephant: the elephants of Lake Manyara. PhD Thesis, University of Oxford.

Eggert LS, Rasner CA, Woodruff DS (2002). The evolution and phylogeography of the African elephant inferred from mitochondrial DNA sequence and nuclear microsatellite markers. *Proc R Soc Lond Ser B* **269**: 1993–2006.

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

Fernando P, Evans BJ, Morales JC, Melnick DJ (2001b). Electrophoresis artefacts: a previously unrecognized cause of error in microsatellite analysis. *Mol Ecol Notes* **1**: 325–328.

Fernando P, Lande R (2000). Molecular genetic and behavioural analysis of social organization in the Asian elephant (*Elephas maximus*). *Behav Ecol Sociobiol* **48**: 84–91.

Fernando P, Pfrender M, Encalada S, Lande R (2000). Mitochondrial DNA variation, phylogeography and population structure of the Asian elephant. *Heredity* **84**: 362–372.

Fernando P, Vidya TNC, Melnick DJ (2001a). Isolation and characterization of tri- and tetranucleotide microsatellite loci in the Asian elephant, *Elephas maximus*. *Mol Ecol Notes* **1**: 232–234.

Fernando P, Vidya TNC, Payne J, Stuewe M, Davison G, Alfred RJ *et al* (2003b). DNA analysis indicates that Asian elephants are native to Borneo and are therefore a high priority for conservation. *PloS Biol* **1**: 110–115.

Fernando P, Vidya TNC, Rajapakse C, Dangolla A, Melnick DJ (2003a). Reliable non-invasive genotyping: fantasy or reality? *J Hered* **94**: 115–123.

Fleischer R, Perry E, Muralidharan K, Stevens E, Wemmer C (2001). Phylogeography of the Asian elephant (*Elephas maximus*) based on mitochondrial DNA. *Evol Int J Org Evol* **55**: 1882–1892.

Frantzen MAJ, Silk JB, Ferguson JWH, Wayne RK, Kohn MH (1998). Empirical evaluation of preservation methods for faecal DNA. *Mol Ecol* **7**: 1423–1428.

Gene Codes Corporation (1999). *SEQUENCHER: A Genetic Analysis Software* Release 3.1.1. Gene Codes Corporation: Ann Arbor.

Goodman S (1997). RSTCALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Mol Ecol* **6**: 881–885.

Hartl G, Kurt F, Hemmer W, Nadlinger R (1995). Electrophoretic and chromosomal variation in captive Asian elephants (*Elephas maximus*). *Zoo Biol* **14**: 87–95.

Hartl G, Kurt F, Tiedemann R, Gmeiner C, Nadlinger K, Mar KU *et al* (1996). Population genetics and systematics of Asian elephants (*Elephas maximus*): a study based on sequence variation at the Cyt b gene of PCR-amplified mitochondrial DNA from hair bulbs. *Z Saugetierkd* **61**: 285–294.

Jukes T, Cantor C (1969). Evolution of protein molecules. In: Munro HN (ed) *Mammalian Protein Metabolism*. Academic Press: New York, pp 21–132.

Kohn M, Wayne RK (1997). Facts from feces revisited. *TREE* **12**: 223–227.

Leberg P (2002). Estimating allelic richness: effects of sample size and bottlenecks. *Mol Ecol* **11**: 2445–2449.

Luikart G, Allendorf F, Cornuet J-M, Sherwin W (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J Hered* **89**: 238–247.

Mantel N (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res* **27**: 209–220.

Maynard Smith J, Haigh J (1974). The hitch-hiking effect of a favourable gene. *Genet Res* **23**: 23–35.

Michalakis Y, Excoffier L (1996). A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**: 1061–1064.

Moss C, Poole J (1983). Relationships and social structure of African elephants. In: Hinde R (ed) *Primate Social Relationships*. Blackwell Science Publications: Boston, pp 315–325.

Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press: New York.

Nozawa K, Shotake T (1990). Genetic differentiation among local populations of Asian elephants. *J Zool Syst Evol Res* **28**: 40–47.

Nyakaana S, Arctander P (1998). Isolation and characterization of microsatellite loci in the African elephant, *Loxodonta africana*. *Mol Ecol* **7**: 1436–1437.

Patton JL, Da Silva MNF, Malcolm JR (1994). Gene genealogy and differentiation among arboreal spiny rats (Rodentia:

Echimyidae) of the Amazon Basin: a test of the riverine barrier hypothesis. *Evol Int J Org Evol* **48**: 1314–1323.

Piry S, Luikart G, Cornuet J-M (1997). *BOTTLENECK: A Program for Detecting Recent Effective Population Size Reductions from Allele Frequency Data*. Laboratoire de Modélisation et Biologie Evolutive: Montpellier.

Ramakrishnan U, Santosh JA, Ramakrishnan U, Sukumar R (1998). The population and conservation status of Asian elephants in the Periyar Tiger Reserve, southern India. *Curr Sci* **74**: 110–113.

Rao TKG, Rajendran CP, Mathew G, John B (2002). Electron spin resonance dating of fault gouge from Desamangalam, Kerala: evidence for Quaternary movement in Palghat gap shear zone. *Proc Indian Acad Sci (Earth and Planetary Sciences)* **111**: 103–113.

Raymond M, Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* **86**: 248–249.

Rice W (1989). Analysing tables of statistical tests. *Evol Int J Org Evol* **43**: 223–225.

Schneider S, Roessli D, Excoffier L (2000). *Arlequin: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Switzerland: Geneva.

Slatkin M (1995). A measure of population subdivision based on microsatellite allele frequency. *Genetics* **139**: 457–462.

StatSoft Inc. (2001). *STATISTICA for Windows (Data Analysis Software System)* StatSoft Inc.: Tulsa.

Subramanian KS, Muraleedharan MP (1985). Origin of the Palghat Gap in South India – a synthesis. *J Geol Soc India* **26**: 28–37.

Subramanyam K, Nayar MP (1974). Vegetation and phytogeography of the Western Ghats. In: Mani MS (ed) *Ecology and Biogeography in India*, Vol. 23. Dr W Junk Publishers: The Hague, pp 178–196.

Sukumar R (1989). *The Asian Elephant: Ecology and Management*. Cambridge University Press: Cambridge.

Sukumar R (2003). *The Living Elephants. Evolutionary Ecology, Behavior, and Conservation*. Oxford University Press: New York.

Sukumar R, Ramesh R, Pant R, Rajagopalan G (1993). A δ^{13}C record of late Quaternary climate change from tropical peats in southern India. *Nature* **364**: 703–706.

Sukumar R, Santiapillai C (1996). *Elephas maximus*: status and distribution. In: Shoshani J, Tassy P (eds) *The Proboscidea: Evolution and Palaeoecology of Elephants and their Relatives*. Oxford University Press: New York, pp 327–331.

Taberlet P, Waits LP, Luikart G (1999). Noninvasive genetic sampling: look before you leap. *TREE* **14**: 323–327.

Weir B, Cockerham C (1984). Estimating F-statistics for the analysis of population structure. *Evol Int J Org Evol* **38**: 1358–1370.

Appendix A

Allele frequencies of the six loci used in this study are given in (Table A1).

Table A1 Allele frequencies of the six loci

Locus	Allele	Nilgiris	Anamalai	Periyar
EMX-1	135	0.560	0.311	0.312
	144	0.000	0.033	0.000
	147	0.004	0.067	0.062
	150	0.436	0.589	0.625
EMX-2	219	0.437	0.216	0.231
	225	0.563	0.784	0.769
EMX-3	238	0.093	0.056	0.083
	254	0.907	0.944	0.917
EMX-4	351	0.309	0.524	0.625
	375	0.272	0.098	0.156
	387	0.419	0.378	0.219
LafMS02	131	0.004	0.000	0.021
	133	0.046	0.125	0.062
	135	0.313	0.400	0.458
	137	0.401	0.400	0.354
	139	0.000	0.075	0.042
	141	0.237	0.000	0.062
LafMS03	135	0.007	0.000	0.000
	137	0.426	0.624	0.667
	139	0.346	0.024	0.071
	149	0.121	0.094	0.095
	151	0.029	0.176	0.143
	155	0.070	0.082	0.024