Molecular Ecology (2004) **13**, 2819–2827 doi: 10.1111/j.1365-294X.2004.02288.x

Crossing the Red Sea: phylogeography of the hamadryas baboon, *Papio hamadryas hamadryas*

BRUCE J. WINNEY,*†‡§ ROBERT L. HAMMOND,*†‡§¶ WILLIAM MACASERO,‡ BENITO FLORES,‡ AHMED BOUG,** VERONIQUE BIQUAND,†† SYLVAIN BIQUAND‡‡ and MICHAEL W. BRUFORD§§
**Institute of Zoology, Zoological Society of London, Regents Park, London NW1 4RY, UK, †Conservation Programmes, Zoological Society of London, Regents Park, London NW1 4RY, UK, ‡King Khalid Wildlife Research Centre, National Commission for Wildlife Conservation and Development, PO Box 61681, Riyadh 11575, Saudi Arabia, **National Commission for Wildlife Conservation and Development, PO Box 61681, Riyadh 11575, Saudi Arabia, ††Université de Rennes I, Station Biologique, F-35380 Paimpont, France, ‡‡Abilis Ecologie, 40 rue de la montagne Ste Genevieve, Paris 75005, France, §§Cardiff School of Biosciences, Main Building, Museum Avenue, PO Box 915, Cardiff CF10 3TL, UK*

Abstract

The hamadryas baboon (*Papio hamadryas hamadryas*) is found both in East Africa and western Arabia and is the only free-ranging nonhuman primate in Arabia. It has been hypothesized that hamadryas baboons colonized Arabia in the recent past and were possibly even transported there by humans. We investigated the phylogeography of hamadryas baboons by sequencing a portion of the control region of mtDNA in 107 baboons from four Saudi Arabian populations and combing these data with published data from Eritrean (African) *P. h. hamadryas*. Analysis grouped sequences into three distinct clades, with clade 1 found only in Arabia, clade 3 found only in Africa, but clade 2 found in both Arabian and African *P. h. hamadryas* and also in the olive baboon, *P. h. anubis*. Patterns of variation within Arabia are neither compatible with the recent colonization of Arabia, implying that baboons were not transported there by humans, nor with a northerly route of colonization of Arabia. We propose that hamadryas baboons reached Arabia via land bridges that have formed periodically during glacial maxima at the straits of Bab el Mandab in the southern Red Sea. We suggest that the genetic differentiation of Arabian from African populations suggests that Arabian populations have a higher conservation status than recognized previously.

Keywords: baboon, colonization, glaciation, *Papio hamadyas hamadryas*, phylogeography, population structure

Received 21 April 2004; revision received 16 June 2004; accepted 16 June 2004

Introduction

Climate fluctuations, particularly since the Pleistocene, have been highly influential in shaping the distribution and population genetics of many modern animals and plants (Hewitt 2000). To date, the majority of studies have focused on European and North American species, but climate change has also had important consequences elsewhere in the world (Trauth *et al*. 2003). The large mammal fauna of Arabia has been influenced by Africa and a number of species have ranges that encompass both Arabia and Africa (e.g. leopard, striped hyena, honey badger, caracal, common genet; Harrison & Bates 1991) or there are endemic Arabian species that are related closely to African species (e.g. Arabian oryx, Saudi gazelle; Harrison & Bates 1991). Africa and Arabia are separated by the narrow and shallow straits of Bab el Mandab in the southern Red Sea and land-bridges are known to have periodically formed during glacial maxima (Rohling *et al*. 1998; Siddall *et al*. 2003). These land-bridges provide a potential route for the colonization of

Correspondence and current address: Bruce Winney, Cancer and Immunogenetics Group, First Floor Laboratory, Department of Clinical Pharmacology, Radcliffe Infirmary, Woodstock Road, Oxford OX2 6HE, UK. Fax: 01865 791712;
E-mail: bruce.winney@clinpharm.ox.ac.uk
¶Current address: Department of Ecology and Evolution, Bâtiment de Biologie, University of Lausanne, Lausane CH-1015, Switzerland.
§Bruce J. Winney and Robert L. Hammond are joint first authors of this study.

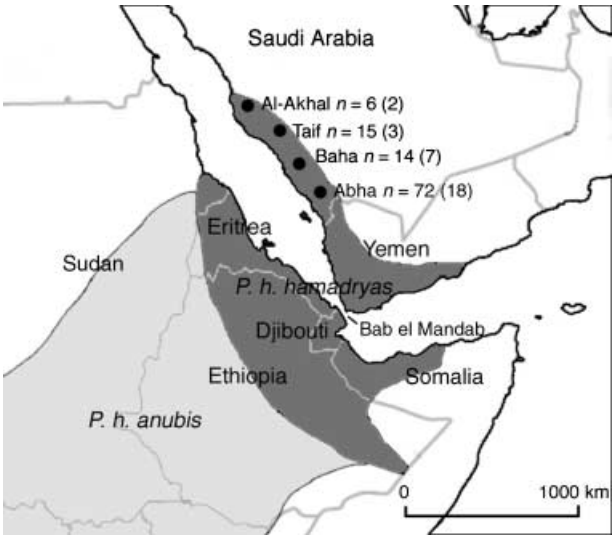

Fig. 1 Map of sampling locations of *P. h. hamadryas* in Saudi Arabia, the approximate range of *P. h. hamadryas* within Arabia and Africa and the eastern range of *P. h. anubis* (after Hapke *et al*. 2001); '*n*' = individuals typed, numbers in brackets indicate number of haplotypes observed.

Arabia by African species (or vice versa) and for genetic exchange between species whose ranges encompass both sides of the Red Sea. Given that sea-level changes in the southern Red Sea are known with high precision (Siddall *et al*. 2003) the comparison of populations in Arabia and east Africa offers a good opportunity to relate biogeography to physical processes.

The hamadryas baboon, *Papio hamadryas hamadryas*, is found in both East Africa and the Arabian Peninsula (Fig. 1) and is the only species of nonhuman primate found in the wild in Arabia (Harrison & Bates 1991; Biquand *et al*. 1992; Kingdon 1997). Morphologically and behaviourally, Arabian hamadryas baboons are very similar to African hamadryas (Kummer *et al*. 1985). It is not known, however, whether this phenotypic similarity is because Arabia has been colonized only in the very recent past or whether baboon populations in Arabia are older and phenotypic similarity has been maintained by stabilizing selection. Recent ancestry is quite possible as sea-levels were low enough 18 000 years ago to expose a land-bridge (Rohling *et al*. 1998; Siddall *et al*. 2003). There is also the intriguing possibility that humans may have transported hamadryas baboons between Arabia and Africa approximately 4500 years ago, as ancient Egyptians are known to have worshipped them, considering them to be incarnations of Thoth, the god responsible for weighing the souls of the dead (Kummer 1995).

A number of hypotheses have been suggested to explain the disjunct range of *P. h. hamadryas* (e.g. Kummer *et al*. 1981). The simplest suggests that in the very recent past African *P. h. hamadryas* colonized Arabia (Kummer 1995). Another, more complex, hypothesis suggests that Arabia was colonized much earlier by an ancestor of modern hamadryas baboons which, after evolving into *P. h. hamadryas* in isolation in Arabia, expanded its range back across the Red Sea into East Africa (T. Shotake, personal communication). In addition to questions of timing there is also the route of colonization to consider. This may have occurred either by a southerly route, via the Straits of Bab el Mandab, or by a northerly route, via Sinai (Kummer 1995). These different hypotheses are expected to give rise to different phylogeographical patterns in mitochondrial DNA (e.g. Avise 1994). If colonization of Arabia was recent, Arabian mtDNA haplotypes should be a subset of African haplotypes. Conversely, if Arabia was the source of African hamadryas one would predict the opposite pattern. Similarly, a northerly route of colonization predicts that populations in the north of the range on either side of the Red Sea should be more similar genetically than those on opposite sides of the straits of Bab el Mandab.

Recent work on Eritrean *P. h. hamadryas* populations (Hapke *et al*. 2001) has shown that patterns of mtDNA variation support behavioural observations (Stammbach 1987) that females, rather than males, disperse from natal groups. Hapke *et al*. (2001) also provided evidence that mtDNA from *P. h. anubis* had introgressed into *P. h. hamadryas*. We have analysed patterns of mtDNA variation in 107 hamadryas baboons from four populations sampled from western Saudi Arabia. Furthermore, we have investigated the evolutionary relationships between African and Arabian hamadryas baboons, by analysing our newly gathered data together with that of Hapke *et al*. (2001).

Materials and methods

Sample collection

Tissue and blood samples were collected from 298 individuals from four sites (Abha, $n = 244$; Baha, $n = 23$; Taif, $n = 25$; Al-Akhal, $n = 6$) in the Asir mountain range in the west of Saudi Arabia (Fig. 1). Samples from Baha, Taif and Al-Akhal were taken opportunistically from the local baboon population; however, in Abha 32% of the sample was trapped as complete, or near-complete, one-male units (OMUs, $n = 7$) as part of an investigation of social structure (Hammond *et al*. in preparation). In this study, the collections from Abha ($n = 72$), Baha ($n = 14$) and Taif ($n = 15$) were subsampled on an *ad hoc* basis for mitochondrial DNA sequencing. In the Abha subsample, 44% of individuals were from known OMUs. As there was no significant difference in frequencies of the two major clades (see Results section) between known and unknown OMUs (Fisher's exact test, $P = 0.427$), all Abha individuals were included in the following analyses. Baboons were trapped and samples collected with the permission of the Saudi

Arabian National Commission for Wildlife Conservation and Development (details of trapping methods are provided in Hammond *et al*. in preparation).

Molecular methods

DNA was extracted from blood and tissue samples using standard SDS/proteinase K digestion, phenol/chloroform extraction and ethanol precipitation (Ausebel *et al*. 1996). The complete control region was amplified by polymerase chain reaction (PCR) from 10 individuals from Abha ($n = 6$), Taif ($n = 2$) and Al-Akhal ($n = 2$) using the primers H15439 (5′-CTGGC GTTCT AACTT AAACT AC-3′) and L22 (5′-GGCAT TTTCA GTGTC TTGCT TTGG-3′) that were designed from a published *P. hamadryas* ssp. sequence (GenBank Accession no. Y18001). From an alignment of these sequences two primers, H15613 (5′-CATAC TTACC CTCAA TCCAT AAATG-3′) and L15910 (5′-AGAAC CAGAT GCCGG ATACA GTTC-3′) were designed that amplified a 345 base pairs (bp) portion of the hypervariable I region (Saccone *et al*. 1991), which encompassed the most variation. This region was amplified and sequenced manually in 107 individuals to give an overall alignment of 279 bp.

Although nuclear transposition of mtDNA sequences (numts) is a common phenomenon (e.g. Collura & Stewart 1995; Thalman *et al*. 2004) numts were not evident within our data set, for the following reasons. First, all PCRs gave single, cleanly amplified, products and no additional bands were seen in sequencing autoradiographs. Second, high levels of sequence variation were found throughout the data set (see Results), which were more compatible with rapid evolution in the mitochondrial control region than the 10 (Ward *et al*. 1991) to fourfold (Lundstrom *et al*. 1992) slower evolution for noncoding regions in the nuclear genome. Third, and most importantly, our sequences aligned with no deletions or insertions to the published *P. hamadryas* mtDNA genome and this sequence was derived from clones of isolated mtDNA (GenBank Accession no. Y18001; Arnason *et al*. 1998).

Analysis

For the Saudi Arabian data, haplotype frequency differences were compared by exact tests using GENEPOP 3.1 (Raymond & Rousset 1995). Pairwise F_{ST} values were estimated between populations based on either haplotype frequencies (Weir & Cockerham 1984) using GENEPOP 3.1 or on Kimura's two-parameter genetic distance (Kimura 1980) with a gamma correction estimating a heterogeneous mutation rate at different sites ($\alpha = 0.16$) using ARLEQUIN 2.0 (Excoffier *et al*. 1992). To investigate recent population dynamics, pairwise sequence mismatch distributions (Slatkin & Hudson 1991) were generated for the four Saudi Arabian populations and a minimum spanning network (Excoffier & Smouse 1994) was estimated using TCS (Clement *et al*. 2000) for the Saudi Arabian haplotypes.

To investigate phylogeographical relationships between Arabian and African hamadryas baboons we compared our Arabian sequences with published sequences from Eritrean *P. h. hamadryas* and *P. h. anubis* (Hapke *et al*. 2001; GENBANK Accession nos AF275383–AF275475) and a sequence from an African baboon of unknown geographical origin (GENBANK Accession no. Y18001 *et al*. 1998). The resulting alignment gave a 168 bp region of overlap. Maximum likelihood estimates of distances, based on Felsenstein's (1984) model of evolution, were used to construct unrooted neighbour-joining trees and maximum likelihood trees using PAUP* 4.0 (Swofford 2001). Both analyses assumed a heterogeneous distribution of base substitutions, for which α (the shape parameter of the gamma distribution of substitution rate) was estimated to be 0.13 (using PAUP* 4.0). The robustness of phylogenetic patterns was tested using 1000 bootstrap replicates for neighbour-joining trees and 100 replicates for maximum likelihood trees. Further unrooted trees were estimated by maximum parsimony, using PAUP* 4.0. Both branch-and-bound and 'fast' stepwise approaches were used to construct trees with 1000 bootstrap replicates. An analysis of molecular variance (AMOVA; Excoffier *et al*. 1992) was used to investigate the partitioning of genetic variation within and between Saudi Arabia and Eritrea.

Results

Genetic structure among Saudi Arabian populations

Sequencing of 279 bp of hypervariable I region of mtDNA d-loop in 107 Saudi Arabian baboons revealed 26 haplotypes (Table 1; GENBANK Accession nos: AY247443–AY247551). Within this region there were 35 variable sites and all substitutions were transitions. Sequencing of a longer 442 bp region in 10 individuals revealed 29 variable positions, of which only three variable sites outside the core 279 bp region, indicating that 90% of the variation was captured within the shorter 279 bp region. For the 279 bp region, 92% of haplotypes were sampled in single populations (Table 1), with only two haplotypes (haplotypes 1 and 2, Table 1) shared by two or more populations. All but three single population haplotypes were rare with individual frequencies less than 10%, but within populations the combined frequency of such haplotypes was high (60% in Abha, 50% in Baha, 60% in Taif and 33% in Al-Akhal). Overall haplotype frequencies were significantly different among sites (exact tests, all $P < 0.001$). Pairwise F_{ST} values calculated from haplotype frequencies varied from 0.029 (Abha–Baha) to 0.292 (Taif–Al-Akhal) and all differed significantly from zero except Abha–Baha. Pairwise F_{ST}

Table 1 *P. h. hamadryas* mtDNA haplotypes frequencies (279 bp of the d-loop region) in four Saudi Arabian populations; haplotypes in bold type (5, 8–10) belong to clade 2, whereas all other haplotypes belong to clade 1

	Haplotype																										
Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	Total
Abha	20	11	7	0	16	0	0	1	1	2	1	1	1	2	1	1	1	1	1	1	1	0	0	0	2	0	72
Baha	4	3	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	2	14
Taif	5	1	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15
Al-Akhal	4	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
Total	33	15	7	9	16	2	2	1	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1	1	2	2	107

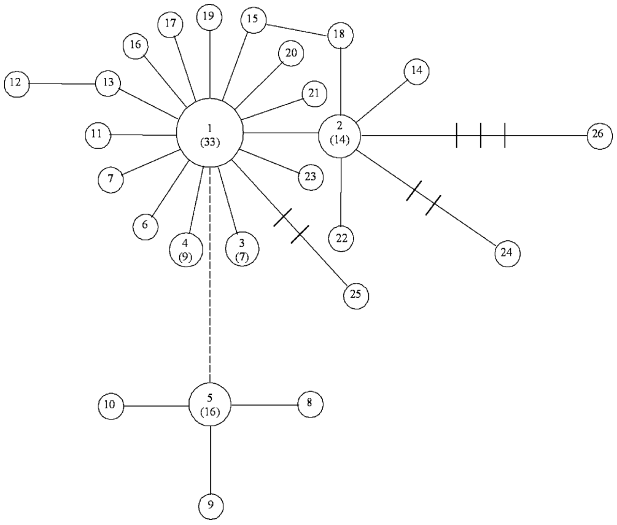

Fig. 2 Spanning network of 26 haplotypes derived from 279 bp of d-loop mtDNA sequence sequenced from Saudi Arabian *P. h. hamadryas*. Clade 1 (haplotypes 5, 8, 9, 10) and clade 2 (all other haplotypes) are separated by 15 substitutions represented by a dotted line. Numbers within nodes but without parentheses refer to haplotype designations in Table 1, whereas those within parentheses are the total number of individuals with that haplotype. Haplotypes where no totals are given were found in one or two individuals. Line lengths correspond to one, two (with two ticks) or three (with three ticks) mutations.

calculated from Kimura's two parameter genetic distances varied from 0.086 (Abha–Al-Akhal) to 0.399 (Taif–Al-Akhal) and all differed significantly from zero, except Abha–Al-Akhal.

Construction of a minimum spanning network (Fig. 2) revealed two highly distinct clades (1 and 2). Clade 1 comprised 22 haplotypes that differed by a maximum of six substitutions (2.2%), whereas clade 2 comprised four haplotypes (5, 8, 9, 10) that differed by a maximum of two substitutions (0.7%). Representative haplotypes from groups 1 (haplotypes 1–4) and 2 (haplotype 5) differed by 15–16 mutations (5.4–5.7%). Pairwise sequence mismatch distributions of the haplotypes formed smooth bell-shaped distributions (not shown) when each group was considered separately. All individuals sampled in Baha, Taif and Al-Akhal had clade 1 mtDNA whereas in Abha 72% of individuals had clade 1 and 28% clade 2 mtDNA. The frequency of clade 2 in Abha was significantly different to that at Baha, Taif, and Baha–Taif–Al-Akhal combined (Fisher's exact tests, all $P < 0.05$). Combined sample size at Baha, Taif and Al-Akhal ($n = 35$) gave more than 80% power to detect clade 2 at a frequency of greater than 5%.

Relationships between Saudi Arabian and Eritrean populations

Alignment of Arabian and Eritrean (Hapke *et al*. 2001) sequences gave a 168 bp region of overlap. Within this region there were 45 variable sites and all were transitions. Phylogenetic analysis using unrooted trees revealed three major clades supported by bootstrap values greater than 70% (Fig. 3). Clade 1, the most common in Saudi Arabia, was found only in Arabia. Clade 2, only found in Abha in Saudi Arabia, matched Hapke *et al*.'s (2001) clade B, which comprised haplotypes from Eritrean *P. h. hamadryas* and Eritrean *P. h. anubis*. Clade 3 matched Hapke *et al*.'s (2001) clade A, which included only Eritrean *P. h. hamadryas*. Importantly, our analysis recovered the same major lineages that were identified from analyses of longer sequences (this study and Hapke *et al*. 2001). This suggests the phylogenetic patterns described here (Fig. 3), although based upon shorter sequences, are robust.

Within these clades mean maximum likelihood genetic distances were 0.016 (range 0.000–0.038) in clade 1, 0.049 (range 0.000–0.177) in clade 2 and 0.023 (range 0.000–0.077) in clade 3. In contrast, the mean genetic distance among clades was three to 10 times greater at 0.163 (range 0.103–0.434). Sequence Y18001 (Arnason *et al*. 1998) fell within clade 1 but had a mean genetic distance of 0.195 from other haplotypes within the clade (range, 0.128–0.240). This distance is greater than that found within clades and is more inline with the distances among clades. We therefore considered Y18001 to lie outside clade 1.

Fig. 3 Unrooted neighbour-joining tree based on 168 bp of d-loop mtDNA sequence from Saudi Arabian *P. h. hamadryas* and from Eritrean *P. h. hamadryas* and *P. h. anubis* (Hapke *et al*. 2001). All methods gave the same major clades. Numbers refer to haplotypes, letters to sample site (A = Abha, B = Baha, T = Taif, N = Al-Akhal) and numbers in parentheses are sample sizes for each haplotype. Eritrean haplotypes are identified by their GenBank Accession nos. Clade 1 includes only Arabian haplotypes, clade 2 includes haplotypes from Abha and Hapke *et al*.'s (2001) clade B and clade 3 corresponds to Hapke *et al*.'s (2001) clade A.

To estimate approximate divergence dates, mutation rates estimated from human and chimp d-loop sequences (Jensen-Seaman & Kidd 2001) were used. The Jensen-Seaman & Kidd (2001) mutation rate is based on a transversion mutation rate converted to a transition mutation rate using a human–chimp transition–transversion (ts/tv) ratio of 25. These data suggest that the ts/tv ratio is higher within the hypervariable I region of *P. hamadryas* as 45 transitional changes and no transversions were observed. In spite of this a ts/tv ratio of 25 was used. Based on mean pairwise distances a date of divergence for all clades from a most recent common ancestor was estimated to be 316–443 kyr, and divergence dates within clades 1, 2 and 3 were estimated as 85–119 kyr, 156–219 kyr and 86–120 kyr, respectively. It must be stressed that these dates are very approximate with unknown but large confidence intervals.

Analysis of molecular variance (AMOVA) with two different structures was used to investigate the partitioning of genetic variation (Table 2). In analysis A, 34% of the variation was accounted for within populations, 29% among populations within either Eritrea or Saudi Arabia and the largest faction of variation (44%) was between Eritrea and

Table 2 Results of analysis of molecular variance (AMOVA). In analysis A genetic variation was partitioned within populations, among populations within Eritrea (E) and Saudi Arabia (SA) and between E and SA. In analysis B populations were subdivided into clades. All theta values were significantly different from zero

Source of variation	Analysis	d.f.	Variance components	Theta	% variation
Between E and SA	A	1	6.11	0.440	44.04
	B	1	5.67	0.408	40.78
Among populations	A	14	3.10	0.398	22.29
Among populations subdivided by clade	B	23	7.38	0.897	53.11
Within populations	A	182	4.67	0.663	33.67
Within clades within populations	B	173	0.85	0.939	6.11

Saudi Arabia. In analysis B, we subdivided populations into their major clades. With this structure, 6% of variation was within haplotypes within populations, 53% among populations (subdivided by haplotype) within either Eritrea or Saudi Arabia and 41% was between Eritrea and Saudi Arabia.

Discussion

Our data show some striking and unexpected results. First, there are two highly divergent clades within Saudi Arabia (clades 1 and 2), but within each clade there is limited sequence divergence. Second, these two clades show an unexpected geographical distribution, with clade 1 found in all sampled Arabian populations, but clade 2 restricted to our southernmost sample, Abha. Finally, the most common clade in Saudi Arabia, clade 1, does not match either of the two major clades described previously from Africa, although clade 2 haplotypes are found in Arabian and African hamadryas and also in African olive baboons, *Papio hamadryas anubis* (Hapke *et al*. 2001).

Population structure within Saudi Arabia

Patterns of genetic variation within Saudi Arabia were compatible with both low and higher levels of gene flow. Limited gene flow is suggested by significant pairwise F_{ST} values, significant differences in allele frequency among populations and the restriction of clade 2 to Abha, in spite of sample sizes giving sufficient power to detect clade 2 north of Abha. The distribution of clade 1 haplotypes, however, suggests there may have been higher levels of gene flow in the past. Within clade 1, haplotype 1 is the most, or secondmost (Taif), common haplotype in all four populations (frequency range: 28–66%), and the similarity in frequency is particularly striking in the larger samples of Abha, Baha and Taif (frequency range: 28–33%). Similarly, haplotype 2 is the second most common clade 1 haplotype in Abha, Baha and Taif.

Female-biased gene flow is expected to lead to low levels of differentiation at mtDNA among populations (e.g. Morin *et al*. 1994; Piertney *et al*. 2001). In hamadryas, dispersal is known to be female-biased (Stammbach 1987) and the lack of mtDNA population structure found in Eritrea (Hapke *et al*. 2001) suggests female-biased gene flow. In contrast to patterns in Eritrea, we found higher levels of genetic differentiation in Arabia. This may reflect intrinsic differences in female dispersal between African and Arabia; however, differences in the geographical scale of sampling in Eritrea and Arabia is a more likely explanation, as the average distance between Arabian samples (178 km) was similar to the maximum distance between samples in Eritrea (202 km). Isolation-by-distance, as found in other female dispersing species (e.g. the greater prairie chicken, Johnson *et al*. 2003), would explain this disparity, and in support Hapke *et al*. (2001) reported a significant positive relationship between genetic and geographical distance within Eritrea.

Relationships between African and Arabian hamadryas baboons

We would like to know the direction, timing, route and frequency of interchanges between hamadryas baboons in Africa and Arabia. The direction of colonization, from Africa to Arabia or vice versa, ultimately collapses into one of timing. This is because baboons almost certainly evolved in Africa, rather than Arabia, given the much greater diversity and range of baboon species in Africa compared to Arabia (Jolly 2001). Our main question is, therefore, how long have hamadryas baboons been present in Arabia? In Kummer's (1995) view the behavioural and morphological similarity between African and Arabian hamadryas (Kummer *et al*. 1981) suggested that Arabian hamadryas colonized Arabia less than 10 000 years ago. Furthermore, a role for human transportation in the establishment of hamadryas in Arabia also implies that the colonization of Arabia was probably within the last 5000 years (Kummer 1995).

Colonization during the last 10 000 years predicts that haplotypes should be shared between Africa and Arabia. Our data do not support this prediction. For clades 1 and 2, the average divergence between Arabian and African sequences is 0.066 and 0.033, respectively, which gives very approximate dates of divergence of 316–443 kyr for clade 1 and 156–219 kyr for clade 2 (minimum levels of divergence of 0.018 for clade 2 give divergence dates of 85–120 kyr). This suggests that hamadryas baboons have been present in Arabia for considerably longer than 10 000 years and that human transportation was not responsible for establishing *P. h. hamadryas* in Arabia. We are cautious to reject of the recent colonization hypothesis, as we only have genetic data from the north of the African range (Hapke *et al*. 2001) and we do not have samples from southern Arabian (Yemeni) populations. We cannot therefore totally rule out the possibility that either clade 1 haplotypes or Arabian clade 2 haplotypes will not be found in as yet unsampled hamadryas populations in the eastern and southern African range. Interestingly, sequence Y18001, from a hamadryas baboon from an unknown location (Arnason *et al*. 1998), is genetically distinct from clades 1, 2 and 3. This suggests that there is indeed more variation in mtDNA within hamadryas baboons and there are likely to be further mtDNA clades.

If clade 1 is absent from throughout the African range, a more ancient colonization of Arabia would be likely. The minimum spanning network for clade 1 is star-like and the mismatch distribution is bell-shaped and smooth. These characteristics suggest that variation within clade 1 is fairly recent (divergence dates, 85–119 kyr) and that the population has been expanding (Slatkin & Hudson 1991; Rogers & Harpending 1992). If hamadryas baboons had been isolated for long periods in Arabia one might expect deeper branches within clade 1, but this is not evident. A counter-argument is that postcolonization populations of hamadryas remained small and localized for a long period, or that the Arabian population has been through a population bottleneck and lost genetic variation (Nei *et al*. 1975). This seems likely, as although hamadryas baboons are arid adapted (Zurovsky & Shkolnik 1982, 1993), their current range within Saudi Arabia is linked tightly to patterns of rainfall, with hamadryas absent from regions with an annual rainfall of less than 100 mm per year (Biquand *et al*. 1992). During glacial maxima climates were much dryer (Yan & Petit-Maire 1994) and it is almost certain that the range of hamadryas in Arabia during the last glacial maximum (18 kyr ago; Siddall *et al*. 2003) was reduced considerably compared to that today, and that consequently populations were much smaller in size. The genealogical patterns we now observe are likely to be signatures of a recent population expansion within Arabia.

Kummer (1995) hypothesized that a northerly route of colonization, via Sinai, was more likely than a southerly route across the Red Sea, via the straits of Bab el Mandeb. His reasons were that the straits, although dry during glacial maxima (Rohling *et al*. 1998; Siddall *et al*. 2003), would be salt flats and therefore an effective barrier to baboon dispersal. Regardless of direction, a recent colonization by a northerly route predicts genetic similarity between populations at the northern edges of the range on either side of the Red Sea. Our analysis does not support this prediction, as Eritrean and northern Saudi Arabian samples (Al-Akhal, Taif and Baha) are genetically distinct from one another. This suggests that a southerly route, via a land bridge, is most probable. Land bridges are thought to have formed five times during the past 500 kyr at approximately 18, 130, 270, 340 and 440 kyr ago (Rohling *et al*. 1998; Siddall *et al*. 2003). Average levels of variation within clades suggests that colonization during the most recent glacial maximum 18 kyr ago is unlikely. If clade 1 has evolved since colonization of Arabia, the most recent common ancestor in Arabia was 85–119 kyr ago, a date which matches closely the second-most recent sea-level low-stand (Rohling *et al*. 1998; Siddall *et al*. 2003). It must be stressed, however, that mutation rate estimates for the control region vary wildly (e.g. Parsons *et al*. 1997; Jazin *et al*. 1998; Pesole *et al*. 1999; Jensen-Seaman & Kidd 2001) and this variation, coupled with the stochastic nature of mutation and lineage sorting, means that any divergence times have large, and unknown, confidence limits. In view of this, any match to known sea-level low-stands must be treated with some skepticism.

The restriction of clade 2 to Abha is unlikely to be an artefact of sampling (see above). One explanation is that clade 2 has been lost from Baha, Taif and Al-Akhal by genetic drift or, if hamadryas baboons colonized Arabia from the south, clade 2 was lost during founder events as the species expanded its range northwards. A further scenario is that baboons with clade 2 mtDNA arrived in Arabia after the initial colonization of Arabia by baboons with ancestral clade 1 mtDNA. Distinguishing between these hypotheses is difficult, but the similarity of Arabian and African clade 2 haplotypes (minimum difference, 0.018) compared to higher variation within clade 1 and clade 1 being found only in Arabia, suggest that a later colonization of Arabia by clade 2 bearing hamadryas baboons is most likely.

Implications for conservation

Conservation is an important global issue and one of growing concern within Saudi Arabia (Abuzinada 2003). Prior to these genetic analyses hamadryas baboons were thought to have been recent colonists of Arabia (Kummer 1995), and therefore perhaps of low conservation status. Although the mtDNA gene tree does not meet the criterion of reciprocal monophyly, as defined by Moritz (1994),

we suggest that our findings raise the conservation status of Arabian hamadryas for two reasons. First, hamadryas baboons have been long-term residents and natural colonists of Arabia. Second, our AMOVA analysis shows that a considerable amount of genetic variation is found within Arabia that is not found within Africa. In addition, hamadryas baboons are the only primate species native to the Arabian Peninsula. Currently, hamadryas populations in Arabia are thriving and the species is numerous enough to be in conflict with human populations (Biquand *et al*. 1994). Unchecked persecution, however, has led to the decline (e.g. Arabian oryx, Stanley-Price 1989; gazelles, Thouless *et al*. 1991) and extinction (e.g. Saudi gazelle, Hammond *et al*. 2001; Arabian ostrich, Robinson & Matthee 1999) of many of Arabia's large vertebrates. A careful management programme is required, therefore, to ensure the long-term survival of hamadryas in Arabia, while minimizing human–baboon conflict.

Acknowledgements

We would like to thank Lori Lawson Handley and two anonymous referees for helpful comments on a previous version of the manuscript. This work was funded by the National Commission for Wildlife Conservation and Development (NCWCD), Saudi Arabia and we would like to thank the secretary general of NCWCD, Dr Abdulaziz Abu Zinada, for his support.

References

Abuzinada AH (2003) The role of protected areas in conserving biological diversity in the kingdom of Saudi Arabia. *Journal of Arid Environments*, **54**, 39–45.

Arnason U, Gullberg A, Janke A (1998) Molecular timing of primate divergences as estimated by two nonprimate calibration points. *Journal of Molecular Evolution*, **47**, 718–727.

Ausebel FM, Brent R, Kingston RE *et al*. (1996) *Current Protocols in Molecular Biology*, vol. 1. John Wiley and Sons Inc., USA.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, London.

Biquand S, Biquand-Guyot V, Boug A, Gautier JP (1992) The distribution of *Papio hamadryas* Saudi Arabia: ecological correlates and human influence. *International Journal of Primatology*, **13**, 223–243.

Biquand S, Boug A, Biquand-Guyot V, Gautier JP (1994) Management of commensal baboons in Saudi Arabia. *Revue d'Ecologie (Terre et Vie)*, **49**, 213–222.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1600.

Collura RV, Stewart C-B (1995) Insertions and duplications of mtDNA in the nuclear genomes of Old World monkeys and hominids. *Nature*, **378**, 485–492.

Excoffier L, Smouse PE (1994) Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. *Genetics*, **136**, 343–359.

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.

Felsenstein J (1984) Distance methods for inferring phylogenies: a justification. *Evolution*, **38**, 16–24.

Hammond RL, Macasero W, Flores B, Mohammed OB, Wacher T, Bruford MW (2001) Phylogenetic reanalysis of the Saudi gazelle and its implications for conservation. *Conservation Biology*, **15**, 1123–1133.

Hapke A, Zinner D, Zischler H (2001) Mitochondrial DNA variation in Eritrean hamadryas baboons *Papio hamadryas hamadryas*: life history influences population genetic structure. *Behavioural Ecology and Sociobiology*, **50**, 483–492.

Harrison DL, Bates PJJ (1991) *The Mammals of Arabia*. Harrison Zoological Museum, Sevenoaks, Kent.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Jazin E, Soodyall H, Jalonen P, Lindholm E, Stoneking M, Gyllensten U (1998) Mitochondrial mutation rate revisited: hot spots and polymorphism. *Nature Genetics*, **18**, 109–110.

Jensen-Seaman MI, Kidd KK (2001) Mitochondrial DNA variation and biogeography of eastern gorillas. *Molecular Ecology*, **10**, 2241–2248.

Johnson JA, Toepfer JE, Dunn PO (2003) Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. *Molecular Ecology*, **12**, 3335–3347.

Jolly CJ (2001) A proper study of mankind: analogies from the papionin monkeys and their implications for human evolution. *Yearbook of Physical Anthropology*, **44**, 177–204.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

Kingdon J (1997) *The Kingdon Field Guide to African Mammals*. Academic Press/Harcourt Brace, New York.

Kummer H (1995) In: *Quest of the Sacred Baboon: a Scientist's Journey*. Princeton University Press, Princeton.

Kummer H, Banaja AA, Abo-Khatwa AN, Ghandour AM (1981) A survey of hamadryas baboons in Saudi Arabia. *Fauna of Saudi Arabia*, **3**, 441–471.

Kummer H, Banaja AA, Abo-Khatwa AN, Ghandour AM (1985) Differences in social behavior between Ethiopian and Arabian hamadryas baboons. *Folia Primatologica*, **45**, 1–8.

Lundstrom R, Tavare S, Ward RH (1992) Estimating substitution rates from molecular-data using the coalescent. *Proceedings of the National Academy of Sciences USA*, **89**, 5961–5965.

Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, Woodruff DS (1994) Kin selection, social structure, gene flow and the evolution of chimpanzees. *Science*, **265**, 1193–1201.

Moritz C (1994) Defining evolutionary significant units for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.

Parsons TJ, Muniec DS, Sullivan K *et al*. (1997) A high observed substitution rate in the human mitochondrial DNA control region. *Nature Genetics*, **15**, 363–368.

Pesole G, Gissi C, De Chirico A, Saccone C (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. *Journal of Molecular Evolution*, **48**, 427–434.

Piertney SB, MacColl ADC, Bacon PJ, Racey PA, Lambin X, Dallas JF (2001) Matrilineal genetic structure and female-mediated gene flow in red grouse (*Lagopus lagopus scoticus*): an analysis using mitochondrial DNA. *Evolution*, **54**, 279–289.

Raymond M, Rousset F (1995) GENEPOP, version 3.1: population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Robinson TJ, Matthee CA (1999) Molecular genetic relationships of the extinct ostrich, *Struthio camelus syriacus*: consequences for ostrich introduction into Saudi Arabia. *Animal Conservation*, **2**, 165–171.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 52–569.

Rohling EJ, Fenton M, Jorissen FJ *et al.* (1998) Magnitudes of sea-level lowstands of the past 500 000 years. *Nature*, **394**, 162–164.

Saccone C, Pesole G, Sbisa E (1991) The main regulatory region of mammalian mitochondrial DNA: structure–function model and evolutionary pattern. *Journal of Molecular Evolution*, **33**, 83–91.

Siddall M, Rohling EJ, Almogi-Labin A *et al.* (2003) Sea-level fluctuations during the last galcial cycle. *Nature*, **423**, 853–858.

Slatkin M, Hudson RE (1991) Pairwise comparisons of mitochondrial DNA in stable and exponentially growing populations. *Genetics*, **129**, 555–562.

Stammbach E (1987) Desert, forest and montane baboons: Multilevel-societies. In: *Primate Societies* (eds Smuts BB, Cheney DL, Seyfarth RM, Wrangham RW, Struhsaker TT). University of Chicago Press, Chicago.

Stanley-Price MR (1989) *Animal Re-Introductions: the Arabian Oryx in Oman*. Cambridge University Press, Cambridge.

Swofford DL (2001) PAUP *4.0: Phylogenetic Analysis Using Parsimony and Other Methods*. Sinauer Associates, Sunderland, MA.

Thalman 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 great apes. *Molecular Ecology*, **13**, 321–335.

Thouless CR, Grainger JG, Shobrak M, Habibi K (1991) Conservation status of gazelles in Saudi-Arabia. *Biological Conservation*, **58**, 85–98.

Trauth MH, Deino AL, Bergner AGN, Strecker MR (2003) East African climate change and orbital forcing during the last 175kyr BP. *Earth and Planetary Science Letters*, **206**, 297–313.

Ward RH, Frazier BL, Dewjager K, Pääbo S (1991) Extensive mitochondrial diversity within a single Amerindian tribe. *Proceedings of the National Academy of Sciences USA*, **88**, 8720–8724.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Yan Z, Petit-Maire N (1994) The last 140ka in the Afro-Asian arid/semi-arid transition zone. *PaleoGeographyraphy, Paleoclimatology and Paleoecology*, **110**, 217–233.

Zurovsky Y, Shkolnik A (1982) Physiological adaptation of hamadryas baboon to the desert. *Israel Journal of Zoology*, **31**, 61–62.

Zurovsky Y, Shkolnik A (1993) Water economy and body-fluid distribution in the hamadryas baboon (*Papio hamadryas*). *Journal of Thermal Biology*, **18**, 153–157.

Molecular work for this study was carried out at the King Khalid Wildlife Research Centre (KKWRC), Saudi Arabia, in a laboratory dedicated to the conservation genetics of Arabian wildlife. KKWRC is managed by the Zoological Society of London on behalf of the Saudi Arabian National Commission for Wildlife Conservation and Development. BJW, RLH and MWB use molecular genetics to investigate evolutionary questions in a wide range of taxa. W. M. and B. F. are technicians at KKWRC, and work on projects involving gazelles, baboons and leopards. S. B., V. B. and A. B. study the ecology and behaviour of Arabian Hamadryas baboons.
