

# Phylogeography of Barbary macaques (*Macaca sylvanus*) and the origin of the Gibraltar colony

Lara Modolo<sup>†‡</sup>, Walter Salzburger<sup>§</sup>, and Robert D. Martin<sup>¶</sup>

<sup>†</sup>Anthropologisches Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland; <sup>§</sup>Lehrstuhl für Zoologie und Evolutionsbiologie, Fakultät für Biologie, Universität Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany; and <sup>¶</sup>The Field Museum, 1400 South Lake Shore Drive, Chicago, IL 60605

Communicated by F. Clark Howell, University of California, Berkeley, CA, March 22, 2005 (received for review March 11, 2004)

The Barbary macaque (*Macaca sylvanus*) is the earliest offshoot of the genus *Macaca* and the only extant African representative, all other species being Asiatic. Once distributed throughout North Africa, *M. sylvanus* is now restricted to isolated forest fragments in Algeria and Morocco. The species is threatened; the maximum total wild population size is estimated at 10,000 individuals. Relationships among surviving wild subpopulations in Algeria (96 samples) and Morocco (116 samples) were examined by using 468-bp sequences from hypervariable region I of the mitochondrial DNA control region. Twenty-four different haplotypes were identified, differing by 1–26 mutational steps (0.2–5.6%) and 1 insertion. With one exception (attributable to secondary introduction in coastal Morocco), Algerian and Moroccan haplotypes are clearly distinct. However, whereas Moroccan subpopulations show little divergence in hypervariable region I sequences and little correspondence with geographical distribution, there is a deep division between two main subpopulations in Algeria and one marked secondary division, with haplotypes generally matching geographical distribution. Accepting an origin of the genus *Macaca* of 5.5 million years ago, the Moroccan population and the two main Algerian subpopulations diverged  $\approx$ 1.6 million years ago. Distinction between Moroccan and Algerian haplotypes permitted analysis of the origin of the Gibraltar colony of Barbary macaques (68 samples; 30% of the population). It is generally held that the present Gibraltar population descended from a dozen individuals imported during World War II. However, the Gibraltar sample was found to include Algerian and Moroccan haplotypes separated by at least 16 mutational steps, revealing a dual origin of the founding females.

conservation | fragmentation | haplotype network | molecular clock | mtDNA

Population genetic studies describe genetic diversity within and among populations and contribute to knowledge about evolutionary mechanisms that change natural and artificial populations (1). Previous research has established crucial features of mtDNA, such as its relatively rapid sequence evolution, lack of recombination, and maternal inheritance (2). These properties make mtDNA particularly valuable for phylogenetic and phylogeographic reconstructions of relatively recent cladogenic events in species with female philopatry, such as macaques. Female philopatry involves a social system in which males disperse while females remain in their natal group. It is characterized by local homogeneity and large interpopulational differences in mtDNA (3).

Macaques are widely distributed in southern and eastern Asia as well as in limited areas of western North Africa. They are classified into at least 19 extant species, with the Barbary macaque (*Macaca sylvanus*) as the only extant African representative (4, 5). Barbary macaques live in multimale, multifemale groups with an approximately balanced adult sex ratio. Natural groups contain an average of  $\approx$ 30 animals living in a highly promiscuous mating system (6). Because of its small numbers and major habitat reduction, *M. sylvanus* has been

formally recognized as a vulnerable taxon by the World Conservation Union (2002 International Union for Conservation of Nature and Natural Resources *Red List of Threatened Species*).

After an inferred origin in northern Africa at least 5.5 million years ago (mya), early macaques presumably spread into Eurasia, and one lineage entered southern Europe by the beginning of the Pliocene (7–9). Fossils of Pliocene and Pleistocene macaques have been discovered throughout southerly regions of Europe and reliably identified as representatives of the *M. sylvanus* lineage (10–13). *M. sylvanus* apparently flourished in Europe only during the warm phases of the Pliocene and Pleistocene, persisting at least until the last interglacial and then eventually retreating from the continent entirely (7). The species was formerly widely distributed in the once entirely forested Maghreb, inhabiting diverse environments ranging from lowland shrub to high-altitude coniferous forests (14). The Maghreb has been continuously subjected to human influence for at least 4,000 years, and the most significant factor in extensive forest destruction has probably been overgrazing by goats. Two decades ago, the maximum population size of *M. sylvanus* was estimated at 20,000 individuals in a few refuges in Morocco and Algeria (14, 15) (map in Fig. 2). Since then, the wild population has apparently decreased to  $\approx$ 10,000 (16). Remaining wild subpopulations are completely isolated in forest fragments separated by extensive intervening built-up areas, such that natural genetic exchange between them is ruled out.

Although there are now three major captive colonies of Barbary macaques in France (Kintzheim and Rocamadour) and Germany (Salem), the only long-established European population is the provisioned free-ranging colony on Gibraltar. The origin of the Gibraltar colony is unclear; some authors have postulated a remnant European population, whereas other anecdotes suggest that animals were imported from North Africa between 711 and 1492 A.D. However, the first written record of Barbary macaques on Gibraltar dates from 1704 (17). The population was subsequently subjected to artificial intervention to regulate group size over a long period. Numbers were at times kept low by removing individual animals and at others restocked with importations from North Africa. The last introduction of new stock occurred during the second World War, reportedly from Morocco (18).

The pronounced philopatric nature of female macaques has been shown in other macaque species to preserve multiple mitochondrial types and can lead to local fixation of deeply divergent lineages (5, 19). This pattern is reinforced by the ever increasing anthropogenic impacts on *M. sylvanus* in North Africa. Habitat destruction has not only reduced overall population size to a very low level but has also promoted marked isolation

Abbreviations: AMOVA, analysis of molecular variance; kya, thousand years ago; mya, million years ago.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. DQ001976–DQ002005).

<sup>†</sup>To whom correspondence should be addressed. E-mail: modolo@aim.unizh.ch.

© 2005 by The National Academy of Sciences of the USA

**Table 1. Names of sample locations (see Fig. 2), sample sizes (N), and haplotype frequencies**

Population	N	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24
Gibraltar	68		24							8							36								
Rif	28		4							20							4								
Middle Atlas	69	4	16	3				1	12		10	1	16	2	3	1									
High Atlas	19		6		6	3			4																
Akfadou	25																25								
Djurdjura	54																						45	9	
Pic d. Singes	12																	7	5						
Kherrata	5																			2	1	1	1		

of the remaining subpopulations. Study of the genetic effects of population fragmentation is therefore of central importance for formulation of long-term conservation guidelines.

In this study, we present a phylogenetic and phylogeographic analysis of the Barbary macaques in North Africa, with special emphasis on the origin of the Gibraltar colony. We analyzed the most variable segment of the control region of mtDNA from seven wild-living populations in Algeria and Morocco and the semicaptive colony on Gibraltar. Our primary aim was to elucidate phylogeographic relationships among the individual Barbary macaque populations and to examine the extent to which habitat fragmentation has left its mark on mtDNA variation. Previous studies indicated reduced genetic variability relative to other macaque species (20). A linearized tree approach was applied, based on molecular clock calibrations, to estimate dates for divergence between lineages. We then reconstructed the geographical origin of the Gibraltar population, while recognizing that some lineages may have already become extinct.

## Materials and Methods

**Sample Collection.** This study includes 90 and 96 samples from wild populations of Barbary macaques in Morocco and Algeria, respectively (map in Fig. 2 and Table 1). Samples from 26 additional early-generation individuals from the Moroccan Middle Atlas originated from outdoor enclosures in Germany and France. Sixty-eight samples were collected in Gibraltar from individually known animals living in six different social groups ( $\approx 30\%$  of the total Gibraltar population).

Collection of blood and hair samples was carried out as far as possible by trapping animals in baited cages and sedating them (Algeria and Gibraltar). When trapping was not possible, fecal samples were collected (Morocco). For the Algerian population, samples were collected from four of the seven remaining wild subpopulations (Akfadou, Djurdjura, Kherrata, and Pic des Singes) and from individual groups at eight different localities (five social groups at Djurdjura and single groups at the other three localities). Samples from Morocco originate from three different areas of distribution, namely from the Rif Mountains, from the Azrou Forest in the Middle Atlas, and from Cascade d'Ouzoud and Ourika Valley in the High Atlas.

**Laboratory Methods.** DNA was extracted from hair and blood samples by using PCI (25:24:1 mix of phenol, chloroform, and isoamyl-alcohol) and chloroform (21). From fecal samples, DNA was extracted with the QIAamp DNA stool kit (Qiagen) according to the manufacturer's protocol. Initially, a fragment of  $\approx 1,200$  bp of the mtDNA was amplified from blood samples and a shorter fragment of 500 bp was amplified from fecal and hair samples. This was then used to sequence 468 nt of hypervariable region I of the control region. PCR amplifications were performed in a 100- $\mu$ l final volume by using 0.06 M Tris, 0.015 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.0015 M  $\text{MgCl}_2$ , 0.78 M dimethyl sulfoxide (DMSO), 0.025 mM each dNTP, 1 mM each primer, and 2.5

units of Promega *Taq* polymerase in an Applied Biosystems 9700 thermal cycler. The primers used were M2Pro (5'-TCC ACC ACC AAC ACC CAA AGC-3'), M282 (5'-AAG GCC AGG ACC AAG CCT-3'), and DLR3 (5'-GAA CCA GAT GTC GGA TAC AG-3') as an internal primer for the sequencing reaction. The reaction cycles consisted of denaturation for 30 s at 95°C, primer annealing for 60 s at 50°C, and extension for 60 s at 72°C. Cycles were repeated 35 times, with the final cycle including an extension reaction of 5 min. The PCR products were purified with the QIAquick PCR purification kit (Qiagen) according to the supplier's instructions. Cycle sequencing reactions were performed by using BigDye Terminator cycle sequencing chemistry (Applied Biosystems). Sequencing products were purified by Sephadex G-50 in CentriSep columns (Princeton Separations), and sequences were analyzed by using an Applied Biosystems 377 sequencing system. All templates were sequenced for both strands.

**Phylogenetic and Population Genetic Analyses.** Sequences were aligned by using the program SEQUENCHER (Gene Codes, Ann Arbor, MI). For phylogenetic analyses, we performed maximum-likelihood, maximum-parsimony, and neighbor-joining analyses in parallel by using the program PAUP\* 4.0b10 (22). MODELTEST 3.06 (23) was run to determine the appropriate model of molecular evolution in a likelihood ratio test framework. Six representatives of three of the four major *Macaca* lineages (*Macaca tonkeana*, *M. nigra*, *M. nemestrina*, *M. mulatta*, *M. fascicularis*, and *M. arctoides*) were taken collectively as the outgroup (4, 24). Gaps were treated as a fifth state in parsimony analyses. Bootstrap analyses were performed with 5,000 replicates for maximum parsimony and neighbor-joining; for maximum likelihood, we performed 100 full heuristic bootstrap replicates. For Bayesian inference, we used the program MRBAYES v.3.0b4 (25). Four Monte Carlo Markov chains were run for 100,000 generations, sampling every 10 generations; the initial 5% of trees were discarded as burn-in.

To infer timing of major branching events for the Barbary macaques, we performed molecular clock analysis, based on a maximum-likelihood tree in which the molecular clock was enforced. Average uncorrected pairwise distances between and among the four major lineages were calculated to infer absolute time scales, based on an inferred minimum date for the initial divergence of *M. sylvanus* from other macaque species at  $\approx 5.5$  mya (7, 8, 19). To examine relationships between the haplotypes observed in the 286 specimens sequenced, we applied a haplotype network approach. To this end, we translated the maximum-likelihood topology into an unrooted network, describing the optimal maximum-likelihood topology under the maximum parsimony criterion to map the number of mutations assigned to each branch of the network.

Hierarchical analysis of molecular variance (AMOVA) (26) was performed to compare levels of genetic diversity within and among four principal Barbary macaque lineages by using ARLEQUIN 2.0 (27) and 1,000 permutations. Both the Gibraltar pop-

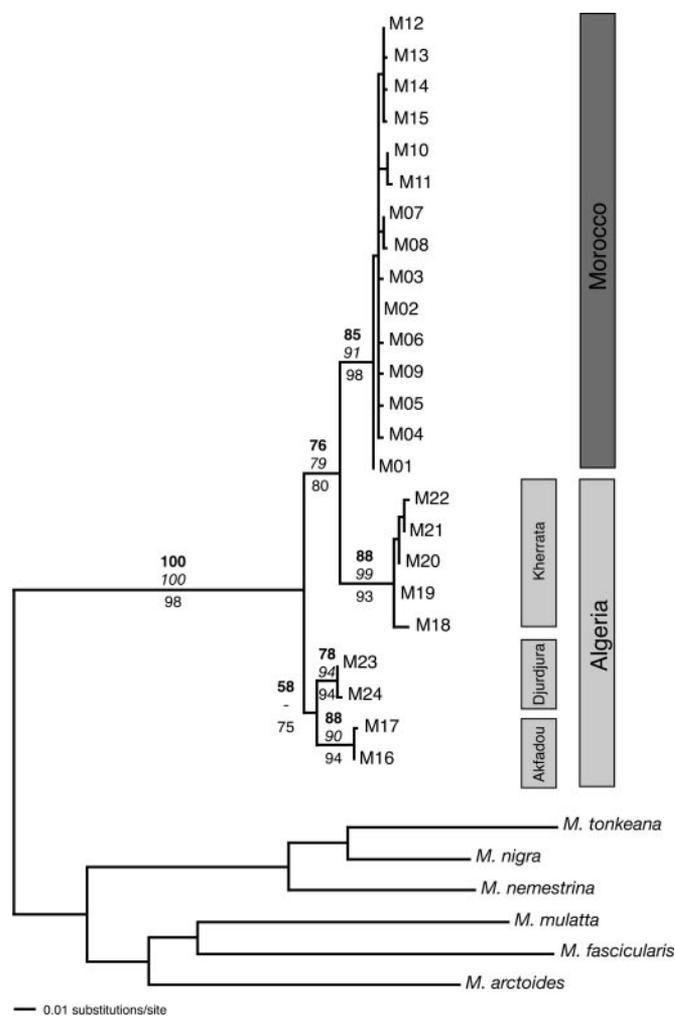
ulation and the Rif group at locality 2 (M16) were excluded from this analysis because they involve artificially introduced animals. For AMOVA, the remaining wild populations were arranged according to geographical distribution into seven groups, three from Morocco and four from Algeria (Table 1). A mismatch analysis (28) was then performed with ARLEQUIN 2.0 to compare the demographic histories of all lineages according to the phylogenetic analysis. Finally, the moment estimator of time to expansion ( $\tau$ ) was used in the equation  $\tau = 2\mu t$  to infer a time scale for the demographic expansion (where  $t$  is the coalescent time in generations, which corresponds to the time of the last common ancestor, and  $\mu$  is the mutation rate for the whole sequence), assuming an average generation time of 5 years (29).

## Results

**Phylogenetic Analyses.** Using COLLAPSE 1.1 (30), we identified 24 different haplotypes within the entire sample of Barbary macaques, differing by 1–26 mutational steps (0.21–5.6%). Haplotype M24, shared by nine individuals of a single social group in Djurdjura, showed an insertion at position 155 in our alignment. Of the total 468 nucleotide positions, 116 were parsimony-informative when including the six outgroup taxa and 35 without the outgroup taxa. The HKY+G model (31) was selected by MODELTEST as the best-fitting model, with a gamma shape correction of 0.21, a transition/transversion ratio of 7.94, and base frequencies of  $A = 0.308$ ,  $C = 0.298$ , and  $G = 0.129$  (the parameters being estimated from the data). All phylogenetic analyses performed resulted in almost identical tree topologies (Fig. 1), identifying four main clusters among the Barbary macaques. The sole Akfadou haplotype (M16) and one aberrant haplotype from Pic des Singes (M17) clustered with the two haplotypes found in Djurdjura (M23/M24). This clade was resolved as sister group to a clade containing haplotypes from Kherrata/Pic des Singes (M18–M22) and Morocco (M1–M15). Generally, there were high bootstrap supports in all analyses, as well as high Bayesian posterior probabilities with respect to the branching order of the lineages.

We calculated average uncorrected pairwise distances between principal lineages in the genus *Macaca* for a molecular clock calibration, using the minimal initial divergence time of 5.5 mya based on the fossil record (7, 19). The average pairwise distance resulted in a mutation rate  $\mu$  of 2.70–3.04% (see Table 4, which is published as supporting information on the PNAS web site). According to the average pairwise distances obtained, the initial split between the Morocco/Kherrata clade and the Algerian Akfadou/Djurdjura subclade was dated at 1.24–1.85 mya, and the separation between Morocco and Kherrata at 1.22–1.62 mya. The basal split among individuals assigned to the Akfadou/Djurdjura subclade was estimated at 0.90 mya. Distances within groups were lower, with estimated maximal divergence times of 64–359 thousand years ago (kya) for the Kherrata subpopulation, 110–282 kya for the Moroccan clade, and about 70–78 kya for the Akfadou and Djurdjura subpopulations. Although the absolute timing of divergences may be questionable [notably because direct calibration from the fossil record may lead to serious underestimation (8)], at least the sequence and relative timing depicted here convincingly reflect the evolutionary history of Barbary macaques.

**Population Genetic Analyses.** In the haplotype network (Fig. 2), three main haplotype clusters emerged. The most abundant haplotype M16, shared by 65 individuals, was found in three different sampling areas, and constitutes [together with one Pic des Singes haplotype found in seven individuals (M17)] a cluster with the Algerian subclade Djurdjura. In the Djurdjura subclade, haplotype M23 was found in all except one social group, which shared haplotype M24. Another Algerian group was resolved including the Kherrata representatives (M19–M22) and a hap-



**Fig. 1.** Maximum-likelihood tree for all 24 haplotypes of Barbary macaques and for 6 outgroup taxa. Bootstrap percentages recorded for maximum-likelihood full heuristic search (100 replicates) and maximum parsimony trees (5,000 replicates with fast stepwise addition option) are shown above major branches in bold and italics, respectively. Bayesian posterior probability values are presented below internodes. The haplotype numbers refer to those in Table 1. The population of Pic des Singes is divided between the Akfadou and Kherrata clades.

lotype found in five individuals of Pic des Singes (M18). Within the Moroccan cluster, diversity was largest in the Middle Atlas population, reflected both by the number of haplotypes and the distances between them. The most abundant and widespread Moroccan haplotype M02 is connected by one step to the ancestral haplotype M01, found only in the Middle Atlas. Four haplotypes of the Moroccan clade were restricted to only two geographical populations (with Gibraltar excluded as a nonwild population), namely the Rif Mountains (M09) and the High Atlas (M04–M05/M08), respectively. The Gibraltar population was resolved into three different haplotypes, clearly suggesting a Moroccan (M02/M09) and Algerian (M16) origin. The Algerian haplotype M16 is separated from the Moroccan haplotype M02 by 22 mutational steps and from the other Moroccan haplotype M09 by 23 steps.

**AMOVA and Mismatch Analyses.** To quantify population genetic structure within and among populations, a hierarchical AMOVA was performed, grouping all but the Gibraltar population and the presumably introduced social group in the Rif (with Algerian



**Table 3. Results of the mismatch analyses**

Clade	$\tau$	Obs.	$\theta_0$	$\theta_1$	Ragged.	$t_{\text{expansion}}$ , kyr
Kherrata	4.79 (1.34–12.29)	2.69	0.0 (0.0–6.3)	5.1 (1.3–3890)	0.101	194 (54–498)
Morocco	1.56 (0.76–1.89)	1.43	0.0 (0.0–0.9)	3338 (18.62–6699)	0.821	63 (31–76)
Djurdjura	3.20 (0.45–4.70)	0.57	0.0 (0.0–0.8)	0.4 (0.0–1749)	0.674	130 (18–191)
Akfadou	0.28 (0.001–0.91)	0.18	0.0 (0.0–0.3)	0.5 (0.0–2399)	0.446	11 (0.040–37)

The expansion parameter ( $\tau$ ), the mismatch observed mean (Obs.), the mutation parameter before ( $\theta_0$ ) and after ( $\theta_1$ ) expansion, and the raggedness index (Ragged.) (48) were calculated for all four major groups of haplotypes according to obtained phylogenies. Values in brackets refer to 95% confidence intervals. The timing of the most important demographic expansion ( $t_{\text{expansion}}$ ) was determined by assuming a mutation rate of 2.90% per million years, based on the results of the average uncorrected pairwise distances. See Fig. 3 for graphs.

supporting information on the PNAS web site), with one peak corresponding to differences between the four lineages, one peak corresponding to differences between Akfadou and Djurdjura, and the third peak corresponding to differences within populations.

## Discussion

**Phylogenetic Reconstruction.** mtDNA data revealed significant genetic structuring among Barbary macaques, showing deep genetic divergence between some lineages (Fig. 1). For the Algerian lineages, this finding concurs with a previous study based on microsatellites (32). Of particular interest is the initial split occurring at  $\approx 1.6$  mya, dividing the Moroccan clade and one major Algerian clade from a second major Algerian clade (Fig. 1). This result is surprising in that Morocco is sister to the Algerian clade (Kherrata/Pic des Singes) that currently shows the greatest geographical separation. This finding might be explained by climatic oscillations influencing patterns of geographic distribution in the past. Otherwise, the tree reconstruction corresponds to geographic relationships, with the single exception of the locality of Pic des Singes. The large mtDNA difference shown by M17/M18 and the haplotypic affiliation to neighboring populations could have resulted from recent isolation of Pic des Singes from previously connected lineages. This is an interesting case, indicating that the single social group in Pic des Singes contains two completely distinct lineages. The results are consistent with the findings that today the small subpopulation of 50 animals inhabits only a relict fragment of a once widespread oak forest (14). Microsatellite data (32) also support a relatively recent separation of the Pic des Singes population from that at Kherrata. Furthermore, the Kherrata population displays the greatest variability despite having the smallest sample size. Similar findings have been reported in a previous study of microsatellites, in which Kherrata was found to have the highest number of private alleles in one social group (32). However, the presence of three markedly divergent mtDNA lineages in Algeria suggests that the observed geographical patterns probably arose through extinction of intermediate populations and physical barriers to dispersal between regions.

Surprisingly, all Moroccan haplotypes are grouped into a fully supported monophyletic clade, in contrast to the paraphyletic Algerian populations. Phylogenetic analyses failed to resolve the Moroccan clade into its three geographical populations, revealing very close relationships among them. The apparently oldest haplotype M01 is exclusively located in the Middle Atlas, from which the central and most abundant haplotype M02 derives. In addition to the Middle Atlas, M02 also occurs in the Rif Mountains and in the High Atlas. These two localities each have one (M09) and three (M04–M05/M08) private haplotypes, respectively, in addition to the central haplotype M02 shared with the Middle Atlas. This presumably indicates that M02 expanded from this region northward into the Rif and southward into the High Atlas.

**Phylogeography from the Perspective of Glacial Events.** Pleistocene climatic changes have shaped patterns of geographical distribution as well as demographic history and genetic diversification (33). For most species, repeated climatic oscillations resulted in range contractions into separate refugia during climatic cooling and drying (34). The coalescent theory (35) predicts that the most widespread and frequent haplotypes indicate the locations of refugial populations. According to the inferred coalescent time for Kherrata and Morocco, haplotypes diverged in isolation  $\approx 213$  kya and 196 kya, before and during the onset of the Riss glaciation. Taking into account the fact that the most frequent and widespread haplotype M02 is present in Morocco, one may speculate that both Morocco and Kherrata (Algeria) were refuges for Barbary macaques during the last glaciation. Barbary macaques from Akfadou and Djurdjura might have diverged from a different lineage  $\approx 74$  kya during the interstadials when climatic conditions were still moderate (36). By the mid-Holocene, habitats of *M. sylvanus* in Morocco and Algeria might have been interconnected through extended forests until the beginning of forest exploitation by human populations, dated  $\approx 2$  kya (37). There is reason to believe that the North African forest was heavily depleted in Roman times as a result of exportation of cedars and oaks. The Arab conquest and immigration of nomads intensified the destruction, with the result that by the end of the 17th century most of the lowlands had been cleared (38). The lack of forests probably presented an obstacle to migration, leading to the geographical separation of the mtDNA haplotypes that we see today. These patterns are reinforced by the marked philopatric nature of female macaques, causing sharp geographical clustering of mtDNA haplotypes (3). Low levels of control region diversity have been seen not only in species with pronounced female philopatry, but also in species known to have undergone population bottlenecks in the recent past (39–41). It has also been argued that reduced genetic variability in *M. sylvanus* compared with other macaques might be the result of successive bottlenecks during withdrawal from a wider Pleistocene expansion to contemporary areas of dispersal (42). The model of sudden expansion for the Moroccan population is supported by the bell-shaped mismatch distribution (see Fig. 3), as well as by the star-like structure of the haplotype network and its short branches (Fig. 2) (43). Moroccan populations perhaps underwent sudden expansion  $\approx 63$  kya, starting from one matriline in the Middle Atlas and extending northward and southward (Table 3). This inference fits with the more favorable climate during this interstadial that permitted some expansion of Mediterranean oak forests (36). Akfadou and Djurdjura showed a skewed distribution due to excess of similar or identical haplotypes, which can arise in two ways. First, it is possible that recent expansion of a young population occurred. If so, then the two subpopulations might have suffered a bottleneck in the past or were recolonized only recently. Alternatively, it is possible that sampling was conducted within the geographical boundaries of a single genetic stock. This can be excluded at least for Djurdjura, where sampling took place in

