Molecular Ecology (2004) **13**, 3425–3435 doi: 10.1111/j.1365-294X.2004.02332.x

Rivers influence the population genetic structure of bonobos (*Pan paniscus*)

J. ERIKSSON,*† G. HOHMANN,* C. BOESCH* and L. VIGILANT*
**Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany, †Department of Animal Ecology, Evolutionary Biology Centre (EBC), Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden*

Abstract

Bonobos are large, highly mobile primates living in the relatively undisturbed, contiguous forest south of the Congo River. Accordingly, gene flow among populations is assumed to be extensive, but may be impeded by large, impassable rivers. We examined mitochondrial DNA control region sequence variation in individuals from five distinct localities separated by rivers in order to estimate relative levels of genetic diversity and assess the extent and pattern of population genetic structure in the bonobo. Diversity estimates for the bonobo exceed those for humans, but are less than those found for the chimpanzee. All regions sampled are significantly differentiated from one another, according to genetic distances estimated as pairwise F_{ST}s, with the greatest differentiation existing between region East and each of the two Northern populations (N and NE) and the least differentiation between regions Central and South. The distribution of nucleotide diversity shows a clear signal of population structure, with some 30% of the variance occurring among geographical regions. However, a geographical patterning of the population structure is not obvious. Namely, mitochondrial haplotypes were shared among all regions excepting the most eastern locality and the phylogenetic analysis revealed a tree in which haplotypes were intermixed with little regard to geographical origin, with the notable exception of the close relationships among the haplotypes found in the east. Nonetheless, genetic distances correlated with geographical distances when the intervening distances were measured around rivers presenting effective current–day barriers, but not when straight-line distances were used, suggesting that rivers are indeed a hindrance to gene flow in this species.

Keywords: bonobo, chimpanzee, Democratic Republic of Congo, HV1, mitochondrial DNA, phylogeography, population structure

Received 11 June 2004; revision received 23 July 2004; accepted 23 July 2004

Introduction

One long-standing view of the formation of species requires the disruption of gene flow by geographical isolation of a previously panmictic population into two or more disparate populations, thus allowing for the accumulation of mutational change and vicariate divergence into separate species (Mayr 1963; Wu & Ting 2004). A similar process of limited gene flow can create population structure within a species. While many large mammals are capable of long-range dispersal across various geographical barriers provided that suitable continuous habitat exists (Lehman & Wayne 1991; Roy *et al*. 1994; but see Forbes & Boyd 1997), bodies of water can present an impenetrable barrier to routine dispersal by several primate species (Colyn *et al*. 1991; Ayers & Clutton-Brock 1992; Telfer *et al*. 2003). Large rivers have therefore been suggested to influence inter- and intraspecific patterns of differentiation for many primates, including the large-bodied great apes (Schwarz 1934; Oates 1988; Grubb 1990).

One of the most striking examples is provided by the distribution of chimpanzees and bonobos in Africa. The bonobo, *Pan paniscus*, occurs only south of the Congo River, while the more widespread chimpanzee, *P. troglodytes*, can be found north of this river in a discontinuous distribution from West to East Africa (Fig. 1). The classic definition of chimpanzee subspecies has riverine barriers separating the

Correspondence: L. Vigilant. Fax: + 49 341 3550299; E-mail: vigilant@eva.mpg.de

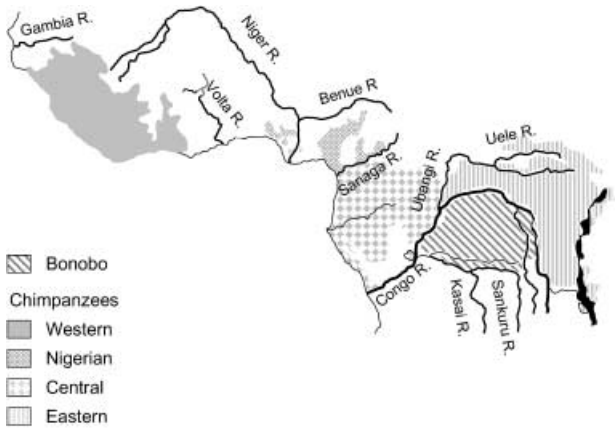

Fig. 1 The approximate current geographical distribution of chimpanzees and bonobos is indicated by shading. Major rivers suggested to have influenced *Pan* distribution are also shown.

geographically defined but morphologically similar subspecies, with the eastern chimpanzee (*P. t. schweinfurthii*) occurring north of the Congo River and east of the Ubangi River, the central chimpanzee (*P. t. troglodytes*) situated west of the Ubangi River and east of the Niger River, and the western chimpanzee (*P. t. verus*) found west of the Niger River to as far west as the Gambia River (Fig. 1) (Schwarz 1934; Osman-Hill 1969).

A connection between the population history and levels of present-day neutral genetic diversity in populations is well established, and patterns of genetic variation observed within species are the result of migration and random genetic drift operating as a function of the species' effective population size over the last millions of years (Slatkin 1995; Rousset 1996; Rousset 1997). Furthermore, when samples of known origin are available, genetic analysis provides an opportunity to compare the distribution of genealogical lineages with geography (Avise 2000) and to investigate a population's past demographic history.

Such studies often rely upon analysis of variation of the mitochondrial DNA (mtDNA) molecule, as the rapid rate of mtDNA sequence evolution means that it is particularly useful for consideration of relatively recent separation events. With the highest rate of evolution of any mitochondrial segment (Brown 1983; Saccone *et al*. 1991), the control region is the segment most often used in analyses of variation within species (Avise 2000). An initial study of mtDNA control region sequence variation using a sampling of eastern, central and western chimpanzees produced a phylogenetic tree with a clear divergence between a clade containing sequences from western chimpanzees and one containing sequences from both eastern and central chimpanzees (Morin *et al*. 1994). However, the limited sampling in regions near rivers considered important for defining subspecies, in combination with apparently high levels of gene flow inferred from the demonstrated similarity of sequences collected from individuals hundreds of kilometres apart, mean that the precise role of rivers as boundaries to gene flow among chimpanzee subspecies is uncertain (Morin *et al*. 1994; Goldberg & Ruvolo 1997; Gagneux *et al*. 2001). The importance of appropriate sampling was underlined by work by Gonder and colleagues suggesting that the smaller Sanaga River, rather than the large Niger River, represented the dividing line between central and western chimpanzees (Gonder *et al*. 1997) and further proposing that chimpanzees from Nigeria are sufficiently phylogenetically distinct to warrant classification as a new, fourth subspecies, *P. t. vellerosus*.

In contrast to the situation with chimpanzees, the demographic, behavioural and ecological components of variation in bonobos have have been little studied (Stanford 1998; Boesch 2002) and little is known about genetic variation of bonobos in the wild (reviewed in: Bradley & Vigilant 2002). While chimpanzees are distributed widely across much of equatorial Africa and occupy a wide range of habitats including tropical lowland forest, savanna/gallery forest and montane forest, bonobos inhabit a more limited habitat range. The majority of bonobos live in areas consisting of dense tropical lowland forest, with the exception of the most southerly distribution which occurs in a forest/savanna mosaic. Studies of genetic diversity in bonobos have either considered captive individuals of unknown geographical origin (Gerloff *et al*. 1999; Reinartz *et al*. 2000) or examined single wild groups habituated to human observation (Hashimoto *et al*. 1996; Gerloff *et al*. 1999) and therefore no comprehensive conclusions concerning genetic variation or population structure in the wild could be drawn.

The generally contiguous forest of the Congo basin in which the bonobos range is crossed by numerous substantial rivers (Fig. 2). In this study we examine mtDNA control region sequence variation in individuals from five distinct localities separated by rivers in order to estimate relative levels of genetic diversity, assess population subdivision and investigate historical demographic changes in the bonobo. The maternally inherited mtDNA molecule is particularly suited for this investigation because the bonobo, in common with chimpanzees and humans, is a species characterized by primarily female dispersal (Kano 1992). In contrast to chimpanzees, during glaciation events in the Pleistocene, bonobos would have been restricted to a single fluvial refuge area south and along the Congo River (Maley 1996; but see Colyn *et al*. 1991). This means that the geographical pattern of variation observed today should be influenced only by physical barriers to gene flow and not be confounded by dispersal from disparate refugia, as is the probable case in chimpanzees (Gonder 2000). Understanding the apportionment of genetic variation in an endangered and relatively unprotected species such as the bonobo is also of practical importance as it may affect

Fig. 2 The shaded area indicates the current estimated geographical distribution of bonobos within the Demographic Republic of Congo. Geographical location of samples used in this study are indicated (N, North; NE, Northeast; C, Central; S, South; E, East) as well as the location and names of the major rivers (1, Congo-Lualaba.; 2, Kasai-Sankuru; 3, Lomami; 4, Salonga-Tshuapa; 5, Lukenie) likely to influence movement of bonobos.

the size and shape of conservation areas and thereby possibly influence management decisions regarding the allocation of resources for conservation.

Materials and methods

Sampling and DNA extraction

Samples were collected from bonobo populations across the geographical range of the species distribution with a focus on regions separated by large rivers (Fig. 2). Bonobo groups were localized by detection of vocalizations during the course of the day and were followed until they constructed night nests. Fresh faeces were collected from underneath the vacated nests early the following morning. In order to minimize cross-contamination and double sampling of individuals, only faeces directly underneath nests separated by a minimum distance of 2 m were collected. For each sample, approximately 5–10 g of faeces were placed a tube containing 40–50 mL of RNAlater (Ambion) and kept at ambient temperatures for up to 3 months in the field, with subsequent storage in the laboratory at –20 °C. The diameter of the home range of a habituated bonobo community at Lomako was estimated at 5 km, and so bonobo nest groups encountered a minimum distance of 10 km apart were assumed to represent distinct social units or 'communities' and are referred to here as separate populations. Three populations were collected in the Central and East regions, respectively, and one in the South. The total number of samples collected from each region was 84 in Central, 43 in South and 23 in East (Fig. 2). Included in this study are an additional five haplotypes from 36 individuals from Lomako, which are here termed the North population (GenBank Accession nos AF137482–AF187486) (Gerloff *et al*. 1999) and seven haplotypes from 17 individuals from Wamba, termed here the Northeast population (Hashimoto *et al*. 1996) (Fig. 2). Sequences derived from captive bonobos were included in the phylogenetic analysis (Gerloff *et al*. 1999; Deinard & Kidd 2000).

A QIAampDNA Stool Kit (Qiagen) was used to extract genomic DNA following a slightly modified version of the protocol provided by the manufacturer. Specifically, 2 mL of faeces–RNALater mixture was first centrifuged for 15 min at 3000 ***g*** and the supernatant discarded, followed by a second centrifugation for 15 min at 500 ***g*** and discard of supernatant. The pellet was then resuspended in 1.6 mL ASL buffer from the kit, vortexed, and incubated for 5 min at room temperature. The subsequent steps followed the manufacturer's protocol, and a final volume of 200 μL was obtained for each extract, aliquoted and stored at –20 °C. Two negative extraction controls were processed along with each set of 10–15 faecal extracts. A polymerase chain reaction (PCR) amplification of a segment of the X–Y homologous amelogenin locus was used in order to check for contamination, assess the success of each extraction and sex each sample (Sullivan *et al*. 1993; Bradley *et al*. 2001). Multiple amelogenin PCRs were conducted for each extract and sex was assigned either as male following a minimum of two observations of the allele derived from the Y chromosome or as female upon four observations of only the allele from the X-chromosome. Extrapolation from the observed error rate of a previous study in which DNAs from faecal samples from known chimpanzees were sexed (Bradley *et al*. 2001) suggests that the accuracy of sex determination using this multiple PCR procedure was > 99%.

Control region sequencing

A 470-base pairs (bp) segment of the first hypervariable control region (HV1) of the mitochondrial genome was amplified using the primers L15996 (Vigilant *et al*. 1989) and H16498 (Kocher & Wilson 1989). PCRs were set up in 50 μL volumes consisting of 1X *Taq* Gold® PCR buffer, 2 mM $MgCl_2$, 0.2 μM of each primer, 0.1 mM each dNTP, 1 U Ampli *Taq* Gold® (Perkin Elmer), 0.4 μg/μL bovine serum albumin (BSA) and 5 μL of DNA extract. The PCR conditions were: 3 min at 97 °C followed by 25 cycles of 30 s at 95 °C, 30 s at 60 °C and 30 s at 72 °C followed by 72 °C for 30 min and 4 °C storage in a Peltier thermal cycler, PTC 200 (MJ Research). The success of the PCR was assessed by gel electrophoresis of 5 μL of the product. After purification

using the QIAquick PCR purification kit (Qiagen) according to the protocol provided by the manufacturer, the products were cycle-sequenced (Kilger & Pääbo 1997) in both directions using the same set of primers as for the amplification. Sequences were aligned by eye using the sequence alignment editor BIOEDIT version 5.0.9 (Hall 1999). Sequences from both primers were consistent, and no additional sequences attributable to nuclear insertions of mtDNA were detected using this set of primers either in this study or in a previous study by Thalmann and coworkers (Thalmann *et al*. 2004) that specifically investigated nuclear inserts of HV1 in the great apes. All haplotype sequences have been deposited in GenBank (Accession nos AJ829446–AJ829473).

Determination of the number of individuals

Despite the precautions mentioned above, some individuals may have been sampled more than once and so samples derived from the same population that were found to have the same sex and mtDNA sequence were genotyped at three nuclear microsatellite loci (D11S2002, D5S1470 and D2S1326) described originally for humans and used in studies of chimpanzees (Vigilant *et al*. 2001). Details of PCR and analytical methods are as described elsewhere (Vigilant *et al*. 2001). When tested using a random subset of individuals ($N = 8$), each microsatellite marker was found to be highly polymorphic in bonobos (D11S2002, allele number (k) = 7, $H_0 = 0.875$; D5S1470, $k = 6$, $H_0 = 1.000$ and D2S1326, $k = 6$, $H_0 = 0.875$). Individuals of the same sex and mtDNA sequence, but with differing microsatellite genotypes, were considered to represent different individuals.

mtDNA diversity, phylogeny and population structure

Haplotype (gene) diversity (h) and its standard deviation (SD) (Nei 1987), mean pairwise sequence difference (MPD) and nucleotide diversity (π) and its standard deviation for each population (Nei & Li 1979) were estimated using the programs ARLEQUIN 2.0 (Schneider *et al*. 2000) and DNASP (Rozas & Rozas 1999).

Phylogenetic analyses were conducted using the quartet-puzzling algorithm implemented in TREE-PUZZLE version 5.0 (Schmidt *et al*. 2002) to generate maximum-likelihood (ML) trees and the program package MEGA2 (Kumar *et al*. 2001) for neighbour-joining (NJ) tree analyses. Trees were calculating using the Tamura–Nei model of sequence evolution (Tamura & Nei 1993) in MEGA2 (Kumar *et al*. 2001) and both the Tamura–Nei model and the the Hasegawa–Kishino–Yano (HKY) (Hasegawa *et al*. 1985) model in TREE-PUZZLE 5.0 (Schmidt *et al*. 2002). In TREE-PUZZLE, an alpha value of 0.32 for the gamma-distributed rate heterogeneity across sites gave the best fit to the data. We ran 1000 puzzling and 1000 bootstrap steps, respectively, to evaluate the support of each node in the ML and NJ trees.

In order to estimate the extent of genetic differentiation of the multiple populations sampled within Central and East regions, pairwise F_{ST} values (Weir & Cockerham 1984; Michalakis & Excoffier 1996) were calculated using TREE-PUZZLE 5.0 (Schmidt *et al*. 2002). Similarly, pairwise F_{ST} statistics were used to estimate the differentiation between each region. As an additional measure of genetic distance between regions, we also calculated Nei's net number of nucleotide differences (D_A) between regions (Nei & Li 1979). The statistical significance of the extent of differentiation was tested using a nonparametric permutation approach (Excoffier *et al*. 1992). In this approach, the null distribution under the hypothesis of no difference between the populations/regions is obtained by permuting haplotypes between populations and the P-value is the proportion of permuted F_{ST} values larger or equal to that observed.

The relationship between regional differentiation and distance between regions was examined by plotting the pairwise F_{ST} values against two different measurements of geographical distance: the shortest straight-line distance between regions and the distance measured around any separating rivers shown in Fig. 2. Geographic distances were estimated using the measure tool in the software Encarta World Atlas 1999 (Microsoft). The significance of associations was determined using a correlation test (Mantel 1967) and 5000 permutations as implemented in the computer program MATMAN version 1.0 (Noldus Information Technology).

In order to further deduce the significance of geographical divisions among local and regional groupings, an analysis of molecular variance (AMOVA (Excoffier *et al*. 1992)) was carried out using ARLEQUIN 2.0 (Schneider *et al*. 2000). AMOVA calculations using haplotype frequency and sequence divergence, respectively, were conducted. This approach is analogous to an ANOVA in which the correlation among haplotype distances, at various hierarchical levels, are used as F-statistic analogues. We used the Excoffier *et al*. (1992) method because, unlike other F_{ST} analogues, it is not sensitive to deviations from the normal distribution of genotypes (Takahata & Palumbi 1985; Hudson *et al*. 1992; Georgiadis *et al*. 1994).

Historical demographic events

In order to compare possible similarities and differences in past demographic events among bonobos, chimpanzees and humans we tested for past population reduction followed by an expansion by calculating Tajima's D as implemented in the software package DNASP (Rozas & Rozas 1999). Under neutrality, the number of nucleotide differences between sequences from a random sample should be equal to the number of differences between the segregating (polymorphic) sites only, but population expansions can cause

significant negative departures of Tajima's *D* from zero. Furthermore, because population size changes are expected to leave detectable patterns in the distribution of genetic differences (Slatkin & Hudson 1991; Rogers & Harpending 1992; Rogers *et al*. 1996), we compared the observed pairwise nucleotide site differences (mismatch distribution) for bonobos, humans and three chimpanzee subspecies. In addition, we calculated and compared some standard diversity values as mentioned above [nucleotide diversity (π) and mean pairwise difference (MPD)]. Calculations were based on the same segment of the HVI for all species (data taken from Vigilant *et al*. 1991; Goldberg & Ruvolo 1997; Gagneux 1998; Gagneux *et al*. 1999).

Results

mtDNA HV1 sequences

We successfully determined the sex and mtDNA HV1 haplotype for 131 of the 150 samples (~87%) collected for this study. Of these, 108 samples had the same sex and mtDNA haplotype as one or more additional samples from the same population and were therefore genotyped using three microsatellite loci. Only 27 of these showed identical three-locus genotypes to one or more samples and were deleted because they possibly represented repeated samples of the same individual, reducing the number of unique individuals analysed to 104 individuals. Use of just three microsatellite loci to distinguish individuals was supported by the finding that in a subsample of 19 individuals genotyped at a total of nine microsatellite loci, all individuals found to be identical at three loci were also identical at the additional six loci, while all that differed at the initial three loci also differed at the additional loci. In addition, the unbiased probability of identity; that is, the chance that two individuals drawn at random from a population with the same allele frequencies as the set of typed individuals would have an identical three-locus genotype, was estimated at 2.3×10^{-5} (Waits *et al*. 2001). We found no pairs of individuals identical for sex, mtDNA haplotype and three-locus genotype across populations. In sum, the data defined a total of 30 HV1 haplotypes of 405 bp in length.

Previously published haplotypes from wild bonobos, five from Lomako and seven from Wamba, were added to the data set (Lomako; $n = 36$, GenBank nos AF137482–AF 137486 and Wamba; $n = 17$, Hashimoto *et al*. 1996). Addition of these sequences shortened the average sequence length to 315 bp, which collapsed the 30 haplotypes found in this study to 28. Because two of these haplotypes are identical with haplotypes from Lomako and because one haplotype was found in both Lomako and Wamba, the final data set contained 37 bonobo HV1 haplotypes of known geographical origin. Sixty-nine of the 316 sequence positions were polymorphic. The polymorphisms consist of 63 transitions and six transversions. Single-base insertion/deletion events occurred at two positions.

Both region-specific and shared haplotypes were identified in bonobos. North shared single haplotypes with Northeast, Central and South, respectively (N1 = NE6, N3 = C18 and N2 = S6). The remaining haplotypes were specific to single regions, and East was the only region for which no shared haplotypes were identified. The majority (75%) of the haplotypes were present in more than one individual, with the most common haplotype, C2 from the Central region, shared by 11 individuals.

mtDNA HV1 diversity

The F_{ST} values of the three populations each sampled within the Central and East regions did not indicate significant differentiation (permutation test; $P > 0.05$). Therefore, subsequent analyses were performed using regional comparisons. The Central region had significantly greater haplotype diversity than any of the other regions (one-way ANOVA $F_{4,35} = 22.92$, $P < 0.01$, $r^2 = 0.72$; Tukey's *post-hoc* test $P < 0.05$) (Table 1). Nucleotide diversity was significantly greater in the North and significantly lower in the East, compared over regions (one-way ANOVA $F_{4,35} = 9.65$, $P < 0.01$, $r^2 = 0.52$; Tukey's *post-hoc* test, $P < 0.05$) compared to the other regions (Table 1). On average, the most similar sequences were found in the East, with a mean pairwise difference of 7.8, while an average of 12.1 differences occurred between sequences from the North.

Phylogenetic analysis

Phylogenetic trees were recontructed using all 37 haplotypes, as well as an additional seven haplotypes identified from captive bonobos, and rooted using nine sequences from chimpanzees of various subspecific affinities. In both neighbour-joining (NJ) and maximum likelihood (ML) tree reconstructions, bonobo genotypes clustered into two major clades (A and B) separated by a net sequence divergence of 3.2% (Fig. 3). The two clades did not correspond absolutely with any obvious geographical pattern. All regions but

Table 1 Genetic variability of mtDNA in regional samples of bonobos

Population	*N*	Hp	$h \pm$ SD	MPD	$\pi \pm$ SD
North	36	5	0.819 ± 0.019	12.1	0.038 ± 0.004
Northeast	17	7	0.853 ± 0.053	9.8	0.031 ± 0.007
Central	63	16	0.923 ± 0.014	10.3	0.032 ± 0.003
South	26	7	0.813 ± 0.036	9.5	0.030 ± 0.002
East	15	5	0.781 ± 0.064	7.8	0.023 ± 0.002

Sample size (N), number of haplotypes (Hp), haplotype (gene) diversity (h), mean pairwise difference (MPD) and nucleotide diversity (π).

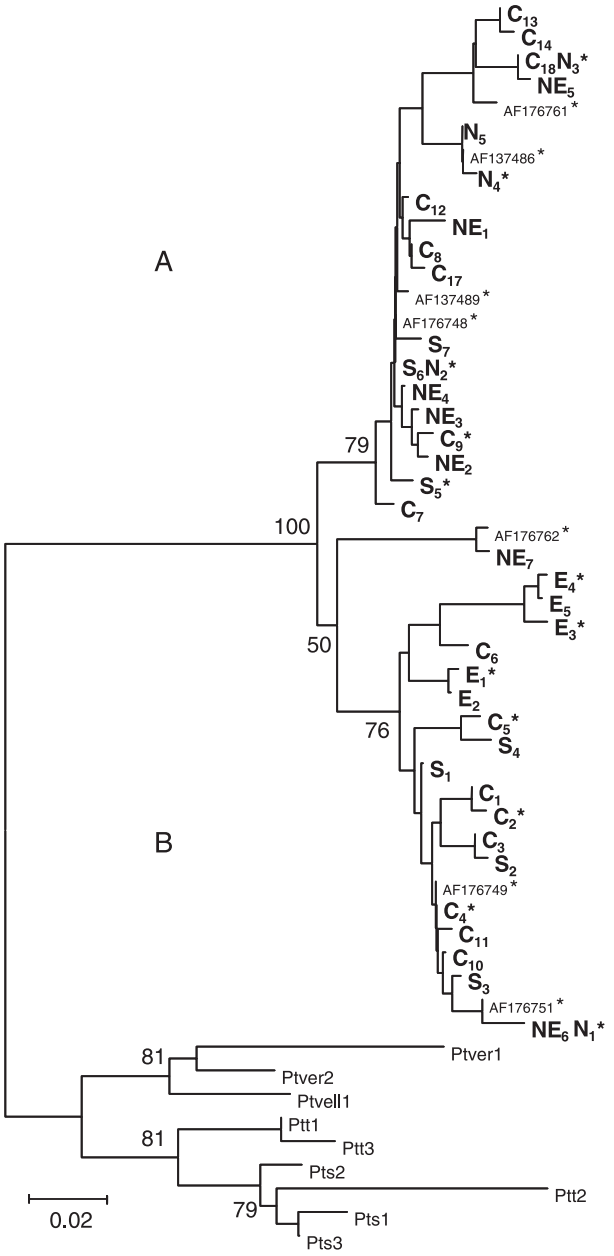

Fig. 3 Neighbour-joining tree with branch lengths generated from mitochondrial HVI haplotypes, rooted using nine sequences from the four subspecies of chimpanzees. Bootstrap support values above 50 for the major branches are shown. Haplotype IDs correspond to sample location as given in Fig. 2. Asterisks (*) indicate haplotypes which have been also observed in captive bonobos, and cases for which a GenBank Accession no. is given are those in which the haplotype has been found only in captive(s). Clades A and B are indicated.

East contained genotypes belonging in both clades. While haplotypes from the Central and South regions were distributed roughly equally in both clades, haplotypes from the North and Northeast regions were found primarily in clade A while haplotypes from the East region occurred only in clade B. The seven haplotypes observed thus far only from captive individuals are distributed throughout the tree, and haplotypes found in both captive and wild individuals are also found throughout the tree.

Geographic division of genetic diversity

In order to reveal the apportionment of the genetic diversity within the entire sample, we conducted a hierarchical AMOVA (Excoffier *et al*. 1992), an analysis based upon variance of genotype frequencies and the number of mutations between genotypes. Most of the variation (~70%) exists within sampled populations (communities) while 30% of the variation is found among different regions (Table 2). All regions are significantly differentiated from one another (permutation test, $P < 0.05$; Table 3). The greatest differentiation was found between region East and each of the two Northern populations (N and NE) while the least differentiation was found between regions Central and South (Table 3).

Genetic distances (pairwise F_{ST}) among the different regions correlated significantly with geographical distance when the distance was measured around the rivers (matrix correlation test $Z = 10\,611.7$, $r = 0.746$, $P = 0.042$; Fig. 4a) but not with the straight-line distance measured directly between populations ($Z = 4924.5$, $r = 0.157$, $P = 0.294$; Fig. 4b).

Historical demographic events

In contrast to the pairwise mismatch distribution of mtDNA haplotypes in humans and eastern chimpanzees (*P. t. schweinfurthii*), which both show a unimodal distribution indicative of a recent population expansion (Rogers & Harpending 1992; Harpending *et al*. 1993; Bertorelle & Slatkin 1995; Goldberg & Ruvolo 1997), the pairwise distribution of bonobo haplotypes was clearly multimodal (Fig. 5). It thus resembled the distributions found in the western (*P. t. verus*) and central (*P. t. troglodytes*) chimpanzees, suggesting that in these taxa, populations have have had a more stable palaeodemographic history.

As would be expected given the results of the mismatch distribution analysis, for bonobos, western chimpanzees and central chimpanzees the values of Tajima's *D*, although negative, were not significantly different from zero and did not suggest a population expansion (Table 4). Comparison of nucleotide and haplotype diversity values reveals that as a species, bonobos have a level of diversity intermediate between that of humans and chimpanzees, and broadly comparable with what is found within a chimpanzee subspecies (Table 4).

Discussion

This study is the first investigation of the amount and apportionment of genetic diversity within bonobos in the

Table 2 AMOVA assessment of population genetic structure in bonobos

Source of variation	d.f.	SSD	% of variation	*P*-value
Among regions	4	360	30.61	< 0.0001
Among populations within regions	4	16	−2.30	0.024
Within populations	151	1046	71.68	< 0.0001

d.f.: Degrees of freedom; SSD: sum of squares.

Table 3 Differentiation between regions. Below the diagonal, pairwise F_{ST} values based on Tamura–Nei distance; above the diagonal, Nei's average nucleotide difference (D_A) between regions

	North	Northeast	Central	South	East
North	—	1.64	3.86	2.13	9.98
Northeast	0.117	—	6.73	4.13	12.8
Central	0.252	0.427	—	0.92	4.5
South	0.138	0.316	0.083	—	6.0
East	0.498	0.638	0.307	0.419	—

wild. The distribution of mitochondrial variation is clearly affected by population structure, as significant differentiation was found among the five regional samples with some 30% of the variance occurring among regions. However, a geographical clustering of the population structure is not obvious from the tree reconstruction. Haplotypes were shared among all regions excepting East, and the phylogenetic analysis revealed a tree in which haplotypes were intermixed with little regard to geographical origin, with the notable exception of the close relationships among the haplotypes found in the East region. These results are reminiscent of those obtained in studies of chimpanzee subspecies, in which individuals sampled up to 1000 km apart were found to have highly similar or identical haplotypes (Goldberg & Ruvolo 1997; Gagneux *et al*. 2001).

Because African apes are poor swimmers in comparison with other mammals of similar body size, rivers have been suggested to influence inter- and intraspecific patterns of genetic differentiation, as in the separation by the Congo River of bonobos and chimpanzees and the apparent division of chimpanzee subspecies by rivers (but see Gagneux *et al*. 1999). As further support, although no geographical structure to mtDNA variation in the eastern chimpanzee subspecies has been detected, the greatest genetic differentiation there was found between haplotypes sampled across rivers (Goldberg & Ruvolo 1997).

The analysis presented here offers some intriguing hints concerning the influence of riverine barriers upon gene flow among bonobo populations. Genetic distance, as measured by pairwise F_{ST} values between regional samples, did correlate with geographical distance when intervening distance was measured around rivers presenting effective

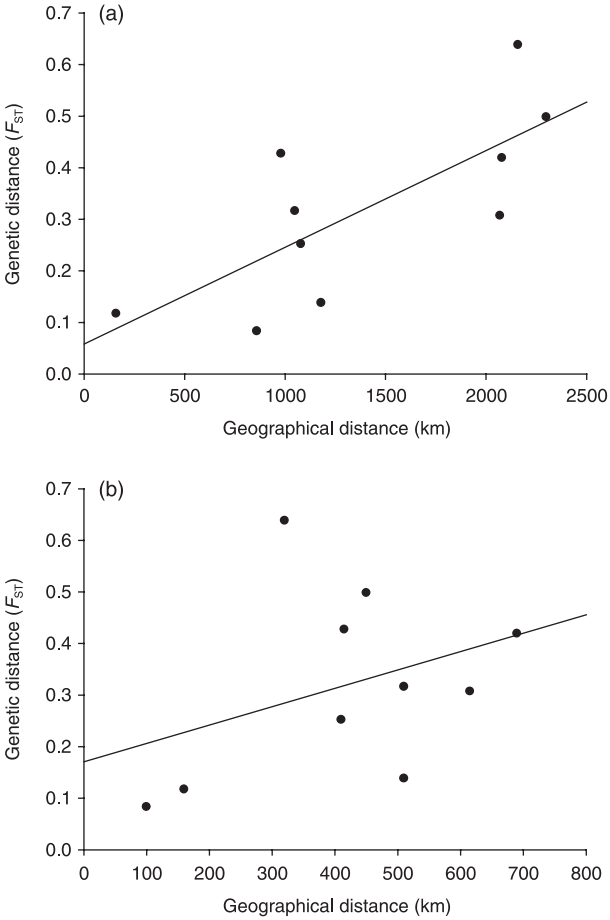

Fig. 4 Plots of genetic distance (pairwise F_{ST}) vs. geographical distance between regional samples of bonobos measured detouring around rivers (a) and as a straight-line distance (b). The relationship in (a) is significant (matrix correlation test $r = 0.746$, $P = 0.04$). The lines indicate the least-squares regression estimates.

current–day barriers. In contrast, no correlation was found when geographical distance was measured directly in a straight line between sampling localities. This result is highlighted by the large genetic differentiation and geographical distance of the East regional sample to all other regions. This suggests that the Lomami River, which extends beyond territory habitable by bonobos (Fig. 2), has presented a particularly effective barrier to bonobos. However, this may not be the case with all of the rivers. A notable example is the case of the Central and South regions,

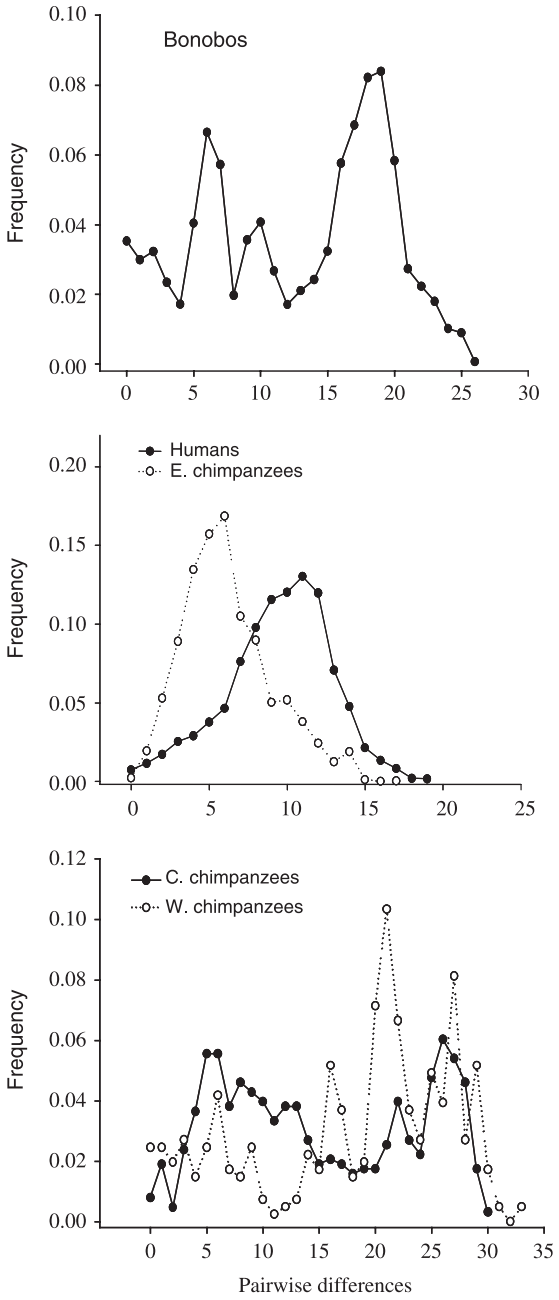

Fig. 5 Mismatch distribution plots describing the frequency of pairwise substitutional differences among haplotypes for bonobos, humans, eastern chimpanzees (*Pan troglodytes schweinfurthii*), central chimpanzees (*P. t. troglodytes*) and western chimpanzees (*P. t. verus*). The number of haplotypes used per taxa is given in Table 4.

which are separated by the Lukenie River but show low levels of genetic differentiation.

Unlike most mammals, bonobos and chimpanzees are characterized by the occurrence of female rather than male dispersal, a feature which in combination with high levels of effective gene flow and/or long-distance dispersal events prevents the establishment of highly structured phylogeographical patterns to mtDNA variation. Instead of ongoing gene flow, an alternative explanation for the lack of a phylogeographical pattern would invoke incomplete lineage sorting and suggest that insufficient time since separation has passed to allow establishment of a pattern. The close relationship between the Central and South populations could be, in this scenario, a result of a recent separation event. An alternative explanation is that rivers in the Congo Basin, although often large and nonseasonal in water volume, are slow-moving and highly meandering due to little change in elevation and can therefore be expected to have altered their course over time. Such changes may result in land masses changing from one side of the river to the other, allowing the exchange of genetic material between previously separated populations. Distinguishing recent separation events or ongoing gene flow is not easy with the use of single genetic locus for analysis, and the difficulties in using present-day genetic variation and geographical features to infer past patterns must be emphasized (Leonard *et al*. 2000).

The amount of mitochondrial diversity within the wild population of bonobos is intermediate between that of humans and chimpanzees, and comparable to what is found within a chimpanzee subspecies. This is in contrast to what was found in a comparison of average nucleotide diversity over 50 noncoding nuclear segments in chimpanzees, bonobos and humans (Yu *et al*. 2003). In that study, bonobos were found to have the lowest diversity, but the sampling of bonobos was limited to nine individuals, some of which were likely to be related. Pairwise mismatch distribution analysis and values for Tajima's *D* suggest that the bonobo population size has been steady over the long term, or at least has not fluctuated in population size to an extent sufficient to cause a bottleneck/expansion effect detectable in the distribution of variation. In this respect, bonobos resemble central and western chimpanzees, rather than eastern chimpanzees which, like humans, show signs of a population expansion during the Pleistocene (Rogers & Harpending 1992; Harpending *et al*. 1993; Bertorelle & Slatkin 1995; Goldberg & Ruvolo 1997; Harpending *et al*. 1998). Because the phylogenetic tree shows little geographical structure, the probable location of a forest refuge during the oscillating arid periods of forest reduction is not easy to deduce. The fact that the Central and South regions contain haplotypes clustering in both major clades while haplotypes from the other regions are limited mainly to a single clade could be suggestive of a Central/Southern source of recolonization after a reduction of habitat, but the areas bordering nearly the entire length of the Congo River are considered probable refugia (Maley 1996).

The lack of a distinct geographical pattern to the mtDNA variation means that this analysis does not provide information that would allow inference of the regional origin of

Population	Hp	π (%)	MPD	*D*	*P*-value
Chimpanzees	130	7.0	20.75	−0.834	NS
Central	31	5.3	16.09	−0.829	NS
Western	32	6.4	19.88	−0.233	NS
Eastern	70	2.5	8.12	−1.947	$P < 0.05$
Bonobo	37	4.6	14.5	−0.461	NS
Human	78	3.0	9.47	−1.770	$0.10 > P > 0.05$

Table 4 Comparison among bonobos, chimpanzees and humans. Number of haplotypes (no. Hp), nucleotide diversity as percentage (π (%)), mean pairwise difference (MPD) and Tajima's D (*D*)

bonobos in captivity. The sequences observed from individuals of captive origin are scattered throughout the phylogenetic tree, but as other sequences obtained from a single locality are also distributed throughout the tree (e.g. Central), it is impossible to infer whether bonobos in captivity are likely to originate from few or many localities. With regard to future conservation efforts in the wild, the mtDNA diversity found within the populations in the Central region encompass almost the entire variation found within the bonobo population as a whole (Fig. 3), which could be used to suggest that one large protected area (e.g. Salonga National Park) in the Central distribution range would be sufficient to maintain the existing bonobo mtDNA variation. However, the fact that a structure to the apportionment of mtDNA diversity is present (30% between regions) in a highly mobile, female-dispersing mammal for which mtDNA variation portrays the part of the genome experiencing maximum gene flow, other loci representing other parts of the genome could show greater structure (e.g. humans, Seielstad *et al*. 1998). Further insights into the population history and the nature of the distribution of genetic variation in bonobos are likely to come from analysis of rapidly evolving variation on the Y-chromosome as well as of sequence variation at multiple autosomal loci (Yu *et al*. 2003; Fischer *et al*. 2004).

Acknowledgements

We are grateful to Mark Stoneking and Mats Bjorklund for useful comments on the manuscript. Annette Abraham, Mirko Bergmann, Bernhard Güntner, Claudia Lang and Heike Siedel assisted in laboratory analyses and Dieter Lukas assisted with the probability of identity calculation. We thank Brenda Bradley and Olaf Thalmann for fruitful discussions and Daniel Stahl for statistical advice. This research was supported partly by the Deutsche Forschungsgemeinschaft (DFG) grant HO1159/6–3 to G. Hohmann, by Uppsala University/Department for Animal Ecology and by the Max Planck Society. We thank the Institute Congolaise pour la Conservation de la Nature and the government of the Democratic Republic of Congo for assisting with and permitting the sampling programme.

References

Avise JC (2000) *Phylogeography*. Harvard University Press, Cambridge, MA.

Ayers JM, Clutton-Brock TH (1992) River boundaries and species range size in Amazonian primates. *American Naturalist*, **140**, 531–537.

Bertorelle G, Slatkin M (1995) The number of segregating sites in expanding human populations, with implications for estimates of demographic parameters. *Molecular Biology and Evolution*, **12**, 887–892.

Boesch C (2002) Behavioural diversity in Pan. In: *Behavioural Diversity in Chimpanzees and Bonobos* (eds Boesch C, Hohmann G, Marchant LF), pp. 1–8. Cambridge University Press, Cambridge.

Bradley BJ, Chambers KE, Vigilant L (2001) Accurate DNA-based sex identification of apes using non-invasive samples. *Conservation Genetics*, **2**, 179–181.

Bradley BJ, Vigilant L (2002) The evolutionary genetics and molecular ecology of chimpanzees and bonobos. In: *Behavioural Diversity in Chimpanzees and Bonobos* (eds Boesch C, Hohmann G, Marchant LF), pp. 259–276. Cambridge University Press, Cambridge.

Brown WM (1983) Evolution of animal mitochondrial DNA. In: *Evolution of Genes and Proteins* (eds Nei M, Koehn RK), pp. 62–88. Sinauer, Sutherland, MA.

Colyn MM, Gautier-Hion A, Verheyen W (1991) A re-appraisal of paleoenvironmental history in central Africa: evidence for a major fluvial refuge in the Zaire Basin. *Journal of Biogeography*, **18**, 403–407.

Deinard AS, Kidd KK (2000) Identifying conservation units within captive chimpanzee populations. *American Journal of Physical Anthropology*, **111**, 25–44.

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.

Fischer A, Wiebe V, Pääbo S, Przeworski M (2004) Evidence for a complex demographic history of chimpanzees. *Molecular Biology and Evolution*, **21**, 799–808.

Forbes S, Boyd D (1997) Genetic structure and migration in native and reintroduced Rocky Mountain wolf populations. *Conservation Biology*, **11**, 1226–1234.

Gagneux P, Gonder MK, Goldberg TL, Morin PA (2001) Gene flow in wild chimpanzee populations: what genetic data tell us about chimpanzee movement over space and time. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, **356**, 889–897.

Gagneux P (1998) *Population Genetics of West African Chimpanzees* (Pan troglodytes verus). PhD Thesis, University of Basel, Switzerland.

Gagneux P, Wills C, Gerloff U *et al*. (1999) Mitochondrial sequences show diverse evolutionary histories of African hominoids. *Proceedings of the National Academy of Sciences USA*, **96**, 5077–5082.

Georgiadis N, Bischof L, Templeton A *et al.* (1994) Structure and history of African elephant populations. I. Eastern and southern Africa. *Journal of Heredity*, **85**, 100–104.

Gerloff U, Hartung B, Fruth B, Hohmann G, Tautz D (1999) Intracommunity relationships, dispersal pattern and paternity success in a wild living community of Bonobos (*Pan paniscus*) determined from DNA analysis of faecal samples. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **266**, 1189–1195.

Goldberg TL, Ruvolo M (1997) Molecular phylogenetics and historical biogeography of east African chimpanzees. *Biological Journal of the Linnean Society*, **61**, 301–324.

Gonder MK (2000) *Evolutionary Genetics of Chimpanzees* (Pan troglodytes) *in Nigeria and Cameroon*. PhD Thesis, City University of New York, New York.

Gonder MK, Oates JF, Disotell TR *et al.* (1997) A new west African chimpanzee subspecies? *Nature*, **388**, 337.

Grubb P (1990) Primate geography in the Afro-tropical forest biome. In: *Vertebrates in the Tropics* (eds Peters G, Hutterer R). Museum Alexander Koenig, Bonn.

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

Harpending HC, Batzer MA, Gurven M *et al.* (1998) Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences USA*, **95**, 1961–1967.

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. *Current Anthropology*, **34**, 483–496.

Hasegawa M, Kishino H, Yano TA (1985) Dating of the human ape-split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hashimoto C, Furuichi T, Takenaka O (1996) Matrilineal kin relationship and social behavior of wild bonobos (*Pan paniscus*): sequencing the D-loop region of mitochondrial DNA. *Primates*, **37**, 305–318.

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics*, **1**, 583–589.

Kano T (1992) *The Last Ape*. Stanford University Press, Stanford, CA.

Kilger C, Pääbo S (1997) Direct DNA sequence determination from total genomic DNA. *Nucleic Acids Research*, **25**, 2032–2034.

Kocher TD, Wilson AC (1989) Sequence evolution of mitochondrial DNA in humans and chimpanzees: control region and a protein-coding region. In: *Evolution of Life: Fossils, Molecules and Culture* (eds Osawa S, Honjo T), pp. 391–413. Springer-Verlag, Tokyo.

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

Lehman N, Wayne RK (1991) Analysis of coyote mitochondrial DNA genotype frequencies — estimation of the effective number of alleles. *Genetics*, **128**, 405–416.

Leonard JA, Wayne RK, Cooper A (2000) Population genetics of ice age brown bears. *Proceedings of the National Academy of Sciences USA*, **97**, 1651–1654.

Maley J (1996) The African rain forest — main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh*, **104B**, 31–73.

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

Mayr E (1963) *Animal Species and Evolution*. Cambridge University Press, Cambridge.

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.

Morin PA, Moore JJ, Chakraborty R *et al.* (1994) Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science*, **265**, 1193–1201.

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

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA*, **76**, 5269–5273.

Oates JF (1988) The distribution of Cercopithecus monkeys in West African forests. In: *A Primate Radiation: Evolutionary Biology of the African Guenons* (eds Gautier-Hion A, Boulier F, Gautier JP, Kingdon J), pp. 79–103. Cambridge University Press, Cambridge.

Osman-Hill WC (1969) The nomenclature, taxonomy, and distribution of chimpanzees. In: *The Chimpanzee*, Vol. 1 (ed. Bourne GH), pp. 22–49. Karger, Basel.

Reinartz GE, Karron JD, Phillips RB, Weber JL (2000) Patterns of microsatellite polymorphism in the range-restricted bonobo (*Pan paniscus*): considerations for interspecific comparison with chimpanzees (*P. troglodytes*). *Molecular Ecology*, **9**, 315–328.

Rogers AR, Fraley AE, Bamshad MJ, Watkins WS, Jorde LB (1996) Mitochondrial mismatch analysis is insensitive to the mutational process. *Molecular Biology and Evolution*, **13**, 895–902.

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

Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics*, **142**, 1357–1362.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Roy MS, Geffen E, Smith D, Ostrander EA, Wayne RK (1994) Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution*, **11**, 553–570.

Rozas J, Rozas R (1999) DNASP, version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.

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.

Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, **18**, 502–504.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN, Version* 2000: A Software for Population Genetics Data Analysis (available at: http://anthro.unige.ch/arlequin).

Schwarz E (1934) On the local races of chimpanzees. *Annals and Magazine of Natural History, London*, **13**, 576–583.

Seielstad MT, Minch E, Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. *Nature Genetics*, **20**, 278–280.

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

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

Stanford CB (1998) The social behavior of chimpanzees and bonobos — empirical evidence and shifting assumptions. *Current Anthropology*, **39**, 399–420.

Sullivan KM, Mannucci A, Kimpton CP, Gill P (1993) A rapid and quantitative DNA sex test — fluorescence-based PCR analysis of X–Y homologous gene amelogenin. *Biotechniques*, **15**, 636–641.

Takahata N, Palumbi SR (1985) Extranuclear differentiation and gene flow in the finite island model. *Genetics*, **109**, 441–457.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Telfer PT, Souquière S, Clifford SL *et al.* (2003) Molecular evidence for deep phylogenetic divergence in *Mandrillus sphinx*. *Molecular Ecology*, **12**, 2019–2024.

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

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

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC (1989) Mitochondrial DNA sequences in single hairs from a southern African population. *Proceedings of the National Academy of Sciences USA*, **86**, 9350–9354.

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. *Science*, **253**, 1503–1507.

Waits LP, Luikart G, Taberlet T (2001) Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology*, **10**, 249–256.

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

Wu C-I, Ting C-T (2004) Genes and speciation. *Nature Reviews Genetics*, **5**, 114–122.

Yu N, Jensen-Seaman MI, Chemnick L *et al.* (2003) Low nucleotide diversity in chimpanzees and bonobos. *Genetics*, **164**, 1511–1518.

J. Eriksson is a student at the Department of Animal Ecology at the University of Uppsala and conducted the study on bonobo mtDNA variation as part of his PhD research in collaboration with the MPI for Evolutionary Anthropology. G. Hohmann is a field researcher interested in the behavioural ecology of great apes and the use of genetic techniques to produce data relevant to the understanding of social behaviour. C. Boesch is a field researcher interested in understanding the factors affecting the evolution of cultures and social structures and the application of genetic, endocrinological and medical research to wild apes. L. Vigilant uses genetic and behavioural data to address topics such as the distribution of paternity or patterns of relatedness among individuals in primate social groups and the implications for the evolution of social behaviour.
