The Complex Evolutionary History of Gorillas: Insights from Genomic Data

O. Thalmann, A. Fischer,* F. Lankester,† S. Pääbo,* and L. Vigilant**

*Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany; and †Limbe Wildlife Centre, Limbe, SWP, Cameroon

Relatively little is known about the evolutionary and demographic histories of gorillas, one of our closest living relatives. In this study, we used samples from both western (*Gorilla gorilla*) and eastern (*Gorilla beringei*) gorillas to infer the timing of the split between these geographically disjunct populations and to elaborate the demographic history of gorillas. Here we present DNA sequences from 16 noncoding autosomal loci from 15 western gorillas and 3 eastern gorillas, including 2 noninvasively sampled free-ranging individuals. We find that the genetic diversity of gorillas is similar to that of chimpanzees but almost twice as high as that of bonobos and humans. A significantly positive Fu & Li's D was observed for western gorillas, suggesting a complex demographic history with a constant, long-term population size and ancestral population structure. Among different population-split scenarios, our data suggest a complex history of western and eastern gorillas including an initial population split at around 0.9–1.6 MYA and subsequent, primarily male-mediated gene flow until approximately 80,000–200,000 years ago. Furthermore, simulations revealed that more gene flow took place from eastern to western gorilla populations than vice versa.

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

Introduction

Over the last decades, population genetic data have helped us to gain a better understanding of the evolutionary history of a variety of species, including the great apes and ourselves (Cann et al. 1987; Ruvolo et al. 1991; Vigilant et al. 1991; Hey 1998; Avise 2000; Gagneux 2002; Finlayson 2005). Thanks to technical advances in molecular genetics, it has become possible to use noninvasively collected samples for genetic studies of wild populations (Taberlet et al. 1996; Constable et al. 2001; Nsubuga et al. 2004; Smith and Morin 2005). This is a considerable advantage because the amount and patterns of genetic variation present in wild populations clarify the factors that have influenced their diversification and may contribute to assessment of the viability of populations.

Although short nuclear markers such as short tandem repeats and segments of the multicopy mitochondrial DNA (mtDNA) have been successfully analyzed from noninvasively collected samples (Taberlet et al. 1997; Warren et al. 2001; Eriksson et al. 2004; Goossens et al. 2006), the utility of such samples for amplification and sequencing of larger (>800 bp) autosomal segments has been little explored. Analysis of multiple noncoding autosomal loci is rapidly becoming a routine approach for population genetics studies (Machado et al. 2002; Pluzhnikov et al. 2002; Broughton and Harrison 2003; Yu et al. 2003; Fischer et al. 2004; Hey and Nielsen 2004; Yu et al. 2004; Fischer et al. 2006) because each independent autosomal locus has its own evolutionary history and reflects as well the history of both sexes. Furthermore, because each single locus could be affected by positive selection (Chen and Li 2001; Lercher and Hurst 2002; Rosenberg and Nordborg 2002), only data from multiple loci from across the genome reveal a comprehensive picture of the history of a population (Hare 2001; Brumfield et al. 2003).

As our closest relatives, African great apes have attracted considerable attention from geneticists. Almost all previous genetic investigations have revealed a higher genetic diversity in gorillas and chimpanzees as compared with humans, regardless of the marker examined (Morin et al. 1994; Ruvolo et al. 1994; Deinard and Kidd 1999; Gagneux et al. 1999; Kaessmann et al. 2001; Noda et al. 2001; Yu et al. 2002; Fischer et al. 2004; Yu et al. 2004; Fischer et al. 2006; but see also Wise et al. 1997). Whereas a wealth of literature exists focusing on our own species and our closest evolutionary relative, the chimpanzee (Cann et al. 1987; Morin et al. 1994; Jorde et al. 1995; Goldberg and Ruvolo 1997; Gagneux 2002; Yu et al. 2003; Harding and McVean 2004; International Human Genome Sequencing Consortium 2004; The Chimpanzee Sequencing and Analysis Consortium 2005; Voight et al. 2005), only a handful of genetic studies have focused intensively on the gorilla (Garner and Ryder 1996; Saltonstall et al. 1998; Jensen-Seaman and Kidd 2001; Jensen-Seaman et al. 2003; Clifford et al. 2004; Yu et al. 2004).

The 2 currently recognized gorilla species (western gorilla, *Gorilla gorilla*, and eastern gorilla, *Gorilla beringei*) (Groves 2001; Grubb et al. 2003) live in equatorial Africa and are separated by about 1,000 km (fig. 1). They inhabit diverse forest habitats, resulting in discernable differences in morphology, ecology, and social and feeding behavior (reviewed in Doran and McNeilage 1998; Taylor and Goldsmith 2003). Whereas eastern gorilla populations are small and fragmented, making any contemporary gene flow between populations very unlikely, western gorillas occupy a rather continuous habitat, enabling gene flow within respective forest blocks.

The majority of genetic investigations of gorillas have been done using the mtDNA hypervariable region I, but, due to the occurrence of insertions of mtDNA fragments into the nuclear genome (numts; Lopez et al. 1994), conclusions derived from analysis of this particular locus should be regarded with caution in gorillas (Jensen-Seaman et al. 2004; Thalmann et al. 2004, 2005). This leaves only 2 investigations (Jensen-Seaman et al. 2003; Yu et al. 2004) of genetic variation of the nuclear genome in gorillas. In the latter study, the authors sequenced 50 noncoding autosomal loci in 15 representatives of the western gorillas. The authors relied upon samples from captive gorillas, and the scarcity of captive eastern gorillas meant that potential genetic differences between western and eastern gorillas could not be investigated. Evidence derived from analysis of coding regions in the mitochondrial genome (COII,

Key words: gorilla, ancestral population structure, demography, population split, nuclear sequences, fecal samples.

E-mail: thalmann@eva.mpg.de.

Mol. Biol. Evol. 24(1):146–158. 2007
doi:10.1093/molbev/msl160
Advance Access publication October 25, 2006

© The Author 2006. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org

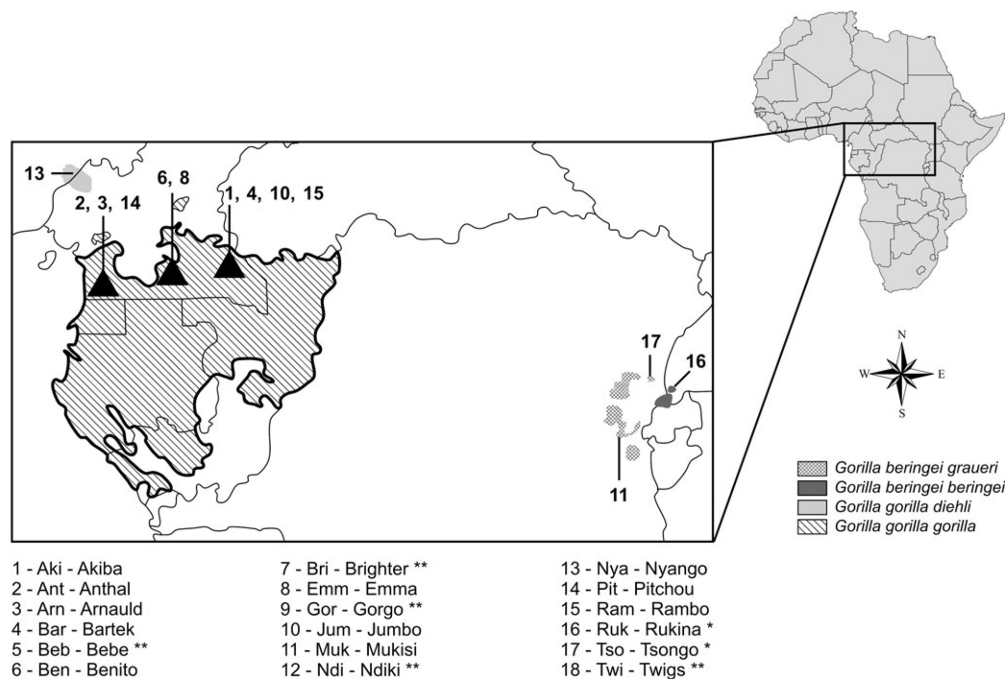

FIG. 1.—Map showing the current geographical distribution of gorillas in Equatorial Africa and their taxonomic classification. The approximate origins of most of the sampled individuals are also shown. Individuals from which fecal samples were obtained are indicated (*) as well as individuals for which localities were unknown (**).

Ruvolo et al. 1994; NADH5, Jensen-Seaman et al. 2003) suggested that the coalescence time of 2 MYA for these genes for western and eastern gorillas is comparable to that found for chimpanzees and bonobos (Ruvolo 1997). Because such time estimates represent the time to the most recent common ancestor of the respective genetic locus and because the coalescence of genes predates the actual population split (Edwards and Beerli 2000; Nichols 2001), there still might be migration between diverging populations until gene flow completely ceases. In the case of chimpanzees and bonobos, it has been suggested that the population split dates back to only about 800,000 years (Fischer et al. 2004; Won and Hey 2005). Furthermore, analyses of a single X-chromosome locus and 8 noncoding nuclear loci revealed a clear reciprocal monophyly for chimpanzees and bonobos but not for gorillas (Kaessmann et al. 2001; Jensen-Seaman et al. 2003), suggesting that the population split between western and eastern gorillas may have occurred recently. Using mitochondrial data, we have recently refined the time to the most recent common ancestor of western and eastern gorillas' mtDNA to ~1.3 MYA (Thalmann et al. 2005), but it remains unclear for how much longer gene flow might have persisted between the 2 gorilla populations.

In the present study, we analyzed noncoding, putatively selectively neutral autosomal loci from 18 individuals of the genus *Gorilla* in order to investigate patterns of genetic diversity in gorillas and to compare these with those of other African great apes. Moreover, we aimed to elaborate on the demographic history of gorillas and to discuss the relationship of the 2 geographically defined gorilla species.

Materials and Methods

Sample Acquisition and DNA Extraction

We extracted DNA from 3 different sample materials: blood, liver, and feces. We obtained the majority of the gorilla blood samples (12 of 15) during routine medical checkups of residents of the Limbe Wildlife Center (LWC), Cameroon, in 2004. These animals represent western gorillas (*G. gorilla*) and include individuals of both currently recognized western gorilla subspecies (11 *Gorilla gorilla gorilla* and 1 *Gorilla gorilla diehli*). Based on the records from the LWC, we could broadly infer the regional geographic origins of most of the Cameroonian gorillas although their exact localities could not be retraced. In addition, we used blood samples from 3 western gorillas from the Leipzig Zoo (Germany) and also a liver sample from a single deceased eastern gorilla (*Gorilla beringei graueri*) from the Zoo Antwerp (Belgium). To increase the sample size of eastern gorillas, we used noninvasively collected fecal samples from 1 eastern lowland gorilla (*G. beringei graueri*) from the area around Mt. Tshiaberimu (Democratic Republic of Congo) as well as from 1 eastern mountain gorilla (*Gorilla beringei beringei*) from the Bwindi Impenetrable National Park (Uganda). Figure 1 shows the current distribution of gorillas in central Africa as well as the inferred geographic origins of the gorilla samples used in this study.

The extraction procedures varied according to sample type. The 12 blood samples from the LWC were preextracted in Limbe by transferring 5 ml fresh blood into 5× volume of cold erythrocyte lysis buffer (EL buffer, Qiagen), inverting and then centrifuging the mixture for 10 min at 1,500 rpm. After discarding the supernatant,

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

Table 1
Polymorphism and Summary Statistics for Each Locus

Locus	Chromosome[a]	Length	Polymorphic Sites	Singletons	π (%)		θ_W (%)		Tajima's *D*		Fu & Li's D	
1	12q13.1	855 (*820*)	4 (*5*)	0 (*1*)	0.082	*0.083*	0.116	*0.147*	−0.741	*−1.128*	1.067	*0.197*
2	6p23	1,059 (*936*)	7	2 (*3*)	0.106	*0.08*	0.164	*0.18*	−1.027	*−1.581*	−0.199	*−0.242*
4	20p12.3	996 (*967*)	7	0	0.152	*0.178*	0.175	*0.175*	−0.378	*0.060*	1.322	*1.310*
5	5q23.3	1,430 (*1,414*)	7	3	0.069	*0.072*	0.122	*0.119*	−1.253	*−1.128*	−0.199	*−0.473*
6	7p21.3	841 (*779*)	4 (*3*)	0	0.189	*0.18*	0.118	*0.093*	1.518	*2.089**	1.067	*0.931*
7	19q12	1,183 (*803*)	0 (*5*)	0	0	*0.098*	0	*0.15*	—	*−0.908*	—	*1.152*
10	4q24	788 (*694*)	9 (*7*)	0	0.317	*0.246*	0.284	*0.245*	0.364	*0.009*	1.442	*1.310*
11	11p14.1	814 (*758*)	4 (*6*)	1 (*3*)	0.104	*0.127*	0.122	*0.191*	−0.382	*−0.910*	−0.007	*−1.327*
13	5q12.1	754 (*729*)	8	0	0.333	*0.374*	0.263	*0.265*	0.792	*1.206*	1.386	*1.374*
14	4q31.3	782 (*759*)	5	1	0.09	*0.083*	0.159	*0.159*	−1.158	*−1.242*	0.227	*0.197*
15	14q23.2	798 (*0*)	4 (*0*)	2 (*0*)	0.084	—	0.124	—	−0.823	—	−1.081	—
16	9p23	796 (*764*)	3 (*5*)	0	0.132	*0.157*	0.094	*0.158*	0.954	*−0.019*	0.792	*1.052*
17	2q12.3	743	4	1	0.123	*0.126*	0.134	*0.13*	−0.193	*−0.067*	−0.007	*−0.039*
20	10p11.21	678 (*0*)	9 (*0*)	2 (*0*)	0.446	—	0.33	—	1.082	—	0.152	—
21	17p12	713	4	0	0.225	*0.218*	0.139	*0.135*	1.554	*1.501*	1.067	*1.052*
22	22q12.3	787 (*753*)	0	0	0	*0*	0	*0*	—	—	—	—
Total		14,017 (*11,632*)	79 (*73*)	12 (*12*)								
Average					0.153	*0.144*	0.147	*0.153*	0.022	*−0.163*	0.502*	*0.500**

NOTE.—Numbers and values obtained from the data set including the fecal samples (further details see text) are indicated with italic font. Loci 1–10 are from Frisse et al. (2001) and loci 11–22 from Voight et al. (2005). $^*P < 0.05$.

[a] The chromosomal location according to the human genome.

we dissolved the remaining pellet in 2× volume of cold EL buffer and repeated the centrifugation and discarding step. Finally, we dissolved the pellet in 320 μl of a tissue lysis buffer (ATL buffer, Qiagen), added 20 μl proteinase K (20 mg/ml, Qiagen), let the mixture incubate at room temperature overnight, and froze the stabilized DNA until transport. Both the Leipzig Zoo samples and the stabilized gorilla DNAs from the LWC were extracted using a commercially available kit according to manufacturer's instructions (QIAamp DNA Mini Kit, Qiagen).

A different extraction procedure (Sambrook et al. 1989) was used to obtain genomic DNA from the gorilla liver sample and included a SDS–Proteinase K digestion step and a Phenol/Chloroform treatment as detailed in Thalmann et al. (2004). The amount of DNA in each extract was quantified with a spectrophotometer (Biophotometer, Eppendorf). Finally, both fecal samples were extracted using the QIAamp DNA Stool kit (Qiagen), as described in Nsubuga et al. (2004), and the DNA quantification was done using a quantitative polymerase chain reaction (PCR) assay (Taq Man, Applied Biosystems) (Morin et al. 2001).

Genomic Markers, PCR Amplification, and Sequencing

We investigated the patterns of genetic diversity in wild gorillas by sequencing 16 autosomal segments located on different chromosomes (table 1). These segments have previously been shown to be noncoding and selectively neutral in other great apes (Frisse et al. 2001; Fischer et al. 2004; Voight et al. 2005). Each nuclear segment was amplified using 2 different primer sets in independent PCRs, an approach that enabled us to avoid null alleles (the nonamplification of 1 allele due to mutations in primer-binding sites). The primers were designed based on the human and chimpanzee genome sequences, and the primer sequences can be obtained upon request from the authors. Each 20 μl PCR reaction mixture consisted of approximately 10 ng of genomic DNA, 1× *SUPER TAQ* PCR buffer (containing $MgCl_2$), an additional 0.875 mM of $MgCl_2$, 0.2 μM each primer, 2.50 μM each dNTP and 0.33 U *SUPER TAQ* (HT Biotechnology, Cambridge, England), which was premixed 2:1 with 1 μg/μl TaqStart monoclonal antibody (BD Biosciences Clontech) for a hot start protocol. For amplifications using DNAs extracted from feces, reaction mixes contained 50–400 pg of template DNA and 16 μg bovine serum albumin (Sigma-Aldrich Inc., Germany). PCR amplifications were performed on a Peltier thermocycler, PTC 200 (MJ Research) including the following steps: initial denaturation at 95 °C for 3 min; 35–45 cycles of denaturation at 94 °C for 15 s, primer annealing at 57–65 °C for 1 min, elongation at 72 °C for 2 min, and a final elongation step of 15 min at 72 °C. The products were visualized under UV light on 1% agarose gels, and products of the expected size were purified with the Millipore MultiScreen Filtration system (Millipore Corp.). In rare cases in which subsequent analysis revealed insertion/deletion polymorphisms, we cloned the respective PCR products using the TOPO TA cloning kit (Invitrogen). A minimum of 10 clones were reamplified in a colony PCR with M13 primers (Kilger et al. 1997), and finally colony PCR products were purified using a Bio Robot 9600 (Qiagen). For all amplifications, we cycle sequenced the purified products using both the PCR primers and additional internal sequencing primers. This strategy enabled us to determine each nucleotide from both DNA strands. In rare cases of irresolvable sequencing problems toward the 5′ or 3′ end on one or the other DNA strand, we obtained these sequence segments by sequencing at least 2 independent PCRs from 1 DNA strand. We analyzed the samples on an automated capillary sequencer DNA Analyzer 3730 (Applied Biosystems).

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

Data Analysis

We used SeqScape Software v2.5 (Applied Biosystems) to 1) visualize the electropherograms, 2) evaluate the quality of each nucleotide in the sequence, 3) detect and validate nucleotide changes, and 4) construct a consensus sequence for each individual on each locus. Polymorphic sites were confirmed manually by examining the electropherograms of each sequence read by eye. We finalized the sequence alignments using Bioedit v7.0.5 (Hall 1999). We tested each polymorphic site within our data set for Hardy–Weinberg equilibrium using Arlequin v3.0 (Schneider et al. 2000; http://anthro.unige.ch/arlequin/).

We estimated standard diversity indices such as π (Nei 1987) and θ_W (Watterson 1975). Under the assumptions of the standard neutral model, we used θ_W to calculate the effective population size (N_e) given the equation $\theta_W = 4N_e\mu$, where μ is the population mutation rate. The mutation rate can be calculated by the equation $\delta = 2T\mu$, where T is the divergence time to a respective outgroup (in this case human) and δ is the average number of nucleotide substitutions per site (Nei 1987) between human and gorillas for each locus. To test each locus for neutrality, we used the HKA test (Hudson et al. 1987), Tajima's *D* (Tajima 1989b), and also Fu & Li's D with an outgroup (Fu and Li 1993). The diversity and the summary statistic calculations were performed with DnaSP v4.0 (Rozas J and Rozas R 1999; http://www.ub.es/dnasp). The haplotypic state of each individual was inferred on each locus by using the program Phase v2.1 (Stephens et al. 2001; Stephens and Donelly 2003), and using those haplotypes, further phylogenetic analyses were performed with MEGA v2 (Kumar et al. 2001) and Network v4.1.0.8 (http://www.fluxus-technology.com). We also used MEGA v2 to calculate genetic distance matrices assuming the Kimura 2-parameter substitution model (Kimura 1980) between each pair of individuals for each locus and transferred the average genetic distance over all loci into a multidimensional scaling analysis (MDS) performed with SPSS v13 (LEAD Technologies Inc., http://www.spss.com). Such an analysis might highlight potential clusters that subsequently can be compared with the origin of the samples, and thus, a potential correlation to population structure might be inferred.

To investigate the demographic history of gorillas in detail, we analyzed all polymorphic loci using the simulation program HKA (http://lifesci.rutgers.edu/~heylab/). We looked for potential significant deviations from neutrality of the observed mean Tajima's *D* value and Fu & Li's D compared with the distribution of values computed based on 10,000 simulated data sets. Furthermore, we conducted a multilocus HKA test with the same program. To explore various scenarios for the population split between western and eastern gorillas, we tested the fit of our data to disparate models. The null model was the simplest but least realistic case: panmixia among gorillas until the present (panmixia model). A second model provided for limited gene flow after the coalescence, followed by an immediate population separation between western and eastern gorillas (scenario 1 in fig. 2, similar to the isolation model proposed by Wakeley and Hey 1997). The third model was similar to the second one but allowed for a more prolonged period of constant gene flow until its final cessation upon population splitting at a certain point in time (scenario 2 in fig. 2, limited gene flow model). The fourth model features an extended period of gene flow following the population split but allows for variation in direction of gene flow (scenario 3 in fig. 2, isolation with migration model). The simulation program WH (Wakeley and Hey 1997) provides a population-split time, estimated based on the variance of fixed and shared polymorphisms under the assumption of an isolation model (scenario 1). Furthermore, we used SIMCOAL v2.1 (Laval and Excoffier 2004) to simulate the timing of a population split under the assumption of the limited gene flow model (scenario 2). In addition to the 2 parameters of interest to be simulated (coalescence time and the time of a population split), running the program requires a set of parameters that are a priori given by the

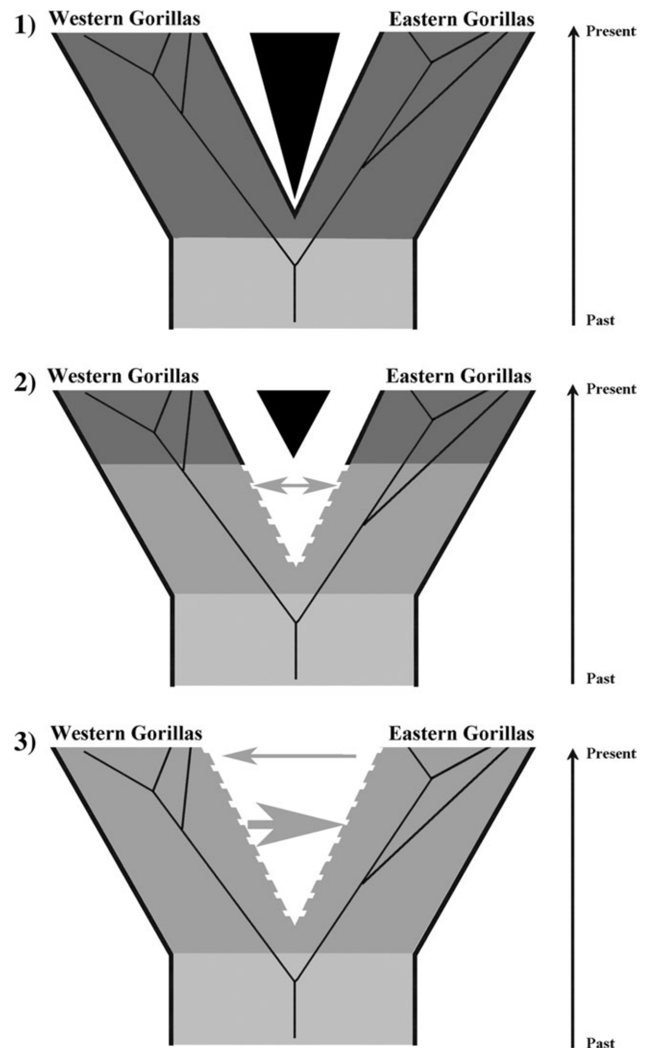

FIG. 2.—Schematic diagram illustrating 3 possible different population-split scenarios for the 2 gorilla species (1, isolation model; 2, limited gene flow model; and 3, isolation with migration model). Bold lines delineate the populations of western and eastern gorillas, and thin lines indicate genetic lineages. The shading within the diagram indicates the degree of population separation: light gray, one population and dark gray, 2 populations without gene flow among them. The dark triangle at the top of diagram 1 and 2 symbolizes a barrier to gene flow. The amount and direction of gene flow is depicted by the thickness and the direction of the gray arrows, respectively.

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

original data and can be summarized as follows. Two demes (western and eastern gorillas) of a particular deme size are represented by the diploid number of chromosomes in a population (2 times N_e). The sample size is the number of sampled genes. A stationary growth rate is assumed and 3 migration matrices are described as follows: first matrix = 0 (no migration for a given period), second matrix = m ($m = M/2N$ ($M = [1 - F_{ST}]/[2F_{ST}]$; N = size of populations [N = 20,000]), third matrix = 0 (fused population – no migration). The number of polymorphic loci was in this case 12 and could be specified along with their respective lengths, mutation rates, and proportions of transitions. The parameters of interest were predefined as historical events: the coalescence time (historical event Y) and the actual population split (historical event X). We first set the time of historical event Y to 80,000 generations ago and systematically decreased the time of the historical event X in every new run (70,000; 65,000 … 5,000). We repeated this procedure defining the time for event Y to 70,000, 60,000, and 50,000 generations ago and adjusting the times for event X accordingly. During each of those runs, 1,000 independent data sets were generated (following Laval G, personal communication), from which we obtained π and F_{ST} values using Arlequin v3.0 (Schneider et al. 2000). The respective median values from each run were plotted in a regression analysis. Finally, we calculated the timings of the 2 historical events given π and F_{ST} values observed in the original data and retained only the time combinations that showed an overlap for both values, π and F_{ST}. Lastly, we used the simulation program IM (Nielsen and Wakeley 2001; Hey and Nielsen 2004) to investigate the isolation with migration model (scenario 3). By applying coalescence simulations and Bayesian computation procedures, IM yields 6 demographic parameters (the population-split time, effective population size for the ancestral and 2 current populations, and migration rates). The posterior probability densities of these parameters are generated by Markov chain Monte Carlo (MCMC) simulations, and 3 of these simulations were run with individual simulations being updated 50 million times. Within each simulation, we used a procedure to swap among 10 heated chains (Metropolis coupling) and observed the sufficient swapping rate while the simulation was running. Each simulation yielded a marginal density histogram for the 6 parameters of interest, and the peaks of the resulting distributions were considered as the maximum likelihood estimates (MLE) of the parameter with credibility intervals equaling the 90% highest posterior density (HPD) intervals. Because the parameter estimates are scaled by the neutral mutation rate (see also Won and Hey 2004), we needed to transform these values into basic demographic parameters by applying a generation time of 15 years for gorillas and the geometric mean of the locus-specific mutation rates. The settings for the prior distributions were identical among the 3 simulations, so we present the average values of the basic demographic parameter values.

Results

Diversity Estimates and Summary Statistics

We successfully amplified 16 noncoding autosomal loci of 800–1,200 bp in length from DNAs of 15 western gorillas and 1 eastern gorilla (GenBank accession numbers AM407431–AM407714) from blood and liver samples. In total, we obtained (excluding indels) 14,017 bp of sequence information for each of the 16 individuals. In addition, we succeeded in amplifying 14 of the 16 loci using DNA from 2 fecal samples (both from eastern gorillas). We hereafter distinguish the results including information from the fecal DNA samples by highlighting it in italic font (see table 1).

Among all gorillas, we identified a total of 79 (*73*) polymorphic sites of which 12 (*12*) were singletons, resulting in diversity estimates across all loci of π: 0.153% (*0.144%*) and θ_W: 0.147% (*0.153%*). When we excluded the eastern gorillas, the diversity estimates for western gorillas decreased slightly (π: 0.151% and θ_W: 0.139%). The summary statistics listed for each population separately can be found in the Supplementary Material online (supplementary table 1). The diversity values varied considerably among the loci (π: 0%–0.45%, table 1), and we attributed this variation to differences in mutation rates across autosomal loci and stochasticity. None of the polymorphic sites deviated significantly from the Hardy–Weinberg equilibrium. When we tested each locus for neutrality, none showed a significant deviation from neutrality consistently over all 3 applied tests. Hence, we concluded that the analyzed loci are indeed neutrally evolving in gorillas.

For each locus, we calculated the average number of nucleotide substitutions per site between human and gorillas, and using a divergence time of 8 Myr (Glazko and Nei 2003), we estimated the average mutation rate over all loci (μ) within gorillas as 1.44×10^{-8} nucleotide changes per site per generation. We used this mutation rate to estimate the diploid effective population size ($\theta_W = 4N_e\mu$) and obtained an estimate of *26,600* for all gorillas. We also calculated the effective population size for western and eastern gorillas separately as 24,100 for western and *13,600* for eastern gorillas.

Demographic Inferences

The demographic history of gorillas was explored using the multilocus approach in the coalescence simulation program HKA. Whereas the mean Tajima's D value over all loci was positive but not significantly different from the values obtained from the simulations, the observed positive mean Fu & Li's D value deviated significantly from the mean value of the simulations (0.502, P = 0.0231 [*0.500, P = 0.0221*]). This result of a significant positive mean Fu & Li's D value indicates a significant deficit of singletons among all gorillas. We also observed a significantly positive Fu & Li's D when we excluded the eastern gorillas and analyzed the western gorillas exclusively (0.444, P = 0.0342). Considering the fact that the autosomal loci included in this study are putatively noncoding and are apparently evolving neutrally, the significantly positive Fu & Li's D in western gorillas might be explained by 2 nonexclusive demographic scenarios. These consist of population substructure within western gorilla populations and/or a very recent or ongoing population size reduction in gorillas. The latter scenario is difficult to test with slowly evolving autosomal markers, and so we looked for signs of a potential correlation between the patterns of genetic

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

diversity and the broadly known geographic origins of the gorillas. However, a phylogenetic analysis using inferred haplotypes for each of the 15 western gorillas and 16 loci did not reveal statistically well-supported branching patterns that can be attributed to spatial population structure (data not shown). We then calculated the pairwise genetic divergence among all 18 gorillas using the mean values estimated from a chromosomal comparison of the common 14 loci. We used the average values for each pairwise comparison over all loci to transform these values into an Eucleadian distance matrix and subsequently used this matrix in a MDS analysis. The MDS plot (fig. 3) reveals a graphical separation of eastern gorillas from western gorillas and also a separation among western gorillas, indicative of structure in the data. However, western gorillas found in the graph in close proximity to each other are not in geographic proximity. Some closely plotted pairs of individuals are separated by several hundred kilometers (Nya–Arn; Jum–Pit), whereas individuals from the same approximate locality are widely spaced in the MDS plot (Arn–Ant; Jum–Ram). This pattern is also evident when performing the analysis in 3 dimensions. This suggests that within gorillas, except for the separation of the eastern from western gorillas, no apparent geographic structure can be detected in our data and that the signal detected with Fu & Li's D test might be a result of population structure prevalent in an ancestral gorilla population, rather than structure linked to current geography.

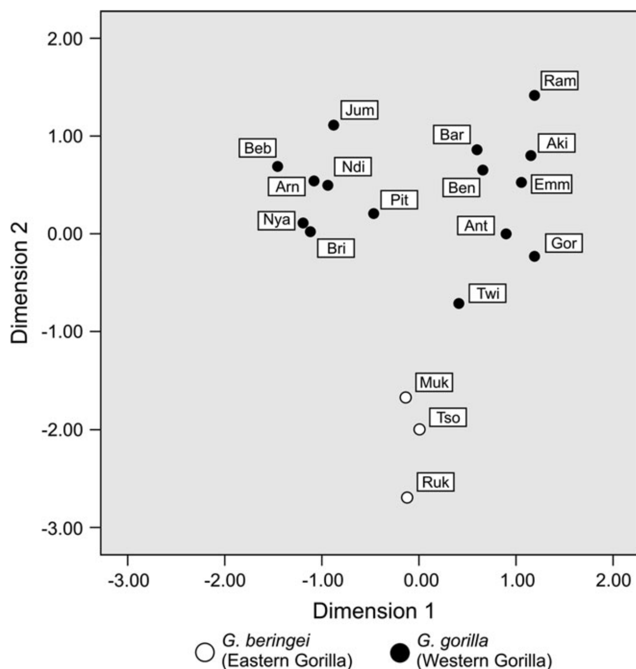

FIG. 3.—Graphical representation of the MDS, which shows the transformed pairwise genetic divergences among individuals in a 2-dimensional space. The graph shows the results of an analysis performed using all 18 gorillas sequenced on 14 nuclear loci. The individuals are coded with respect to their names (first 3 letters). The observed stress value was 0.302 and $R^2 = 0.67$. In an analysis including 3 dimensions, the stress value was reduced to 0.18, but the relationships of the individuals remain the same.

Divergence Assessment between Western and Eastern Gorillas

The MDS plot (fig. 3) also suggests that some western gorillas are about as distant from each other as western gorillas are from eastern gorillas. To investigate the divergence between western and eastern gorillas in more depth, we used the data set including all 18 gorillas and 14 loci. We calculated the frequencies of all pairwise genetic divergences among all individuals (average divergence = 0.14%). A comparison of the frequencies obtained within western gorillas with those obtained between western and eastern gorillas, respectively, showed 2 partially overlapping distributions (supplementary fig. 1, Supplementary Material online). Although the mean values of these 2 overlapping frequency distributions were significantly different (2-tailed *t*-test: -13.082, $P < 0.0001$), it does suggest that 2 randomly chosen western gorillas are not much more genetically similar to each other than are any single eastern gorilla and any western gorilla. Because gene flow may reduce the number of polymorphisms fixed within groups and enhances the number of shared polymorphisms, we estimated those values for western and eastern gorillas and found that shared polymorphisms (7) outnumbered fixed polymorphisms (2). Nonetheless, the significant F_{ST} value of 0.38 and the geographic separation of the closest populations from the western and eastern range of gorillas cause us to reject the null model of complete panmixia between western and eastern gorillas. Using the program WH (Wakeley and Hey 1997), we investigated scenario 1 (fig. 2) and estimated the time of a population split without subsequent gene flow to have occurred around 132,000 years ago. However, the high variance on the estimated population-split time (132,000; standard deviation 65,384) together with the low probability (0.0642) of the fit of the data make the isolation model also quite unlikely. Furthermore, the rejection of this model gets additional support from the results of another simulation. Using SIMCOAL and data from all polymorphic loci (Laval and Excoffier 2004), we modeled a scenario similar to the isolation model by applying 150,000 years as a population-split time (similar to the value estimated with WH). The simulations produced a π value of 0.08% and an F_{ST} of 0.201. These values are quite different from the π (0.131%) and F_{ST} values obtained from the original data used in the simulation, and hence, we concluded that more complex scenarios are necessary to explain the observed data. SIMCOAL was then used to simulate a limited gene flow model (scenario 2, fig. 2) in which gene flow between western and eastern gorillas ceased X generations ago. Before this event, both populations shared migrants at a rate estimated from $m = M/2N = 0.0000187$. A more distant historical event (Y generations ago) describes the end of panmixia in an ancestral gorilla population. By using the equations obtained from the regressions (minimum $R^2 = 0.86$) and applying our observed values of π for western gorillas and F_{ST} with $\pm 5\%$ tolerance intervals, we estimated that panmixia existed until 60,000–70,000 generations ago and that reduced gene flow between western and eastern gorillas persisted until 10,900–15,400 generations ago. Assuming a generation time of 15 years, these

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

estimates correspond to an initial split of 0.9–1.05 MYA and gene flow until ~164,000–230,000 years ago. By using the program IM, we were able to refine 6 demographic parameters estimated in MCMC simulations and test for the isolation with migration model (scenario 3 in fig. 2). The marginal posterior density distributions for these parameters are depicted in figure 4. We present the average values of the MLE and 90% HPD intervals based on 3 independent simulations, but results for each of the 3 individual simulations are also shown in table 2. The average values of MLE of the effective population sizes of western, eastern, and their ancestral population were 17,700, 2,900, and 42,000, respectively. The MLE values for the ancestral population showed broad 90% HPD intervals, and also its posterior distribution appeared to be tailed and nonzero at the upper limit of the parameter space. Interestingly, the average distribution of the population-split time shows a bimodal curve (fig. 4), a pattern that was persistent in each of the 3 individual simulations. Although the first and also highest MLE peak occurred at a split time of 77,700 years, a second MLE peak corresponds to a population-split time of 1.6 MYA. The simulations also suggested ongoing gene flow between western and eastern gorillas subsequent to the population split. Specifically, both marginal posterior distributions of the migration parameter have a peak different from zero, indicating that gene flow has occurred in both directions. When the MLE for the migration parameters are transformed into the population migration rates; $M = 2N_1m_1 = m_1*(\theta_1/2)$, eastern to western gorillas (0.35), and $M = 2N_2m_2 = m_2*(\theta_2/2)$, western to eastern gorillas (0.141), it became apparent that more than twice as much gene flow was found from eastern to western gorillas, than vice versa. By examining each locus individually, we performed a more in-depth analysis of the posterior distributions for the numbers and times of migration events and found evidence at each locus for migration in both directions (supplementary table 2, Supplementary Material online). The majority of loci (7 out of 12) showed more migration events into the western gorillas, whereas for only 3 out of 12 loci were more migrations into eastern gorillas observed. The average number of migration events over all loci was 3.2 in the direction of western gorillas (range: 1–5) and 2.5 migration events into eastern gorillas (range: 1–4). Considering the times of migration events, the average values over all loci were 316,000 (151,000–727,000) and 359,000 (151,000–590,000) years ago for gene flow into western and eastern gorillas, respectively. These estimates are older than the MLE value for the population-split time of western and eastern gorillas estimated with IM but are in very good agreement with the results obtained with SIMCOAL.

Discussion

The Utility of Noninvasively Collected Samples

In this study, we had access to records that allowed us to infer at least the broad geographical location of most Cameroonian gorillas (fig. 1). In most cases, blood samples are unavailable from free-ranging animals and are obtained from captive animals, whose geographic origins are usually only crudely known (Fischer et al. 2004; Yu et al. 2004). However, in order to draw a comprehensive picture of genetic diversity in any species, knowledge of the exact geographical origin of the sampled individuals is necessary. A solution to this dilemma is the application of noninvasively collected samples from wild animals. Reasons for using these types of samples include ethical as well as practical considerations (Constable et al. 1995; Taberlet et al. 1996; Kohn and Wayne 1997; Seddon et al. 2005). Although some precautions need to be undertaken when working with noninvasively collected samples (Morin et al. 2001; Fernando et al. 2003; see also Knapp 2005; Lukas and Vigilant 2005), improvements in technical methods have made noninvasively collected samples ideal for a variety of population genetic studies (Bellemain and Taberlet 2004; Nsubuga et al. 2004; Bradley et al. 2005; Prugh et al. 2005; Smith and Morin 2005). Because, in previous studies, the amplified segments rarely exceeded 500 bp in length, the utility of noninvasively collected samples for sequencing projects of genomic DNA is rather unappreciated. In this study, successful amplifications of several nuclear loci ranging between 800 and 1,200 bp in length from 2 noninvasively collected samples enabled us to extend our analysis to all currently recognized gorilla subspecies and demonstrates the utility of such samples for studies of nuclear DNA segments.

Patterns of Genetic Diversity within Gorillas

We are aware that our 3 eastern gorillas do not by far provide a comprehensive picture of diversity within this taxa, and hence, we encourage further genetic investigations, including additional noninvasively collected samples of known geographic origin. This is of particular interest in light of recent multilocus sequencing analyses of chimpanzees, which showed 1.2-fold to 1.6-fold differences in diversity levels between populations within a single species (Yu et al. 2003; Fischer et al. 2004, 2006) and analyses on worldwide human populations revealing similar patterns (Yu et al. 2002; Voight et al. 2005).

Our results confirm that gorillas have 2-fold higher diversity than humans and a diversity comparable to what has been reported within the genus *Pan* (Kaessmann et al. 2001; Yu et al. 2003, 2004; Fischer et al. 2004). Under the assumption that populations have diverged rather recently ($t \leq 2N_e$ generations), sequences sampled in one population are often more similar to sequences from another population than to sequences sampled in the same population (Tajima 1983). Thus, given this assumption, although our sampling of *G. gorilla* is almost exclusively composed of Cameroonian gorillas (fig. 1), we may have detected most of the nuclear genetic diversity present among extant western gorillas. This suggests that additional samples from populations in close proximity to our sampling area such as Gabon, Equatorial Guinea, or Democratic Republic of Congo would only marginally change the diversity indices we have observed in western gorillas.

Complex Demographic History of Western Gorillas

Previous studies have shown that summary statistics allow inference of the demographic history of species from

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

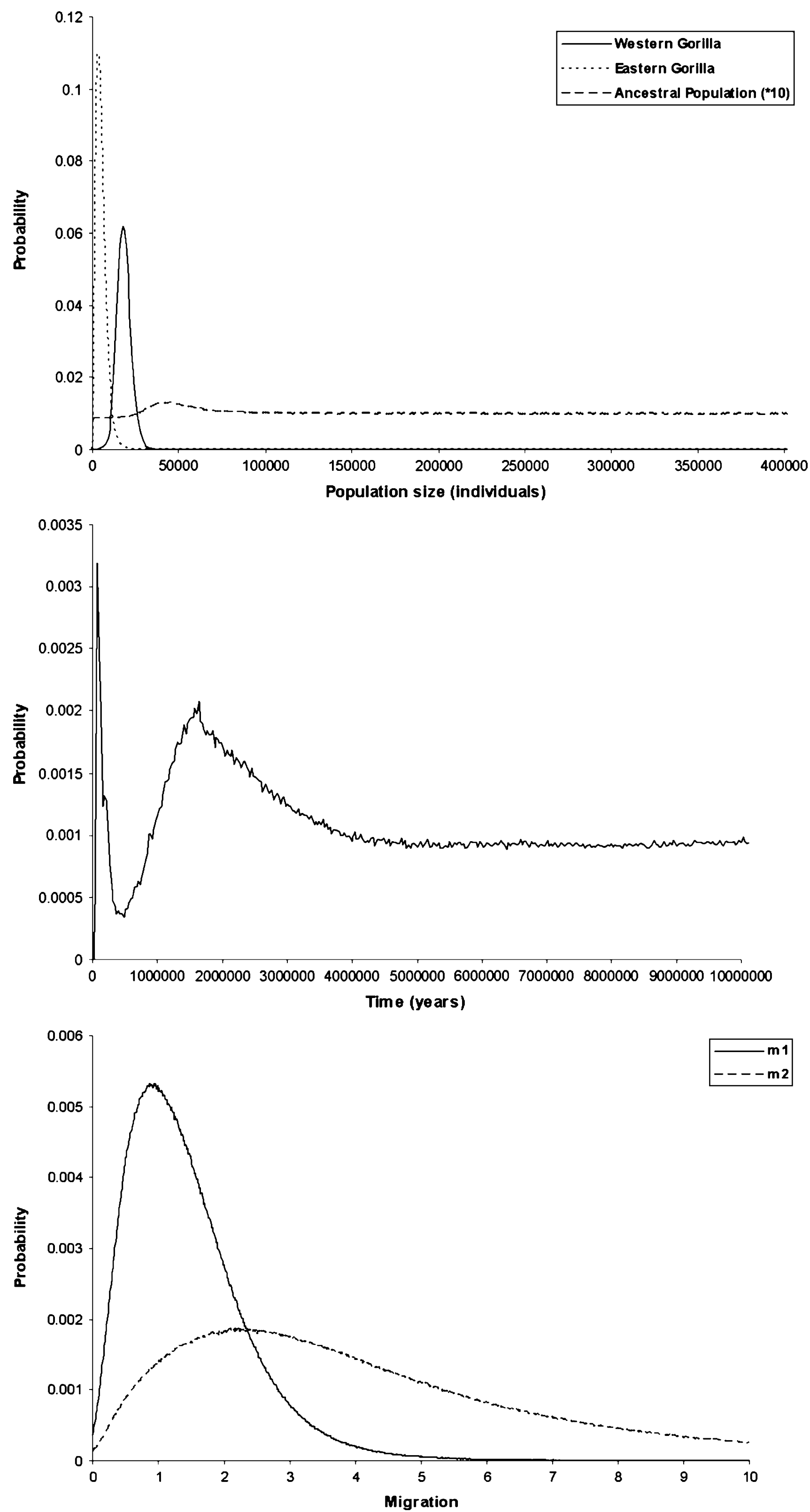

FIG. 4.—Marginal posterior probability distributions for the 6 parameters modeled in 3 independent IM simulations. All input values are average values over the 3 simulations. Whereas for population size and time the values are already transformed into basic demographic units consisting of the number of individuals and years, respectively, the migration graph depicts the actual parameter values estimated with IM.

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

Table 2
MLE and the 90% HPD Intervals of IM Parameters

		θ_1	θ_2	$\theta_{Ancestral}$	m_1	m_2	t	N_{e1}	N_{e2}	$N_{e\ Ancestral}$	$2N_1m_1$	$2N_2m_2$	t (years)
Simulation I	MLE	0.767	0.133	1.817	0.885	2.175	0.07	17,500	3,000	41,600	0.339	0.145	96,000
	Lower 90% HPD	0.417	0.051	1.091[a]	0.125	0.275[a]	0.05[a]	9,500	1,200	24,900[a]			68,600[a]
	Upper 90% HPD	1.117	0.358	26.911[a]	2.655	7.505[a]	19.99[a]	25,500	8,200	615,300[a]			27,424,400[a]
Simulation II	MLE	0.794	0.133	1.79	0.955	2.275	0.05	18,200	3,000	40,900	0.379	0.151	68,600
	Lower 90% HPD	0.444	0.051	1.144[a]	0.135	0.265	0.03[a]	10,200	1,200	26,200[a]			41,200[a]
	Upper 90% HPD	1.144	0.358	26.911[a]	2.675	7.465	19.99[a]	26,200	8,200	615,300[a]			27,424,400[a]
Simulation III	MLE	0.767	0.112	1.898	0.865	2.295	0.05	17,500	2,600	43,400	0.332	0.129	68,600
	Lower 90% HPD	0.364	0.051	0.983[a]	0.125	0.275	0.03[a]	8,300	1,200	22,500[a]			41,200[a]
	Upper 90% HPD	1.225	0.439	26.911[a]	2.685	7.535	19.99[a]	28,000	10,000	615,300[a]			27,424,400[a]
Average values		0.776	0.126	1.835	0.902	2.248	0.057	17,700	2,900	42,000	0.350	0.141	77,700

NOTE.—Results were obtained from 3 independent simulations using 15 *Gorilla gorilla* representing population 1 and 3 *Gorilla beringei* individuals defined as population 2, running for 50 million updates each.

[a] Some of the posterior density distributions that appear to be not complete, meaning nonzero at the upper limit and thus having a tail. For further details see the text.

sequence data obtained from multiple autosomal loci (Pluzhnikov et al. 2002; Broughton and Harrison 2003; Fischer et al. 2006). In this study, we obtained a significantly positive Fu & Li's D value (Fu and Li 1993). One reasonable explanation for such a pattern is population structure among the sampled individuals. This would be a surprising result though, given the fact that our sampling consists of a rather region-specific representation of western gorillas. Furthermore, our results suggest that the pattern of genetic variation detected in western gorillas is not correlated with the current geographical distribution of the sampled gorillas. This finding leads us, in concordance with Yu et al. (2004), to invoke the presence of population structure in an ancestral western gorilla population. It has been shown (e.g., Tajima 1989a) that the number of segregating sites increases in a structured population as compared with a randomly mating population as does the effective population size (but for detailed discussion see also Whitlock and Barton 1997). The N_e value observed in western gorillas (N_e = 24,100) is within the range of other forest-living African hominoids such as chimpanzees: 18,500–47,000 and bonobos: 12,300–45,600 (Kaessmann, Wiebe, et al. 1999; Yu et al. 2003; Fischer et al. 2004, 2006) but exceeds that of humans (Kaessmann et al. 1999; Yu et al. 2002) by a factor of 2.5. We could envision an evolutionary scenario in which an ancestral gorilla population was structured, but climatic changes during the Quaternary led to subsequent intermixing of those demes. Similar scenarios have previously been proposed for a variety of species (reviewed in Hewitt 2000; Taberlet and Cheddadi 2002). Moreover, Hofreiter et al. (2004) used ancient DNA analyses to suggest population histories characterized by expansion and contraction phases for several extinct European species living in the late Pleistocene, including Neanderthals. During the last 2 Myr, glacial and interglacial oscillations have severely affected the global climate (deMenocal 1995; Petit et al. 1999 and references therein). In Africa, arid (glacial) periods were characterized by a contraction of rain forest into fragmentary distributed forest patches (Dupont et al. 2000, 2001), resulting in a retreat of forest-living species (Haffer 1982; Mayr and O'Hara 1986) such as western gorillas. Postglacial forest recovery allowed surviving individuals to spread and intermix with genetically divergent lineages. Such a process ultimately results in the maintenance of relatively high levels of genetic diversity and also the preservation of rather "old" mutations within a population. An alternative explanation for a significantly positive Fu & Li's D value includes a recent reduction in population size as might have occurred in great apes due to anthropogenic pressure, but the detection of such an event would require rapidly evolving genetic markers (Goossens et al. 2006). Although gorillas are under threat of extinction and possibly in decline for the last few decades (Walsh et al. 2003), this is too recent to be the explanation of the pattern seen in our data.

Population Split of Western and Eastern Gorillas

Climatic oscillations might also have facilitated interspecific split events. A concurrent population split between western and eastern gorillas and between chimpanzees and bonobos has been suggested (Ruvolo et al. 1994) and hypothesized to be associated with a common biogeographic event (Jensen-Seaman 2000). Whereas the genetic separation of chimpanzees and bonobos has been maintained by a natural barrier, the Congo River (Yu et al. 2003; Eriksson et al. 2004), such a barrier did not exist between western and eastern gorillas. Population intermixing expedited by climatic and ecological changes could have happened until very recently. There is ample evidence from museum collections to show that a few decades ago, gorillas lived in regions from which they are now extinct (see also Hofreiter et al. 2003). Using mitochondrial data, we estimated a coalescence time of about 1.3 Myr between western and eastern gorillas (Thalmann et al. 2005). After testing different models, the data presented here best fit a scenario in which panmixia in an ancestral gorilla population stopped around 0.9–1.6 MYA, a timing which might have coincided with the formation of the Great Rift Valley (Beadle 1981). Such a large geological change could also have had an impact on the divergence of other species such as chimpanzees and bonobos. Furthermore, under the assumption of a limited

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

gene flow model, we can estimate the period of ongoing reduced gene flow between western and eastern gorillas as lasting until a population split ~160,000–230,000 years ago. It is possible, however, that environmental changes in Africa during the Pleistocene allowed more recent gene flow, as suggested by the most recent estimation for population-split time (77,700 years ago), calculated with IM (Nielsen and Wakeley 2001). Using this simulation program, we discerned patterns of gene flow persisting until about 150,000 years ago among the 2 gorilla populations. Although gene flow occurred in both directions, our results suggest that the majority was from eastern toward western gorilla populations. Under the assumption of a metapopulation model, one would expect more genetic influx into peripheral populations such as the eastern gorilla populations from central populations rather than genetic efflux out of eastern gorilla populations; however, population expansion and contraction in correspondence to climatic changes might have facilitated gene flow from the periphery into the center (Hofreiter et al. 2004). Furthermore, gene flow from smaller, peripheral groups into a bigger more central group has been reported for a small remnant population of western gorillas, the Cross River gorillas (*G. gorilla diehli*) (Bergl and Vigilant Forthcoming).

A discrepancy was observed between the simulated N_e values and N_e values calculated using θ_W (Watterson 1975). The effective population sizes for both western and eastern gorillas were higher in the latter calculations (17,700 vs. 24,100 and 2,900 vs. 13,600, respectively), a pattern that was also found when IM simulations were applied to populations of chimpanzees and bonobos (Won and Hey 2004). The authors attributed this to be an effect of the greater variance of N_e estimates based on diversity indices (Won and Hey 2004, p. 303). Another factor, low sample size, might influence values obtained for eastern gorillas. This is a major drawback for running such complex simulations, potentially also causing the high 90% HPD values for the N_e estimates for the ancestral population and split time as well as the bimodality in the posterior distribution of the split time parameter. Thus, it remains unclear how much this has affected the parameter estimates, advocating additional analysis using a larger sampling of eastern gorillas. However, several lines of evidence suggest a very complex population-split scenario, with a close relationship of western and eastern gorillas persisting until the very recent past, similar to what has been recently proposed for the split of humans and chimpanzees (Patterson et al. 2006).

It is difficult to determine whether currently disjunct populations represent different species or subspecies as these terms resist useful biological definition (e.g., Hey et al. 2003). As previously proposed by Groves (2001; see also Grubb et al. 2003), the western and eastern gorillas are currently recognized as 2 distinct species. Considering the recent common history of both taxa and their high genetic similarity, it could be suggested that return to a one-species classification of gorillas was warranted. In fact, the level of divergence in gorillas is comparable to that found among chimpanzee populations (Fischer et al. 2006).

Nuclear DNA sequences from western and eastern gorillas are not reciprocally monophyletic (data here and in Jensen-Seaman et al. 2003), in contrast to mtDNA sequences from these 2 groups (Ruvolo et al. 1994; Thalmann et al. 2004). This is most readily explained by the expected differences in genealogical depth of autosomal and mitochondrial loci. Because the probability of coalescence increases with decreasing effective population size (Kingman 1982), mtDNA, which has a 4-fold lower effective population size than autosomal DNA, coalesces much faster (Moore 1995; Palumbi et al. 2001). However, the discrepancy among nuclear and mitochondrial phylogenies could at least be partly a result of male-mediated gene flow between western and eastern gorillas (Jensen-Seaman et al. 2001), a scenario that has previously been invoked for other species (Tosi et al. 2000; Pidancier et al. 2006). Such a hypothesis gets additional support from some limited analyses of Y-chromosome variation, showing a single shared haplotype in western and eastern gorillas (Burrows and Ryder 1997; Altheide 2002). Although both male and female gorillas disperse (Doran and McNeilage 1998; Stokes et al. 2003; Watts 2003), field observations suggest that male gorillas migrate farther than females (Yamagiwa 1986; Yamagiwa and Mwanza 1994). Although a recent study by Bradley et al. (2004) suggested that on a local scale, male western gorillas form "dispersed male networks" in which related males lead nearby groups, little is known about the patterns of lone male gorilla dispersal. A scenario similar to that in other social animals in which males occasionally transfer their genes over longer distances while dispersing (Fabiani et al. 2003; Mech and Boitani 2003; Robertson et al. 2006) would help to explain why identical alleles at certain nuclear loci are found in western and eastern gorillas.

In conclusion, we propose, based on genetic evidence, a demographic scenario for western gorilla evolution with ancestral population substructure and phases of intermixing driven by climatic oscillations in Africa during the Quaternary. A potential population-split scenario between western and eastern gorillas includes the existence of a single ancestral gorilla population until 0.9–1.6 MYA followed by reduced, possibly male-mediated gene flow until a population split at around 77,700 years ago. Refined simulations suggest that the gene flow was predominantly from eastern to western gorilla populations.

Supplementary Material

Supplementary figure 1 and tables 1 and 2 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are grateful to the 3 reviewers and especially the Associate Editor for their detailed comments and suggestions. We thank the following organizations for their cooperation: LWC, Cameroon; Le Ministère de l'Environnement et des Forêts du Cameroun; Wildlife Conservation Society; Uganda Wildlife Authority; Uganda National Council of Science and Technology; Institute of Tropical Forest Conservation, Uganda; Institut Congolais pour la Conservation de la Nature; and Dian Fossey Gorilla Fund. We thank M. Robbins, A. Nsubuga, C. Kusamba Zacharie, H. Cirhuza, V. Smith, J. L. Sunderland-Groves, and Dr K. Leus

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

(Center for Research and Conservation of the Royal Zoological Society of Antwerp) for providing samples and assistance. We thank A. Abraham, A. Burkhard, B. Hoeffner, E. Kreuz, R. Schultz, H. Siedel, A. Weihmann, and B. Woltering for laboratory assistance and computational support. We appreciate fruitful discussions and helpful comments on the manuscript from M. Arandjelovic, C. Boesch, J. Eriksson, M. Hofreiter, D. Lukas, and D. Morgan. For her intensive help, we are very obliged to S. Ptak. We are grateful to J. Hey and G. Laval for their helpful advice with the programs IM and SIMCOAL v2, respectively, and we also thank D. Stahl for statistical advice. This work was financially supported by the Max Planck Society and the Deutsche Forschungsgemeinschaft (VI 229/2-1).

Literature Cited

Altheide TK. 2002. Comparative population genetics of the Hominoidea: an investigation of locus-specific and genome-wide influence [dissertation]. Department of Ecology and Evolutionary Biology, University of Arizona.

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

Beadle LC. 1981. The inland waters of tropical Africa. London: Longman.

Bellemain E, Taberlet P. 2004. Improved noninvasive genotyping method: application to brown bear (*Ursus arctos*) faeces. Mol Ecol Notes. 4:519–522.

Bergl RA, Vigilant L. Forthcoming. Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*). Mol Ecol.

Bradley BJ, Doran-Sheehy DM, Lukas D, Boesch C, Vigilant L. 2004. Dispersed male networks in western gorillas. Curr Biol. 14:510.

Bradley BJ, Robbins MM, Williamson EA, Steklis HD, Steklis NG, Eckhardt N, Boesch C, Vigilant L. 2005. Mountain gorilla tug-of-war: silverbacks have limited control over reproduction in multimale groups. Proc Natl Acad Sci USA. 102: 9418–9423.

Broughton RE, Harrison RG. 2003. Nuclear gene genealogies reveal historical, demographic and selective factors associated with speciation in field crickets. Genetics. 163: 1389–1401.

Brumfield RT, Beerli P, Nickerson DA, Edwards SV. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol Evol. 18:249–256.

Burrows W, Ryder OA. 1997. Y-chromosome variation in great apes. Nature. 385:125–126.

Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. Nature. 325:31–36.

Chen F-C, Li W-H. 2001. Genomic divergence between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. Am J Hum Genet. 68:444–456.

Clifford SL, Anthony NM, Bawe-Johnson M, et al. 2004. Mitochondrial DNA phylogeography of western lowland gorillas (*Gorilla gorilla gorilla*). Mol Ecol. 13:1551–1565.

Constable JJ, Packer C, Collins DA, Pusey AE. 1995. Nuclear DNA from primate dung. Nature. 373:393.

Constable JL, Ashley MV, Goodall J, Pusey AE. 2001. Noninvasive paternity assignment in Gombe chimpanzees. Mol Ecol. 10:1279–1300.

Deinard A, Kidd K. 1999. Evolution of a HOXB6 intergenic region within the great apes and humans. J Hum Evol. 36: 687–703.

deMenocal PB. 1995. Plio-Pleistocene African climate. Science. 270:53–59.

Doran DM, McNeilage A. 1998. Gorilla ecology and behavior. Evol Anthropol. 6:120–131.

Dupont LM, Donner B, Schneider R, Wefer G. 2001. Mid-Pleistocene environmental change in tropical Africa began as early as 1.05 Ma. Geology. 29:195–198.

Dupont LM, Jahns S, Marret F, Ning S. 2000. Vegetation change in equatorial West Africa: time-slices for the last 150 ka. Palaeogeogr Palaeocl. 155:95–122.

Edwards SV, Beerli P. 2000. Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution. 54:1839–1854.

Eriksson J, Hohmann G, Boesch C, Vigilant L. 2004. Rivers influence the population genetic structure of bonobos (*Pan paniscus*). Mol Ecol. 13:3425–3435.

Fabiani A, Hoelzel AR, Galimberti F, Muelbert MMC. 2003. Long-range paternal gene flow in the southern elephant seal. Science. 299:676.

Fernando P, Vidya TNC, Rajapakse C, Dangolla A, Melnick DJ. 2003. Reliable noninvasive genotyping: fantasy or reality? J Hered. 94:115–123.

Finlayson C. 2005. Biogeography and evolution of the genus *Homo*. Trends Ecol Evol. 20:457.

Fischer A, Pollack J, Thalmann O, Nickel B, Pääbo S. 2006. Demographic history and genetic differentiation among apes. Curr Biol. 16:1133–1138.

Fischer A, Wiebe V, Pääbo S, Przeworski M. 2004. Evidence for a complex demographic history of chimpanzees. Mol Biol Evol. 21:799–808.

Frisse L, Hudson RR, Bartoszewicz A, Wall JD, Donfack J, Di Rienzo A. 2001. Gene conversion and different population histories may explain the contrast between polymorphism and linkage disequilibrium levels. Am J Hum Genet. 69:831–843.

Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. Genetics. 133:693–709.

Gagneux P. 2002. The genus *Pan*: population genetics of an endangered outgroup. Trends Genet. 18:327–330.

Gagneux P, Wills C, Gerloff U, Tautz D, Morin PA, Boesch C, Fruth B, Hohmann G, Ryder OA, Woodruff DS. 1999. Mitochondrial sequences show diverse evolutionary histories of African hominoids. Proc Natl Acad Sci USA. 96:5077–5082.

Garner KJ, Ryder OA. 1996. Mitochondrial DNA diversity in gorillas. Mol Phylogenet Evol. 6:39–48.

Glazko GV, Nei M. 2003. Estimation of divergence times for major lineages of primate species. Mol Biol Evol. 20:424–434.

Goldberg TL, Ruvolo M. 1997. The geographic apportionment of mitochondrial genetic diversity in east African chimpanzees, *Pan troglodytes schweinfurthii*. Mol Biol Evol. 14: 976–984.

Goossens B, Chikhi L, Ancrenaz M, Lackman-Ancrenaz I, Andau P, Bruford MW. 2006. Genetic signature of anthropogenic population collapse in orang-utans. PLoS Biol. 4.

Groves CP. 2001. Primate taxonomy. Washington: Smithsonian Institution Press.

Grubb P, Butynski T, Oates J, Beader S, Disotell T, Groves C, Struhsaker T. 2003. Assessment of the diversity in African primates. Int J Primatol. 24:1301–1357.

Haffer J. 1982. General aspects of the refuge theory. In: Prance GT, editor. Biological diversification in the tropics. New York: Columbia University Press. p. 6–24.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 41:95–98.

Harding RM, McVean G. 2004. A structured ancestral population for the evolution of modern humans. Curr Opin Genet Dev. 14:667.

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

Hare MP. 2001. Prospects for nuclear gene phylogeography. Trends Ecol Evol. 16:700–706.
Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. Nature. 405:907–913.
Hey J. 1998. Population genetics and human origins-haplotypes are key! Trends Genet. 14:303–304.
Hey J, Nielsen R. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics. 167:747–760.
Hey J, Waples RS, Arnold ML, Butlin RK, Harrison RG. 2003. Understanding and confronting species uncertainty in biology and conservation. Trends Ecol Evol. 18:597–603.
Hofreiter M, Serre D, Rohland N, Rabeder G, Nagel D, Conard N, Munzel S, Pääbo S. 2004. Lack of phylogeography in European mammals before the last glaciation. Proc Natl Acad Sci USA. 101:12963–12968.
Hofreiter M, Siedel H, Vigilant L. 2003. Mitochondrial DNA sequence from an enigmatic gorilla population (*Gorilla gorilla uellensis*). Am J Phys Anthropol. 121:361–368.
Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics. 116:153–159.
International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. Nature. 431:931.
Jensen-Seaman M. 2000. Evolutionary genetics of gorillas [dissertation]. Yale University.
Jensen-Seaman M, Kidd K. 2001. Mitochondrial DNA variation and biogeography of eastern gorillas. Mol Ecol. 10:2241–2247.
Jensen-Seaman MI, Deinard AS, Kidd KK. 2003. Mitochondrial and nuclear DNA estimates of divergence between western and eastern gorillas. In: Taylor AB, Goldsmith ML, editors. Gorilla biology: a multidisciplinary perspective. Cambridge (MA): Cambridge University Press. p. 247–268.
Jensen-Seaman MI, Sarmiento EE, Deinard AS, Kidd KK. 2004. Nuclear integrations of mitochondrial DNA in gorillas. Am J Primatol. 63:139–147.
Jorde LB, Bamshad MJ, Watkins WS, Zenger R, Fraley AE, Krakowiak PA, Carpenter KD, Soodyall H, Jenkins T, Rogers AR. 1995. Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. Am J Hum Genet. 57:523–538.
Kaessmann H, Heissig F, von Haeseler A, Pääbo S. 1999. DNA sequence variation in a non-coding region of low recombination on the human X chromosome. Nat Genet. 22:78.
Kaessmann H, Wiebe V, Pääbo S. 1999. Extensive nuclear DNA sequence diversity among chimpanzees. Science. 286: 1159–1162.
Kaessmann H, Wiebe V, Weiss G, Pääbo S. 2001. Great ape DNA sequences reveal a reduced diversity and an expansion in humans. Nat Genet. 27:155–156.
Kilger C, Krings M, Poinar H, Pääbo S. 1997. "Colony sequencing": direct sequencing of plasmid DNA from bacterial colonies. Biotechniques. 22:412–418.
Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 16:111–120.
Kingman JFC. 1982. On the genealogy of large populations. J Appl Prob. 19:27–43.
Knapp LA. 2005. Facts, faeces and setting standards for the study of MHC genes using noninvasive samples. Mol Ecol. 14:1597–1599.
Kohn MH, Wayne RK. 1997. Facts from feces revisited. Trends Ecol Evol. 12:223–227.
Kumar S, Tamura K, Jakobsen IB, Nei M. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics. 17:1244–1245.
Laval G, Excoffier L. 2004. SIMCOAL 2.0: a program to simulate genomic diversity over large recombining regions in a subdivided population with a complex history. Bioinformatics. 20:2485–2487.
Lercher MJ, Hurst LD. 2002. Human SNP variability and mutation rate are higher in regions of high recombination. Trends Genet. 18:337–340.
Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ. 1994. Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J Mol Evol. 39:174–190.
Lukas D, Vigilant L. 2005. Reply: facts, faeces and setting standards for the study of MHC genes using noninvasive samples. Mol Ecol. 14:1601–1602.
Machado CA, Kliman RM, Markert JA, Hey J. 2002. Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. Mol Biol Evol. 19:472–488.
Mayr E, O'Hara RJ. 1986. The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. Evolution. 40:55–67.
Mech LD, Boitani L. 2003. Wolves: behavior, ecology, and conservation. Chicago (IL): The University of Chicago Press.
Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution. 49:718–726.
Morin PA, Chambers KE, Boesch C, Vigilant L. 2001. Quantitative PCR analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). Mol Ecol. 10:1835–1844.
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.
Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.
Nichols R. 2001. Gene trees and species trees are not the same. Trends Ecol Evol. 16:358–364.
Nielsen R, Hey J. 2001. Distinguishing migration from isolation. A Markov chain Monte Carlo approach. Genetics. 158: 885–896.
Noda R, Kim CG, Takenaka O, Ferrell RE, Tanoue T, Hayasaka I, Ueda S, Ishida T, Saitou N. 2001. Mitochondrial 16S rRNA sequence diversity of Hominoids. J Hered. 92:490–496.
Nsubuga AM, Robbins MM, Roeder AD, Morin PA, Boesch C, Vigilant L. 2004. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. Mol Ecol. 13:2089–2094.
Palumbi S, Cipriano F, Hare M. 2001. Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. Evolution. 55:859–868.
Patterson N, Richter DJ, Gnerre S, Lander ES, Reich D. 2006. Genetic evidence for complex speciation of humans and chimpanzees. Nature. 441:1103–1108.
Petit JR, Jouzel J, Raynaud D, et al. 1999. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature. 399:429–436.
Pidancier N, Jordan S, Luikart G, Taberlet P. 2006. Evolutionary history of the genus *Capra* (Mammalia, Artiodactyla): discordance between mitochondrial DNA and Y-chromosome phylogenies. Mol Phylogenet Evol. 40:739–749.
Pluzhnikov A, Di Rienzo A, Hudson RR. 2002. Inferences about human demography based on multilocus analyses of noncoding sequences. Genetics. 161:1209–1218.
Prugh LR, Ritland CE, Arthur SM, Krebs CJ. 2005. Monitoring coyote population dynamics by genotyping faeces. Mol Ecol. 14:1585–1596.
Robertson BC, Chilvers BL, Duignan PJ, Wilkinson IS, Gemmell NJ. 2006. Dispersal of breeding, adult male *Phocarctos hookeri*:

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016

implications for disease transmission, population management and species recovery. Biol Conserv. 127:227–236.
Rosenberg NA, Nordborg M. 2002. Genealogical trees, coalescence theory and the analysis of genetic polymorphisms. Nat Rev Genet. 3:380–390.
Rozas J, Rozas R. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics. 15:174–175.
Ruvolo M. 1997. Molecular phylogeny of the hominoids: inferences from multiple independent DNA sequence data sets. Mol Biol Evol. 14:248–265.
Ruvolo M, Disotell T, Allard M, Brown W, Honeycutt R. 1991. Resolution of the African hominoid trichotomy by use of a mitochondrial gene sequence. Proc Natl Acad Sci USA. 88: 1570–1574.
Ruvolo M, Pan D, Zehr S, Goldberg T, Disotell TR, von Dornum M. 1994. Gene trees and hominoid phylogeny. Proc Natl Acad Sci USA. 91:8900–8904.
Saltonstall K, Amato J, Powell J. 1998. Mitochondrial DNA variability in Grauer's gorillas of Kahuzi-Biega National Park. J Hered. 89:129–135.
Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. New York: Cold Spring Harbor Laboratory Press.
Schneider S, Roessli D, Excoffier L. 2000. Arlequin, version 2.0: a software for population genetics data analysis. Geneva (Switzerland): Department of Anthropology and Ecology, University of Geneva.
Seddon JM, Parker HG, Ostrander EA, Ellegren H. 2005. SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. Mol Ecol. 14:503–511.
Smith S, Morin PA. 2005. Optimal storage conditions for highly dilute DNA samples: a role for trehalose as a preserving agent. J Forensic Sci. 50:1–8.
Stephens M, Donelly P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet. 73:1162–1169.
Stephens M, Smith NJ, Donelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 68:978–989.
Stokes EJ, Parnell RJ, Olejniczak C. 2003. Female dispersal and reproductive success in wild western lowland gorillas (*Gorilla gorilla gorilla*). Behav Ecol Sociobiol. 54:329–339.
Taberlet P, Camarra JJ, Griffin S, Uhrés E, Hanotte O, Waits LP, Dubois-Paganon C, Burke T, Bouvet J. 1997. Noninvasive genetic tracking of the endangered Pyrenean brown bear population. Mol Ecol. 6:869–876.
Taberlet P, Cheddadi R. 2002. Ecology. Quaternary refugia and persistence of biodiversity. Science. 297:2009–2010.
Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Res. 24:3189–3194.
Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics. 105:437–460.
Tajima F. 1989a. DNA polymorphism in a subdivided population: the expected number of segregating sites in the two-subpopulation model. Genetics. 123:229–240.
Tajima F. 1989b. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123: 585–595.
Taylor AB, Goldsmith ML. 2003. Gorilla biology: a multidisciplinary perspective. Cambridge: Cambridge University Press.
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. Mol Ecol. 13:321–335.
Thalmann O, Serre D, Hofreiter M, Lukas D, Eriksson J, Vigilant L. 2005. Nuclear insertions help and hinder inference of the evolutionary history of gorilla mtDNA. Mol Ecol. 14: 179–188.
The Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature. 437:69–87.
Tosi AJ, Moreales JC, Melnick DJ. 2000. Comparison of Y chromosome and mtDNA phylogenies leads to unique inferences of macaque evolutionary history. Mol Phylogenet Evol. 17: 133–144.
Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. 1991. African populations and the evolution of human mitochondrial DNA. Science. 253:1503–1507.
Voight BF, Adams AM, Frisse LA, Qian Y, Hudson RR, Di Rienzo A. 2005. Interrogating multiple aspects of variation in a full resequencing data set to infer human population size changes. Proc Natl Acad Sci USA. 102:18508–18513.
Wakeley J, Hey J. 1997. Estimating ancestral population parameters. Genetics. 145:847–855.
Walsh PD, Abernethy KA, Bermejo M, et al. 2003. Catastrophic ape decline in western equatorial Africa. Nature. 422:611–614.
Warren KS, Verschoor EJ, Langenhuijzen S, Heriyanto, Swan RA, Vigilant L, Heeney JL. 2001. Speciation and intrasubspecific variation of Bornean orangutans, *Pongo pygmaeus pygmaeus*. Mol Biol Evol. 18:472–480.
Watterson GA. 1975. On the number of segregating sites. Theor Popul Biol. 7:256–276.
Watts DP. 2003. Gorilla social relationships: a comparative overview. In: Taylor AB, Goldsmith ML, editors. Gorilla biology: a multidisciplinary perspective. New York: Cambridge University Press.
Whitlock MC, Barton NH. 1997. The effective size of a subdivided population. Genetics. 146:427–441.
Wise C, Sraml M, Rubinsztein D, Easteal S. 1997. Comparative nuclear and mitochondrial genome diversity in humans and chimpanzees. Mol Biol Evol. 14:707–716.
Won Y-J, Hey J. 2005. Divergence population genetics of chimpanzees. Mol Biol Evol. 22:297–307.
Yamagiwa J. 1986. Activity rhythm and the ranging of a solitary male mountain gorilla (*Gorilla gorilla beringei*). Primates. 27:273–282.
Yamagiwa J, Mwanza N. 1994. Day-journey length and daily diet of solitary male gorillas in lowland and highland habitats. Int J Primatol. 15:207–224.
Yu N, Chen F-C, Ota S, Jorde LB, Pamilo P, Patthy L, Ramsay M, Jenkins T, Shyue S-K, Li WH. 2002. Larger genetic differences within Africans than between Africans and Eurasians. Genetics. 161:269–274.
Yu N, Jensen-Seaman MI, Chemnick L, Kidd JR, Deinard AS, Ryder O, Kidd KK, Li WH. 2003. Low nucleotide diversity in chimpanzees and bonobos. Genetics. 164:1511–1518.
Yu N, Jensen-Seaman MI, Chemnick L, Ryder O, Li WH. 2004. Nucleotide diversity in Gorillas. Genetics. 166:1375–1383.

Jody Hey, Associate Editor

Accepted September 29, 2006

Downloaded from http://mbe.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on March 24, 2016