Molecular Ecology (2007) **16**, 501–516 doi: 10.1111/j.1365-294X.2006.03159.x

Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (*Gorilla gorilla diehli*)

RICHARD A. BERGL*† and LINDA VIGILANT‡

**Anthropology Department, City University of New York Graduate Center, 365 Fifth Avenue, New York, NY 10016, USA, †New York Consortium in Evolutionary Primatology, ‡Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany*

Abstract

Recently developed methods of individual-based analysis of genetic data allow an unprecedented opportunity to understand the relationships among fragmented populations. By defining population structure and identifying migrant individuals, such analyses can provide a framework to aid in evaluating the threats posed by inbreeding and reduced genetic variability as a consequence of limited gene flow among fragments. Here we investigate population structure in the critically endangered Cross River gorilla (*Gorilla gorilla diehli*) by applying a suite of individual-based analyses to data obtained from between one-quarter and one-third of the estimated total population through the use of noninvasively collected DNA samples. The population structure inferred using data from 11 autosomal microsatellite loci was broadly consistent with geography and habitat fragmentation, but showed no simple isolation-by-distance effects. In contrast to previous field surveys, which suggested that all gorilla localities were isolated from one another, we infer low levels of gene flow and identify migrants between habitat fragments as well as individuals of admixed ancestry, suggesting persistent recent reproductive contact between many of the localities. These results are encouraging for the conservation of the Cross River gorilla population. Conservation efforts should strive to maintain connectivity between subpopulations that are still in migratory contact and attempt to restore connectivity where it has been lost.

Keywords: conservation genetics, *Gorilla gorilla diehli*, microsatellite genotyping, migration, noninvasive sampling, population structure

Received 13 May 2006; revision received 3 August 2006; accepted 15 September 2006

Introduction

Understanding the structure of threatened populations, particularly those that exist in degraded or fragmented habitats, is crucial for their effective conservation (Lande & Barrowclough 1987; Simberloff 1988; Hanski & Gilpin 1997; Taylor & Dizon 1999; Kraaijeveld-Smit *et al*. 2005). When small populations become fragmented and migration between subpopulations decreases or is eliminated, consequent increases in inbreeding and loss of genetic diversity can have serious negative effects on the long-term viability of population fragments and by extension, the population as a whole (Keller *et al*. 1994; Lacy 1997; Bjilsma *et al*. 2000; Sherwin & Moritz 2000; Coulon *et al*. 2004). Determining which subpopulations are in migratory contact with each other can highlight important dispersal corridors as well as identify isolated areas, thereby suggesting priority areas for conservation.

Researchers increasingly are turning to genetic methods for the inference of population structure (e.g. Ciofi & Bruford 1999; Paetkau *et al*. 1999; Maudet *et al*. 2002; Eggert *et al*. 2004; Piggott *et al*. 2005; Proctor *et al*. 2005; Aspi *et al*. 2006). Approaches employing multiple loci in the nuclear genome provide the most comprehensive description of

Correspondence: Richard A. Bergl, Fax: 1-212-772-5423; E-mail: rbergl@gc.cuny.edu

the effects of dispersal and subdivision upon the genetic variation of populations. Individual-based Bayesian approaches (see reviews in Manel *et al*. 2003; Beaumont & Rannala 2004; Bertorelle *et al*. 2004; Pearse & Crandall 2004) allow inference of population structure, gene flow and demographic history with greater precision than previous approaches which relied upon idealized population models and summary statistics (e.g. F_{ST}, Pearse & Crandall 2004).

Despite the potential of these methods for use in a broad conservation context, their application has often been limited to taxa that are commercially valuable (e.g. chickens, Rosenberg *et al*. 2001; bluefin tuna, Carlsson *et al*. 2004), from which high quality samples are easily acquired (e.g. Bavarian red deer, Kuehn *et al*. 2003; grand skinks, Berry *et al*. 2004; Norway rats, Abdelkrim *et al*. 2005), or which have high public profiles (e.g. mountain lions, Ernest *et al*. 2003; grizzly bears, Proctor *et al*. 2005; Finnish wolves, Aspi *et al*. 2006). Research on population-scale patterns of genetic diversity in wild primates has been limited (DiFiore 2003), but includes a study revealing population subdivision and demographic history in Bornean orangutans (Goossens *et al*. 2005, 2006).

Here we describe population structure and patterns of migration in the Cross River gorilla, an elusive and understudied primate, as inferred from nuclear microsatellite markers applied to DNA derived from a large collection of noninvasive samples. These gorillas are located at least 200 km northwest of other gorilla populations and are largely restricted to rugged highland areas straddling the Nigeria–Cameroon border. Recently revived as a distinct subspecies, *Gorilla gorilla diehli* (Sarmiento & Oates 2000; Groves 2001), the Cross River gorilla is one of Africa's most critically endangered primates (IUCN 2005). Recent surveys suggest that the total population likely numbers fewer than 300 individuals and is fragmented into as many as 10 separate localities with limited potential for reproductive contact and unknown population structure (Oates *et al*. 2003; Sunderland-Groves *et al*. 2003; Sunderland-Groves & Jaff 2004). This population is under intense threat from bushmeat hunting and habitat loss and fragmentation. Despite its distinctiveness and high degree of threat, little is known of the Cross River gorilla beyond cranial morphology, basic distribution and single-site studies of feeding ecology (Sarmiento & Oates 2000; Oates *et al*. 2003; Stumpf *et al*. 2003).

Materials and methods

Study area and sample collection

The gorillas occupy extremely rugged terrain ranging in altitude from less than 200 m above sea level (m a.s.l.) in the lowlands of Takamanda Forest Reserve in Cameroon and Cross River National Park in Nigeria, to above 2000 m a.s.l. near Kagwene Mountain in Cameroon. Their habitat is highly seasonal with marked dry and rainy seasons. The combination of wide altitudinal range and marked seasonality creates a varied habitat, including lowland forest, dry forest, submontane and montane forest and montane grassland (Oates *et al*. 2003, 2004; Bergl *et al*. in press). The gorillas appear concentrated in 10 areas surrounded by considerable human activity: three in Nigeria, six in Cameroon, and one which spans the border between the two countries (Fig. 1).

Between December 2002 and September 2004, faecal samples ($N = 322$) were collected from gorilla night nests and trails during intensive nest searches and reconnaissance walks at all known Cross River gorilla localities. We used flame-sterilized spatulas or disposable sterile tongue depressors to place ~1 cm^3 of faeces in RNALater (Ambion) solution at an approximately 1:7 ratio. All samples were estimated to be less than 5 days old when collected. Samples were stored at –20 °C following storage in the field for up to one month at ambient temperature.

DNA was extracted from faecal samples with a QIAamp DNA stool kit (QIAGEN) according to the manufacturer's instructions with small modifications (Nsubuga *et al*. 2004). Since DNA yield from field-collected faecal samples is unpredictable and low DNA concentration can lead to errors in subsequent analyses, we quantified the amount of DNA in each sample using real-time polymerase chain reaction (PCR). Samples which yielded sufficient quantities of usable DNA (Morin *et al*. 2001) were selected for genotyping via the PCR at 11 autosomal microsatellite loci polymorphic in gorillas (D1s550, D2s1326, D4s1627, D5s1457, D5s1470, D7s817, D7s2204, D8s1106, D10s1432, D16s2624, vWF; Bradley *et al*. 2000). Microsatellite loci were amplified in 20 μl reactions consisting of 1.25x Super-*Taq* buffer (HT Biotechnology), an additional 0.8 mM $MgCl_2$ for a final total concentration of 2.3 mM, 250 nM each primer with the fluorescent label (FAM, HEX or NED) on the forward primer, 250 μM each dNTP,16 μg BSA, 0.33 U Super*Taq* (HT Biotechnology, previously mixed 2:1 with *Taq* Start Antibody (BD Biosciences)), 2–4 μl DNA template and 10.5–12.5 μl ultra-pure H20. PCRs were conducted in a PTC-200 thermocycler (MJ Research) under the following conditions: initial denaturation of 3 min at 94 °C; 45 cycles of 30 s at 94 °C, 30 s at 55–60 °C, and 30 s at 72 °C; and a final extension for 30 min at 72 °C. To control for allelic dropout (stochastic nonamplification of one allele), multiple PCR replicates were performed according to the concentration of DNA in each sample (Morin *et al*. 2001). In practice, DNA yields were such that most homozygote genotypes were confirmed by a minimum of seven replicates (Taberlet *et al*. 1996; Taberlet & Luikart 1999). All heterozygotes were observed in a minimum of two separate reactions. Sex of putative migrant individuals (see below) was determined through typing of the amelogenin locus (Bradley *et al*. 2001).

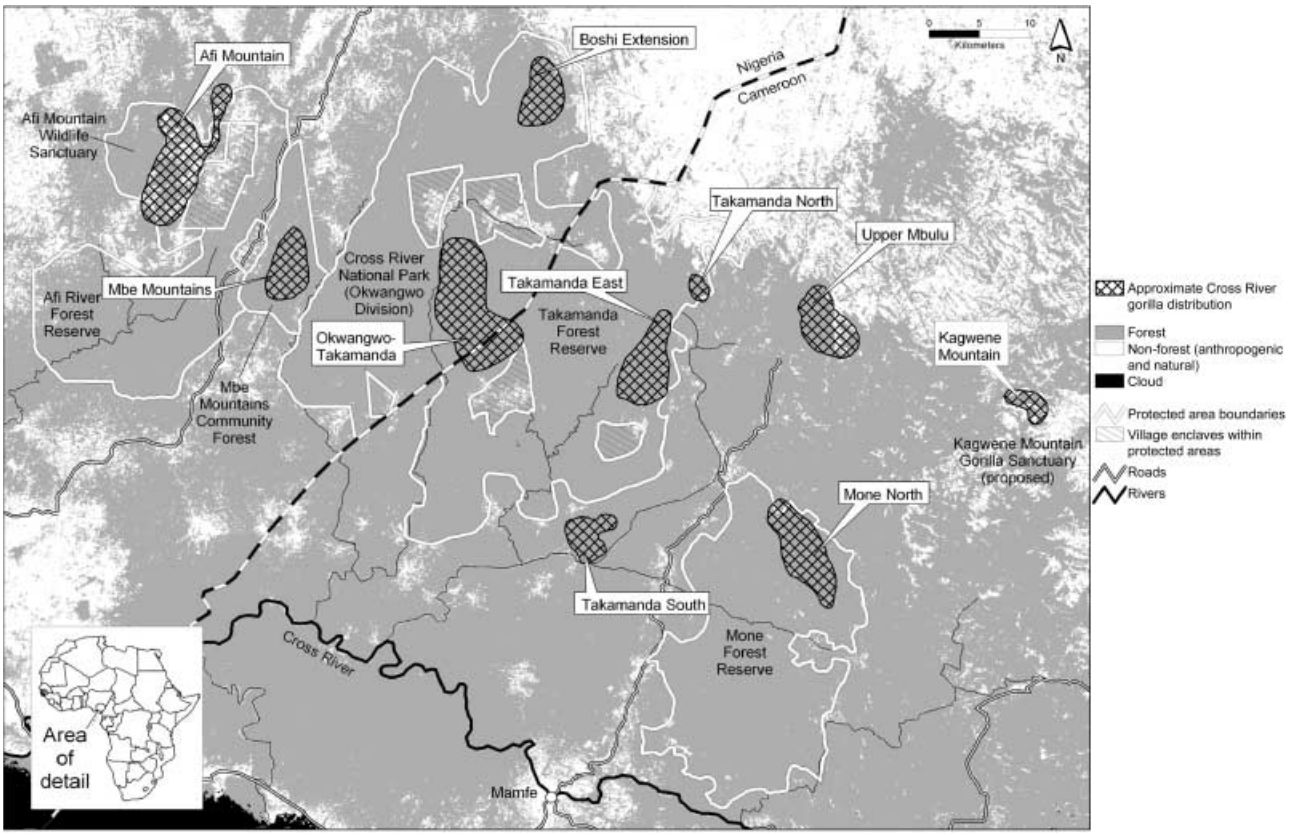

Fig. 1 Approximate distribution of the Cross River gorilla in Nigeria and Cameroon (newly discovered putative gorilla locality to the southeast not shown). Distribution of forest and nonforest based on Landsat 7 data from January 2003 (R. Bergl unpublished data). Gorilla locality names are indicated by text boxes.

PCR products were electrophoresed on an ABI 3100 genetic analyser and alleles sized using GENESCAN software (Applied Biosystems).

Discrimination of individuals

In order to ensure that duplicate samples (i.e. gorillas that were sampled more than once) were not included in the analysis, only unique multilocus genotypes were used. Unique genotypes were identified using the programs IDENTITY 1.0 (Wagner & Sefc 1999) and CERVUS (Marshall *et al*. 1998). Samples that had mismatches at up to two loci were re-examined for possible genotyping errors or allelic dropout. We calculated the probability of identity statistic, $P_{(ID)}$, the probability that two different individuals will share the same multilocus genotype at a given number of loci (Paetkau & Strobeck 1994). We used a more conservative variant of $P_{(ID)}$, $P_{(ID\text{-}sibs)}$ (the probability that a pair of siblings will share the same genotype; Waits *et al*. 2001) to ensure that the loci used could reliably discriminate related individuals. Allele frequencies and private alleles were calculated using CONVERT (Glaubitz 2004).

Linkage, null alleles, and Hardy—Weinberg equilibrium

Data were tested for deviations from linkage disequilibrium and Hardy–Weinberg equilibrium using GENEPOP 3.4 (Raymond & Rousset 1995). For loci with fewer than four alleles, the complete enumeration method (Louis & Dempster 1987) was used. In all other cases, tests employed the Markov chain method of Guo & Thompson (1992). The presence of null alleles, stuttering and small allele dominance was tested using MICROCHECKER (van Oosterhout *et al*. 2004). Data were tested both as a single population and by sampling locality. Significance values for multiple comparisons were adjusted by Bonferroni correction (Rice 1989).

Population structure

Two different Bayesian analyses were used to investigate population subdivision in the Cross River gorilla. First, we used the model-based clustering method implemented in STRUCTURE 2.1 (Pritchard *et al*. 2000) to determine the optimal number of genetic clusters present in the population. STRUCTURE divides sampled individuals into a number of clusters (K) independent of locality information (i.e. based

only on multilocus genotypic data), so as to minimize deviations from Hardy–Weinberg and linkage equilibrium. The program uses a Markov chain Monte Carlo (MCMC) procedure to estimate $P(X \mid K)$, the posterior probability that the data fit the hypothesis of K clusters. The program also calculates the fractional membership of each individual in each cluster (Q).

Real and simulated data have shown that it is not straightforward to determine the optimal value of K when complex population structure is present (Pritchard *et al*. 2000; Pritchard & Wen 2004; Evanno *et al*. 2005; McRae *et al*. 2005), so we took two approaches in choosing K. First, we calculated ΔK, a measure of the second order rate of change in the likelihood of K (Evanno *et al*. 2005). Simulation studies examining a range of patterns of genetic subdivision have shown that the modal value of ΔK corresponds to the most pronounced genetic subdivision present in the data. Second, we also compared posterior probabilities for the values of K with the highest $P(X \mid K)$ using a Mann–Whitney U test, following the approach of Rosenberg *et al*. (2001). We conducted 20 independent runs for each K between 1 and 10 using the admixture model and correlated allele frequencies. Exploratory STRUCTURE runs demonstrated that a burn-in period of 500 000 steps, followed by 10^6 steps of data collection, was sufficient to ensure convergence of the MCMC.

A potential drawback of STRUCTURE's clustering technique is that it assumes all potential source populations have been sampled. This can lead to misassignment of migrant individuals and individuals with migrant ancestry from unsampled populations. We performed an exclusion test (Cornuet *et al*. 1999) in GENECLASS 2.0 using population simulations in order to statistically test whether one or more of the sampled localities could be ruled out as the area of origin for each individual. The probability of individual genotypes coming from each locality was calculated by comparing individual genotypes to 10 000 simulated individuals per locality. We selected the simulation method introduced by Paetkau *et al*. (2004) as it is more representative of real population processes than other methods (e.g. Rannala & Mountain 1997; Cornuet *et al*. 1999) which have been shown to produce an inflated rate of type I errors (Paetkau *et al*. 2004; Piry *et al*. 2004). This resampling method models populations via sampling of hypothetical gametes (as opposed to alleles), which are then combined to create simulated individuals.

We also estimated F_{ST} using θ (Weir & Cockerham 1984). F_{ST} estimates were calculated for all sampling locality pairs in ARLEQUIN 2.0 (Schneider *et al*. 2000). Significance of pairwise F_{ST} values was calculated using permutation tests ($N = 1000$) and Bonferroni corrected for multiple comparisons. We tested for correlation between F_{ST} ($F_{ST}/(1 - F_{ST})$; Rousset 1997) and geographical distance using Mantel tests implemented in the ISOLDE extension of GENEPOP.

Detection of migrants

The two Bayesian approaches, STRUCTURE and GENECLASS 2.0, were also used to identify migrants and those individuals with migrant or mixed ancestry. Though prior population information is not used in the clustering approach taken by STRUCTURE when investigating population structure, it can be incorporated when attempting to determine which individuals are not residents of their sampled population. When the USEPOPINFO option of STRUCTURE is employed, the program assumes an initially high probability that each individual is a resident of its sampling locality. Using prior population information greatly assists the clustering process and allows the program to calculate posterior probabilities that individuals belong to their sampled locality/cluster. STRUCTURE was run this way with the previously inferred STRUCTURE cluster memberships ($K = 3$, from clustering analysis without population information) used as prior population information. When limited information about migration is present in the data, this method of analysis can be sensitive to the migration rate (MIGPRIOR) assigned as an initial condition (Pritchard & Wen 2004). Therefore, we conducted runs using a range of values (0.001–0.1) for MIGPRIOR as suggested by Pritchard *et al*. (2000). Choice of MIGPRIOR did not substantively affect the program output, so only results for MIGPRIOR = 0.05 are presented here. Burn-in and run length were the same as for runs without prior population information.

We selected the 'detect migrants' function in GENECLASS 2.0 as it is explicitly designed to identify first generation migrants (Paetkau *et al*. 2004; Piry *et al*. 2004), i.e. individuals born in a population other than the one in which they were sampled. Given the relatively long generation time in gorillas and the recent and ongoing fragmentation of the Cross River gorillas' habitat, data on recent migration will be most representative of current population processes. GENECLASS 2.0 uses a suite of likelihood-based statistics, in combination with resampling methods, to calculate probabilities that individuals are first generation migrants. We used two different likelihood-based test statistics to identify migrant individuals. L_h, the likelihood of finding a given individual in the population in which it was sampled, is the most appropriate statistic to use when all potential source populations have not been sampled (Paetkau *et al*. 2004; Piry *et al*. 2004). However, L_h lacks power when compared to other estimators (Paetkau *et al*. 2004), and may cause migrant individuals whose source populations have not been sampled to go undetected. Thus, we also used L_h/L_{max}, the ratio of L_h to the greatest likelihood among all sampled populations (Paetkau *et al*. 2004), which has greater power, but is again most informative when all source populations have been sampled. With L_h/L_{max}, migrants from unsampled populations can be misclassified as

residents, since L_h and L_{max} will have similar values. We employed the Bayesian criterion of Rannala & Mountain (1997) in combination with the resampling method of Paetkau *et al*. (2004; described above) to determine the critical value of the test statistic (L_h or L_h/L_{max}) beyond which individuals were assumed to be migrants. We selected an alpha level of 0.01 to determine critical values, as simulated data have shown this level to represent an appropriate balance between stringency and power (Paetkau *et al*. 2004).

Population history

The relative likelihoods of two models of population structure, pure drift vs. immigration-drift equilibrium, were calculated with the program 2MOD (Ciofi *et al*. 1999). In metapopulations adhering to the pure drift model, allele frequencies in each member population are solely the product of random changes, and the effect of migration between populations is negligible. Conversely, in a situation of immigration-drift equilibrium, population allele frequencies are the result of a balance between gene flow and genetic drift. The 2MOD program uses an MCMC procedure to compare likelihoods of the two scenarios and produce probabilities of the data fitting each model. The MCMC simulation was run for 100 000 iterations, with the initial 10% of data discarded to avoid dependence on starting conditions. Probabilities of each model were calculated using both sampling locality and STRUCTURE cluster as the population unit. The probability that two alleles are identical by descent (F) was calculated as a relative measure of the effect of drift on individual populations. F was determined via density estimation using estimates of F from each step of the MCMC.

Results

Identification of unique individuals and descriptive statistics

Of the 322 samples collected, DNA quantification revealed that 184 (57.1%) contained sufficient DNA for further analyses. Rain and humidity, in combination with the steep terrain and cryptic behaviour of the gorillas, prevented the collection of consistently high quality samples from all sampling localities. The small number and poor quality of samples from three Cross River gorilla localities (Okwangwo-Takamanda, Takamada East and Takamanda North) resulted in these areas being unrepresented in the analysis. Individual multilocus genotypes were on average 97.4% complete. Genotypes with more than five missing alleles were excluded from further analysis, except for one individual with 10 missing alleles (54.5% complete) from an under-sampled area which was included as it could be identified as unique. Of the total set of genotypes, 71 (38.6%) were determined to represent unique individuals and constitute approximately one-quarter of the total estimated population (Table 1). Six pairs of samples from within the subset of 71 had mismatches at only one or two loci. These nearly matched individuals were always from the same locality. Since these pairs could not be definitively determined to represent a single individual, all analyses were conducted both with and without the potentially duplicated individuals. There was no significant difference in the results between the two sets of genotypes, so the results for the full set of 71 individuals are presented here.

Probability of identity calculations showed that the power of the loci used to discriminate between individuals was high. $P_{(ID\text{-}sibs)}$, the most conservative of the $P_{(ID)}$ estimators, was 1.32×10^{-4}, indicating that the probability that

Table 1 Faecal sample collection effort and number of individuals genotyped by locality

Locality	No. collected	No. genotyped	No. of unique individuals	Estimated population size*
Afi Mountain	73	34	18	25–30
Upper Mbulu	42	9	5	20–30
Takamanda South	35	21	10	15–25
Boshi Extension	16	11	6	20–25
Kagwene Mountain	79	62	15	20–30
Mbe Mountains	30	24	7	24–32
Mone North	38	23	10	25–35
Okwangwo-Takamanda	1	0	NA	20–30
Takamanda East	7	0	NA	20–30
Takamanda North	2	0	NA	15–25
Total	322	184	71	204–292

*Oates *et al*. (2003), Sunderland-Groves & Jaff (2004).

even related individuals would have the same genotype is extremely low. Thus, different samples which produced duplicate multilocus genotypes (and were excluded from the analysis) can be assumed with high confidence to represent the same individual.

When all sampling localities were pooled, no evidence of significant linkage disequilibrium (LD) was found ($P < 0.05$ after Bonferroni correction), but two loci showed evidence of nonamplifying 'null' alleles. However, when the sampling localities were tested individually, no null alleles were detected. This suggests that evidence of null alleles in the global test is the result of genetic structure in the population, not systematic nonamplification of an allele. Further tests for errors in the data showed no evidence for stuttering or small allele dominance.

When individual localities were examined for HWE, only one locus at one locality (Afi Mountain) was significantly out of equilibrium after Bonferroni correction. Such mild deviations from HWE are to be expected given the likely presence of related individuals in the sample (Bourgain *et al*. 2004; Lukas *et al*. 2004). As with tests for null alleles, significant deviations from HWE (Bonferroni corrected $P < 0.05$) were observed at two loci when all samples were tested as a single population. These results are consistent with a Wahlund effect (disequilibrium caused by treating several separate populations as one; Wahlund 1928) arising from the presence of population substructure. Overall, the results suggest that the individual sampling localities are at equilibrium.

Population structure

Calculation of ΔK from the STRUCTURE output produced a modal value of the statistic at $K = 3$ (Fig. 2). While the largest value of ΔK was at $K = 3$, a second mode was present at $K = 5$. The height of the modal values of ΔK indicates the strength of the population subdivision signal (Evanno *et al*. 2005), here suggesting deep subdivision at $K = 3$, and less pronounced differentiation at $K = 5$.

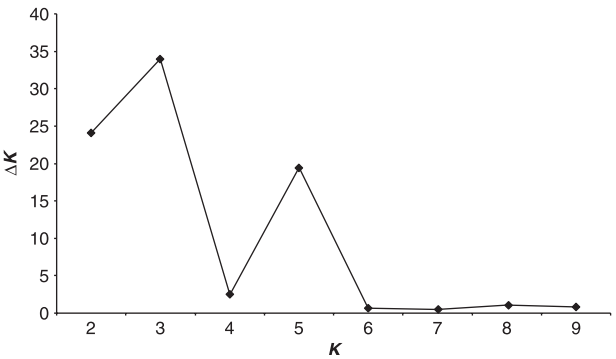

Fig. 2 ΔK (a measure of the rate of change in the STRUCTURE likelihood function) values as a function of K, the number of putative populations.

Examination of Ln P(X | K) values from the program also suggested a level of subdivision at $K = 5$. As has been reported with other data sets (Rosenberg *et al*. 2001; Evanno *et al*. 2005), variance in Ln P(X | K) increased at higher values of K. This variance prevented easy identification of the highest likelihood K. Clustering patterns of the highest likelihoods were found at $K = 5$, 6, and 7, though which of these solutions was most likely varied on a run-to-run basis. When median values of Ln P(X | K) were compared, $K = 5$ had the highest likelihood overall (Mann–Whitney U test, 2-tailed $P < 0.05$).

Thus, though there is evidence for population subdivision at both $K = 3$ and $K = 5$, $K = 3$ appears to be the optimal solution for the following reasons. In cases where the STRUCTURE program finds clustering solutions with similar probabilities at different values of K, the lowest value is typically the most accurate (Pritchard *et al*. 2000; Pritchard & Wen 2004). The presence of related individuals in the sample, as is likely the case in our sample, can also lead to overestimation of the true value of K (Pritchard & Wen 2004) and has been seen in other studies (e.g. Berry *et al*. 2004). Additionally, the model of correlated allele frequencies employed in the analysis, though it allows for the differentiation of closely related populations, is more likely to overestimate K than other models (Pritchard *et al*. 2000). We therefore chose the three cluster solution ($K = 3$) as the hierarchical level which best describes the genetic subdivision in our sample of the Cross River gorilla population.

All runs at $K = 3$ produced identical clustering solutions (Fig. 3) with similar values of cluster membership Q for all individuals within localities. The three clusters correspond roughly to geography, with the majority of individuals from each sampling locality clustering together. A Western cluster is present, with all individuals from Afi Mountain assigned to this cluster. A limited number of Afi individuals also have partial inferred ancestry (mean $Q = 0.07$, range 0.008–0.45) in a Central cluster. The majority of gorillas from Upper Mbulu also belong to the Western cluster and most Boshi Extension individuals show significant fractional membership (mean $Q = 0.25$, range 0.06–0.58) there.

A Central cluster is formed by individuals from Takamanda South, Boshi extension, Mbe Mountains, and Mone North. Mbe Mountains and Mone North individuals have ancestry almost exclusively in this cluster. One Mbe individual is assigned to the Western cluster, and one Mone North individual has partial ancestry ($Q = 0.38$) there. The majority of Takamanda South individuals assign to this cluster, though three have full or partial membership (mean $Q = 0.86$, range 0.80–0.97) in the Western Cluster. Boshi individuals also assign primarily to this cluster, though as mentioned above, most individuals have some ancestry in the Western cluster. All of the Kagwene Mountain gorillas assign

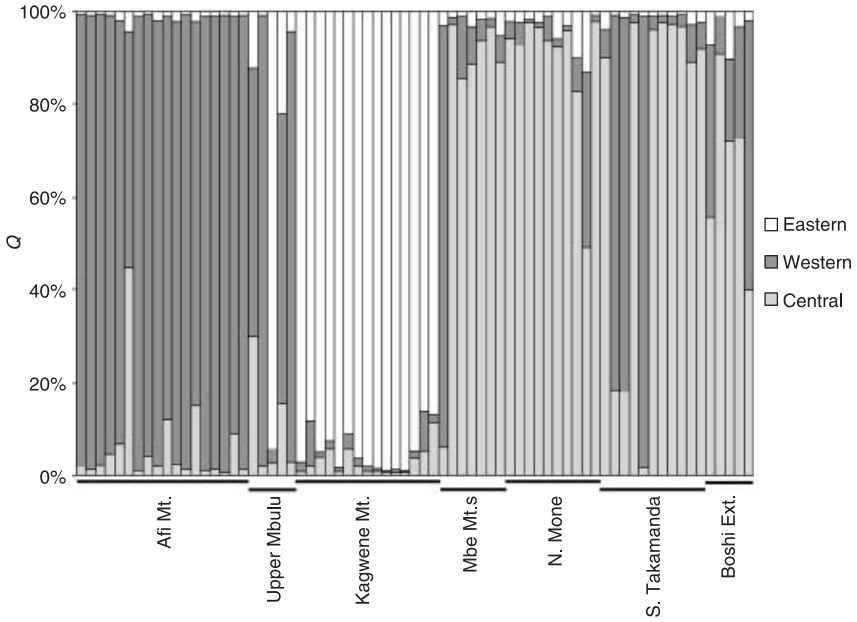

Fig. 3 Proportional membership (*Q*) of each gorilla in the three clusters identified by STRUCTURE. Each gorilla is represented by a single vertical bar. The locality of origin for each individual is indicated below.

unambiguously to an Eastern cluster. Little evidence of ancestry from other clusters is found in this locality. One individual from Upper Mbulu also assigns to this cluster.

At $K = 5$, more fine-grained population structure is present, but subdivision still closely follows geography. The major change is that both Boshi Extension and Takamanda South assign to their own clusters, while the overall pattern of clustering remains the same. As at $K = 3$, a level of genetic affinity remains between Afi, Boshi, and Upper Mbulu, with four Afi individuals assigning to the Boshi cluster. In a minority of runs, Takamanda South, Mbe Mountains and Mone North each had slightly different fractional membership in the same two clusters. Kagwene Mountain remained distinct from the other clusters and showed little evidence of gene flow from other areas.

The rate at which individuals correctly assign to their sampled locality can also be used as an assessment of population genetic structure (Manel *et al*. 2005). The exclusion test with resampling produced an accurate assignment rate of 80.3% when the locality of highest probability was considered (Fig. 4). The majority of misassignments were either to localities belonging to the same cluster, or represented individuals identified in subsequent analyses as migrants. In 73.6% of correct assignments, additional localities other than that of highest probability could not be ruled out as the source population. The pattern of locality assignments that could not be excluded mirrors the results of the STRUCTURE analysis. Generally, localities which could not be rejected as having nonsignificant assignment probabilities were those that according to STRUCTURE had substantial ancestry from, or clustered with, the locality in question. For example, 77.8% of Afi Mountain individuals

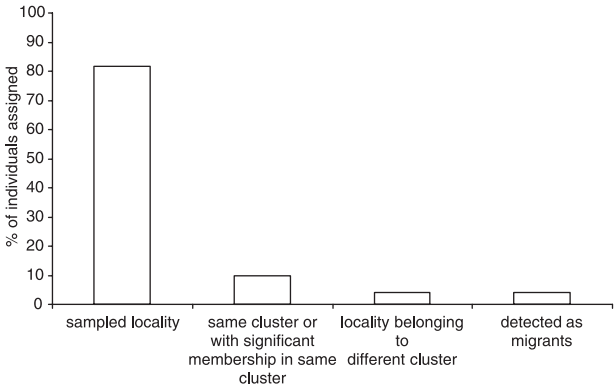

Fig. 4 Distribution of highest probability assignments as determined using the resampling procedure in GENECLASS 2.0.

cannot be excluded from assignment to Upper Mbulu. Both localities were grouped together by STRUCTURE's clustering algorithm. GENECLASS also assigned at the same frequency (77.8%) nontrivial probabilities to Boshi Extension as the area of origin of Afi individuals. This result is concordant with the STRUCTURE analysis where Boshi Extension has substantial ancestry in the Western cluster. Results for other localities are similarly clear-cut: 91.7% of misassignments of Takamanda South individuals are to members of STRUCTURE's Central cluster, while the figure for the Mbe Mountains is 75%. Kagwene Mountain individuals have only 12 potential misassignments, and the majority (70%) of these are very low probability ($P = 0.05$–0.08). Indeed, misassignments in general, though they could not be statistically excluded, had low probabilities (mean $P = 0.16$).

Table 2 Pairwise F_{ST} values (above diagonal) and significance (below diagonal) for each Cross River gorilla locality. Significant values indicated with * (Bonferroni corrected $P < 0.05$)

	Afi Mountain	Upper Mbulu	Takamanda South	Boshi Extension	Kagwene Mountain	Mbe Mountains	Mone North
Afi Mountain	—	0.04744	0.11843	0.14204	0.24214	0.10499	0.12121
Upper Mbulu	NS	—	0.11251	0.11746	0.14556	0.05988	0.106
Takamanda South	*	*	—	0.14835	0.19705	0.07706	0.10445
Boshi Extension	*	NS	*	—	0.20454	0.09153	0.14189
Kagwene Mountain	*	*	*	*	—	0.2079	0.18603
Mbe Mountains	*	NS	*	*	*	—	0.08738
Mone North	*	*	*	*	*	*	—

Pairwise F_{ST} comparisons suggest the same pattern of relationships as the Bayesian methods (Table 2). Within-cluster F_{ST} values were almost uniformly lower than between-cluster comparisons, and peripheral localities show greater differentiation than those in the centre of the gorillas' range. However, somewhat contrary to the individual-based analyses, most localities show significant divergence from one another. The exception is Upper Mbulu, which as in the Bayesian analyses shows affinity with both Afi Mountain and Boshi Extension. Mantel tests for correlation between genetic and geographical distance were all nonsignificant ($P < 0.05$).

Detection of migrants and admixed individuals

The detection of migrants procedures in both STRUCTURE and GENECLASS produced very similar results (Table 3). Using both previously determined STRUCTURE cluster and geographical sampling locality as prior population information, STRUCTURE identified 3 individuals (B22-3, B20-2, and B3-5) as potential migrants, or of migrant ancestry ($P = 0.97$, 0.748, 0.441). All other individuals had high probabilities of being residents ($P > 0.882$ at $K = 3$) with the probability of residency rapidly increasing to 0.99. Each of the apparent migrants was an individual strongly cross-assigned (highest $Q > 0.90$) to a nonhome cluster in the analysis without population information. Two of the three putative migrants are assigned to clusters adjacent to the sampled locality; sample B3-5 from the Mbe Mountains (Central cluster) is identified as a migrant from the Western cluster and sample B22-3 collected from Upper Mbulu (Western cluster) assigns to the Eastern cluster.

GENECLASS identified four individuals as migrants ($P < 0.01$), three with the L_h/L_{max} ratio (B22-3, B20-2, and B3-5), and two with L_h (B20-2 and B17-11, B20-2 was selected using both likelihood methods). The individuals identified with L_h/L_{max} are the same individuals identified by STRUCTURE as having high probability of being migrants. B22-3 (detected with the L_h/L_{max} ratio) assigned strongly to a particular locality in the exclusion test, indicating a seemingly clear population of origin. B3-5 assigned equally to two localities, though one (Afi Mountain) is more likely given its geographical proximity to the collection locality. These individual assignments are in accord with the clustering results from STRUCTURE. The migrants identified with L_h had consistently low assignment probabilities for all localities in the exclusion test, suggesting that these individuals represent migrants from unsampled areas. Only one of these samples (B20-2) was also selected by STRUCTURE. Unlike STRUCTURE, GENECLASS does not assume that all source populations have been sampled and so it is more apt to detect migrants from such unsampled populations. B22-3 and B20-2 were identified as females, while B3-5 and B17-11 were identified as males.

STRUCTURE also identified a number of individuals not readily classified as migrants, but not clearly assigned as residents either, suggesting that these individuals are the products of admixture between localities. In STRUCTURE, potentially admixed individuals are those that do not assign with the majority of individuals from their locality, or which have values of Q that indicated nontrivial membership in more than one cluster. Ranking and plotting individual Q-values following the approach of Beaumont *et al*. (2001) allowed delineation of a set of samples that did not clearly group in any one cluster (Fig. 5). Clear breaks are present at $Q = 0.8$ and $Q = 0.2$. Individuals with Q scores falling exclusively above and below these values assign strongly to one cluster. We defined individuals with Q-values from 0.8 to 0.2 as potentially admixed (Lecis *et al*. 2006; Vähä & Primmer 2006), resulting in seven samples of mixed ancestry (although eight individuals fall within the admixture range, one was identified by previous analyses as a migrant). These admixed gorillas had substantially lower migrant probabilities than did putative migrants, but their probabilities were still higher (mean = 0.066) than the average value for all resident individuals (mean = 0.009). Two individuals that assigned with a cluster different to the majority cluster for their sampling locality were also identified.

GENECLASS is not specifically designed to identify admixed individuals. However, 10 gorillas (B12-4, B2-3, B22-6, B27-8,

Table 3 Results of migrant detection analyses

Sample	Geographic origin	STRUCTURE Q (C/W/E clusters; no prior population information, $K = 3$)	GENECLASS locality of highest probability assignment–exclusion test	GENECLASS highest assignment probability	GENECLASS F_0 migrant probability (L_h; L_h/L_{max} indicated with ^, *$P < 0.01$)	STRUCTURE migrant probability	Final migrant/admixture classification
B22-3	Upper Mbulu	0.026/0.030/0.944	Kagwene Mountain	0.788	0.993^*	0.970	MS
B20-2	Takamanda South	0.017/0.973/0.009	Upper Mbulu	0.078	0.995^*	0.748	MU
B17-11	Mone North	0.492/0.380/0.129	Boshi Extension	0.058	0.994*	0.054	MU
B3-5	Mbe Mountains	0.063/0.907/0.031	Afi Mountain/Upper Mbulu	0.3498/0.317	0.997^*	0.441	MS/MU
B12-4	Upper Mbulu	0.298/0.581/0.121	Boshi Extension	0.172	0.984	0.117	AD
B2-3	Afi Mountain	0.447/0.509/0.044	Mbe Mountains	0.067	0.960	0.059	AD
B22-6	Upper Mbulu	0.156/0.624/0.220	Upper Mbulu	0.303	0.704	0.034	AD
B27-8	Boshi Extension	0.558/0.370/0.072	Boshi Extension	0.439	0.574	0.020	AD
B27-10	Boshi Extension	0.722/0.176/0.102	Boshi Extension	0.910	0.099	0.007	—
B27-11	Boshi Extension	0.728/0.239/0.032	Boshi Extension	0.999	0.001	0.003	—
B28-1	Boshi Extension	0.398/0.584/0.018	Boshi Extension	0.835	0.172	0.009	—
B16-8	Mone North	0.975/0.008/0.016	Mone North	0.081	0.92	0.001	—
B1-5	Afi Mountain	0.020/0.974/0.006	Afi/Boshi	0.8174/0.8143	0.188	0.002	—
B4-10	Afi Mountain	0.024/0.955/0.021	Afi/Upper Mbulu	0.4934/0.4498	0.517	0.005	—
B6-6	Afi Mountain	0.014/0.976/0.010	Afi/Upper Mbulu	0.5904/0.5919	0.405	0.001	—
B27-9	Boshi Extension	0.909/0.080/0.011	Boshi Extension	0.062	0.942	0.010	—
B19-10	Takamanda South	0.185/0.803/0.012	Takamanda South	0.664	0.343	0.007	—
B21-11	Takamanda South	0.892/0.080/0.028	Awurri/Boshi/Mone North	0.224/0.202/0.197	0.767	0.003	—
B19-11	Takamanda South	0.183/0.803/0.014	Takamanda South	0.803	0.205	0.007	—

MS, migrant whose source locality was determined; MU, migrant whose source locality could not be determined; AD, admixed individual.

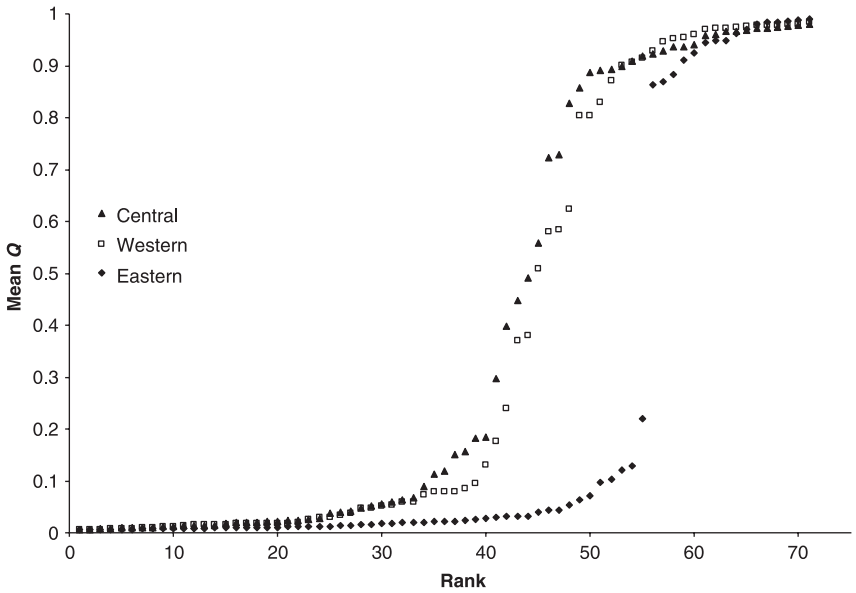

Fig. 5 Ranked mean *Q* (proportional membership in each cluster) for each individual in each cluster. Admixed individuals have values between 0.2 and 0.8.

B16-8, B1-5, B4-10, B6-6, B27-9, B21-11) assigned to areas other than their home locality, had low assignment probabilities for all localities, or had similar probabilities for several areas. Though these samples may represent individuals who cannot be accurately assigned due to a lack of information in the data, the low or similar probabilities could also be indicative of admixed ancestry.

We used a conservative approach and classified individuals as admixed if both programs suggested evidence of migrant ancestry. Thus, four individuals with mixed *Q*, low (< 0.50) or ambiguous GENECLASS assignment probabilities, and high (> 0.50) probabilities of being an F_0 migrant were identified as having migrant ancestry (B12-4, B2-3, B22-6, B27-8). Based on our criteria, we could not definitively identify any other animals as admixed. However, since these criteria require admixed *Q* in STRUCTURE, individuals with ancestry from more than one locality within the same cluster will not be identified as admixed. Six more individuals (B16-8, B1-5, B4-10, B6-6, B27-9, B21-11), though not showing admixture in STRUCTURE, did not clearly assign with their sampling locality in the GENECLASS analysis and may represent additional individuals of migrant ancestry.

Population history

The probabilities of the two models of gene flow (pure drift or immigration-drift equilibrium) were calculated using the coalescent-based approach of Ciofi & Bruford (1999). When all sampling localities were considered as separate populations, neither the pure drift model, nor the model of immigration-drift equilibrium was substantially more likely [P (immigration-drift equilibrium) = 0.524, Bayes factor = 1.1]. However, when STRUCTURE cluster (*K* = 3) was used as the unit of population, the value of *P* (immigration-drift equilibrium) increased to 0.906 (Bayes factor = 9.6). An immigration-drift equilibrium model would not be in accord with the population structure and migrant analyses which showed relatively greater genetic isolation of the extreme Eastern and Western localities. However, when *F* values were examined, the greater isolation of the Western and Eastern clusters was evident (Fig. 6). The probability of genes being identical by descent was low in the Central cluster (*F* = 0.05, 95% highest posterior density (HPD) range: 0.026–0.104). In contrast, the Western and Eastern cluster had higher *F* values (*F* = 0.16, HPD range: 0.098–0.254 and *F* = 0.19, HPD range: 0.125–0.292, respectively), suggesting that, relative to the Central cluster, these areas are more influenced by drift. This is consistent with our other analyses which did not detect migration into either the Western or Eastern clusters.

Discussion

Population structure of the Cross River gorilla

Through the use of noninvasively collected genetic data, we were able to detect previously unknown population structure in the Cross River gorilla. Our analysis suggests that three subpopulations are present; namely, a large central subpopulation consisting of the majority of known Cross River gorilla localities, and two peripheral subpopulations represented by the gorillas of Afi and Kagwene mountains, respectively. This pattern of subdivision corresponds largely to patterns of habitat fragmentation. The constituent localities of the central subpopulation are all (with the exception of Mone North which is separated by a small road and scattered farm land) connected by

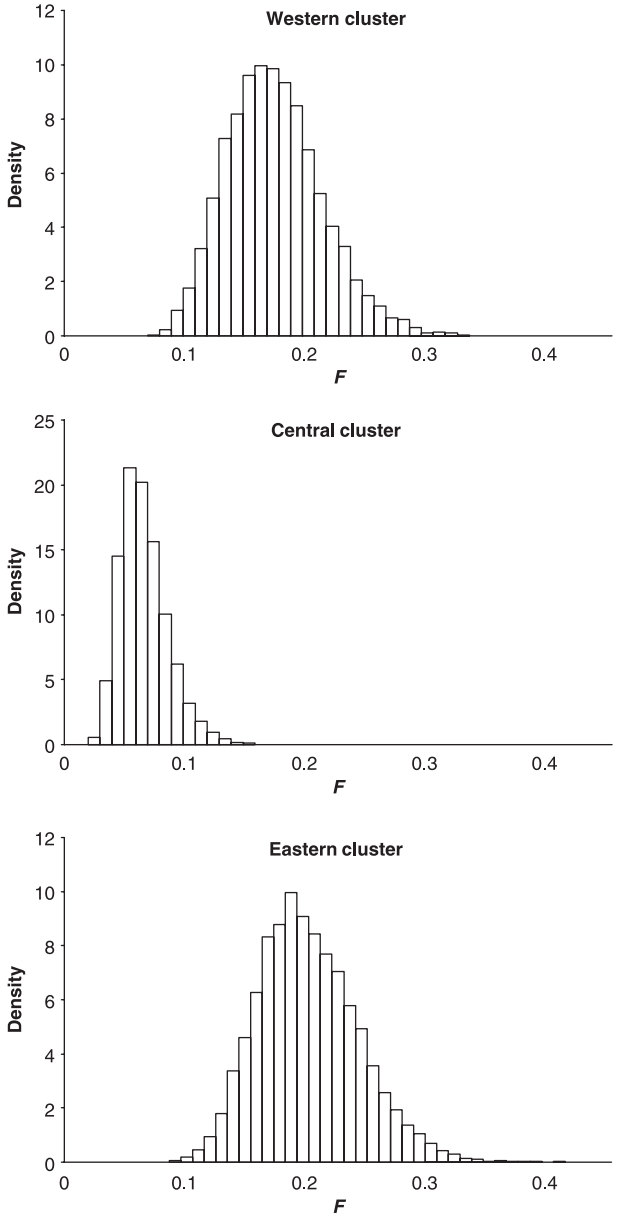

Fig. 6 Density plots of *F*, the probability that two alleles are identical by descent, for each Cross River gorilla cluster.

continuous forested lowland habitat. In contrast, Afi Mountain is almost totally isolated from other gorilla areas by farmland and a frequently travelled highway. Likewise, Kagwene Mountain, though connected to Upper Mbulu by a narrow forest corridor, is largely surrounded by substantial areas of montane grassland and farmland.

Our results imply that habitat discontinuities such as roads and farmland play a larger role in genetic substructuring of population than linear distance. Such relationships have also been observed in other large mammal taxa living in fragmented habitats including bears (Proctor *et al*. 2005) and orangutans (Goossens *et al*. 2005). Several anecdotal reports of Cross River gorilla presence in lowland areas exist from both historical (Mansfield 1908) and contemporary accounts (Fay 1987; Bucknell & Groves 2002; Bergl unpublished data), and western gorillas (*Gorilla gorilla* ssp. *gorilla*) have been reported to cross roads. The presence of migrants and forest corridors between the central localities also hints at the genetic affinities of the three unsampled Cross River gorilla localities. Though we were unable to collect usable samples from these areas, the pattern of relationships between sampled localities suggests that these unsampled areas would also belong to the central population.

One relationship that is not in accord with geography is the apparent genetic affinity between the Afi Mountain and Upper Mbulu gorillas. These localities are separated by of 50 km and several intervening gorilla localities, yet exhibit considerable genetic similarity in all analyses. This association could be observed if a historically continuous subpopulation including both areas had historically existed across the northern limit of the gorillas' range. As this subpopulation became fragmented by human activity and habitat loss, Afi Mountain and Upper Mbulu could have retained similar allele frequencies by chance, resulting in the association seen in our analyses. Alternatively, there may have been a historical division of the population into northern and southern groups and the current pattern of genetic similarity is the result of further subdivision. However, the sample size from Upper Mbulu is small, and so does not warrant excessive interpretation. We do not suggest that Afi Mountain and Upper Mbulu form a contemporary population.

While the genetic data strongly point to the presence of three subpopulations in the Cross River gorilla, and we were able to identify a number of migrant and admixed individuals in the Cross River population, there is also evidence of ongoing differentiation. The weak signal for division into five subpopulations, combined with the high frequency of accurate assignments, suggests that fragmentation of the gorilla population is not static. Although individual animals do move through lowland areas, the rate at which this movement occurs is likely decreasing, mediated by continued human activity in the region. The degree and duration of fragmentation is probably not yet extreme enough to create genetically isolated populations at all localities. However, such isolation is likely if current conditions continue.

F_{ST}-based analysis of population structure

Though many of the pairwise comparisons of F_{ST} were significant, F_{ST} values showed the same patterns of similarity between localities as the individual-based analyses. This observation supports the contention that F_{ST} and its various estimators are better used as a relative comparison

of differentiation when considering a network of populations (Nei 1973, 1987; Neigel 2002). Since F_{ST} requires a priori delineation of populations, and since this division may be somewhat ad hoc, the pattern of similarity is potentially more important than whether specific pairwise comparisons are significant. F_{ST} estimates may also be biased in situations where, as with the Cross River gorilla, populations are small, have experienced recent reductions in size, or have become recently fragmented and do not adhere to models of population structure upon which the estimator is based (Pearse & Crandall 2004). In particular, the likely presence of closely related individuals and parent–offspring pairs within samples from each locality will increase apparent levels of differentiation between localities and bias estimates of differentiation. Thus, F_{ST} is perhaps best applied to situations in which the units of analysis can be more readily defined beforehand. In situations where prior knowledge of population structure is limited, individual-based approaches such as those presented here are likely to produce results that are more biologically meaningful (Pearse & Crandall 2004).

Migration in the Cross River gorilla population

Prior research on this population had concluded, based on habitat discontinuities, hunting pressure in the lowlands, and absence of gorilla signs in intermediary areas, that each gorilla locality was effectively isolated from all others (Oates *et al.* 2003; Sunderland-Groves *et al.* 2003), and that gene flow between localities was therefore minimal (though some local hunters reported occasional movement of gorillas outside population nuclei, Bucknell & Groves 2002). The ranges of these gorillas were presumed to be limited to hilly and mountainous regions which are less affected by human disturbance. Though the gorillas are indeed concentrated in these inaccessible areas, our genetic data suggest that interchange of animal continues or has recently occurred between many of the gorilla localities. Our data show that some animals have migrated between localities within the current generation. We were able to identify as migrants four individuals (two males and two females) from four different localities. Two migrants (one male and one female) each moved from one of the two most isolated localities (Kagwene and Afi mountains) into the nearest neighbouring locality (Upper Mbulu and Mbe Mountains, respectively).

Neither of these two source localities appears to be the recipients of migrants. Both areas are genetically quite homogenous, though one Afi individual apparently does have admixed ancestry. In undisturbed gorilla populations, where there are relatively high densities of gorillas, animals that disperse from their natal group are likely to find other groups or individuals quite readily. However, in areas such as Kagwene and Afi mountains, where there appear to be only single social groups or very small communities of gorillas (Oates *et al.* 2003; Bergl unpublished data), opportunities for dispersing individuals to find new groups will be limited. Thus, dispersers will be under relatively greater pressure to travel long distances or cross barriers such as roads and disturbed forest, potentially resulting in large dispersal distances. This may explain the substantial distance (approximately 10 km) crossed by the migrants from these areas and why no migrants into these areas were found.

The other two migrants detected were collected in Takamanda South and Mone North, but could not be definitively assigned to a source population. Likely source localities for these migrants, based on geographical proximity, are among the unsampled gorilla areas. The presence of these individuals in the data set highlights the need for assessing whether all potential sources of migrants have been sampled in these types of analyses. If we had assumed that our data set represented all prospective source populations, at least one migrant would have gone undetected.

The identification of individuals with admixed ancestry was not as clear cut as the detection of migrants. Four gorillas were identified by both the STRUCTURE and GENECLASS analyses as having a high probability of migrant ancestry. Eleven other individuals showed evidence of admixture in results from one of the two programs. Thus, while we can only definitively identify four admixed individuals, several others may in fact have recent ancestry in an area other than their source locality. The presence of admixed individuals suggests that not only are gorillas able to move between localities on occasion, but that they are also able to reproduce in the new area. Such dispersal and subsequent reproduction is crucial for the maintenance of long-term genetic health in small, fragmented populations (Gilpin & Soulé 1986; Lacy 1997; Cunningham & Moritz 1998; Bjilsma *et al.* 2000).

Population history

Though our analyses of population structure indicate the presence of at least three subpopulations within the Cross River gorilla, the most probable model of gene flow was one of immigration-drift equilibrium. Such a model would predict a greater similarity in allele frequencies between subpopulations than we actually observed. The disparity between the two sets of analyses, in combination with high *F* values in the isolated subpopulations, suggests that the signal for immigration-drift equilibrium is an artefact of historical panmixia. In recent history, the Afi and Kagwene subpopulations were likely in much more regular contact with the central subpopulation. If this contact was still present, *F* values would be similar in each of the subpopulations. This was not what we observed. The peripheral areas have substantially higher levels of *F*, illustrating that drift plays a much larger role in determining allele frequencies in

these areas than in the central population. Migration from Kagwene and Afi into the central population, in combination with the unidirectional nature of this movement, is apparently not sufficient to overwhelm the effect of drift in the isolated areas and maintain genetic similarity across the gorillas' range. While allele frequencies best fit an equilibrium model, it appears that the isolation has simply not been maintained long enough for frequencies to shift to a pure drift model. Given the high *F* values and small population sizes in the peripheral areas, the influence of drift would be pronounced. That allele frequencies in the population as a whole do not support a pure drift model suggests that population fragmentation has been quite recent. These results emphasize the impact human disturbance has had on connectivity within the Cross River population in the recent past.

Conservation implications

The genetic population structure we detected in Cross River gorillas has important implications for the conservation and management of this critically endangered primate. Overall, our analyses suggest that the situation facing this population is not as dire as had been assumed. We documented reproductive connectivity during the current generation between several localities, including the most peripheral population nuclei, and genetic similarity between most of the sampled localities. Using conservative criteria, approximately 11% (8 of 71) of individuals were inferred to be migrants or to have recent ancestry from more than one locality. Though migration levels are low, they were considerably higher than prior nongenetic studies (Oates *et al*. 2003; Sunderland-Groves *et al*. 2003) had concluded. Even areas that had been assumed (Oates *et al*. 2003; Sunderland-Groves *et al*. 2003) to be isolated from other gorilla localities were determined to be in occasional migratory contact with at least the nearest locality. Unfortunately, equivalent data from other African ape populations are not available, so it is not possible to quantitatively compare levels of migration and genetic structure in the Cross River population to those of other less-fragmented populations. However, even in the absence of a comparative data set, our data suggest that the overall extent of migration across the Cross River population is substantially lower now than in recent history.

Conservation efforts must therefore, focus on the maintenance, and if possible, expansion, of forest connectivity between gorilla localities. While such actions present a challenge for both conservation biologists and wildlife managers, the situation allows for some optimism given that substantial habitat remains between many of the areas. Unlike the Cross River gorillas, some other threatened primate populations (e.g. orangutans, Goossens *et al*. 2005) exist in very highly fragmented habitats, necessitating substantial restoration of corridors. Only two of the gorilla localities are separated by habitat discontinuities (i.e. roads separating Afi Mountain and Mone North from the central portion of the gorillas' range), and in these cases, forested habitat abuts the discontinuity. In all other cases, at least a narrow forest corridor (and in many cases large continuous areas of forest) connects the population nuclei.

Much of the forest which constitutes Cross River gorilla habitat is already legally protected, and several additional areas are currently being considered for protected area status. Yet key corridors, such as those between Kagwene Mountain and Upper Mbulu, and between Mone North and the other members of the central subpopulation currently have no legal status. Revision of the status of these areas needs to be considered.

Besides habitat loss, other human activities such as bushmeat hunting in particular, can also limit migration. Extensive hunting in lowland areas is likely the main cause of the gorillas' current distribution in the highlands. This pressure needs to be relaxed in order to allow migration though, and potentially recolonization of, lowland habitat. Control of hunting is particularly important in the centre of the gorillas' range, which contains the largest concentration of gorillas and substantial areas (approximately 1300 km^2) of continuous forest. Conservation in this area presents a particular challenge, as it consists of a continuous forest block spanning the Nigeria–Cameroon border and the need for a cooperatively managed *trans* boundary protected area has been suggested (Sunderland-Groves & Jaff 2004).

Though connectivity between many of the gorilla localities remains, it is possible that in the near future some areas may become totally isolated. In such a case, the use of translocation as a management strategy should be considered. A discussion of the feasibility of translocation in the context of the Cross River gorilla population is beyond the scope of the current research. However, our results could help to guide future translocation projects, if such efforts were to be undertaken. Translocation efforts would be most effective in increasing levels of genetic diversity if animals are moved among each of the three genetically identifiable subpopulations (Crandall *et al*. 2000; Keller & Waller 2002). Our analysis of the demographic history of the Cross River gorilla suggests that the population was at immigration-drift equilibrium in the recent past, and that current subdivision is a product of human-induced fragmentation. Translocation between localities would therefore restore natural patterns of gene flow (Goossens *et al*. 2002) and avoid concerns associated with admixture of genetically distinct populations (IUCN 1998; Moritz 1999). Translocations could be particularly effective in increasing genetic diversity in the Afi Mountain and Kagwene Mountain localities given their small size, apparent lack of immigrant individuals and greater susceptibility to drift.

While the genetic data offer important insights for the management of the Cross River gorilla population, they are only one source of information about the threats facing these animals. Many other factors including bushmeat hunting, demographics, habitat loss and fragmentation, logistics, politics, financing, and feasibility must be considered when developing a comprehensive conservation plan for these animals.

Acknowledgements

Financial support for this project was provided by the National Geographic Society Conservation Trust, Conservation International, Lincoln Park Zoo, Primate Conservation Inc, and the Max Planck Society. We thank the National Parks Service (Nigeria), the Cross River State Forestry Commission (Nigeria), and the Ministry of Environment (Cameroon) for their allowing R.A.B. to conduct field work in areas under their management. We thank J. Oates for help developing and implementing the project and providing helpful comments on the manuscript. We also thank J. Sunderland-Groves, M. Ashu, E. Nwufoh, A. Mbong, C. Ransom, N. Mkpe, M. Tabeh, P. Jenkins, L. Gadsby, the Wildlife Conservation Society and local assistants too numerous to mention for assistance in the field. Thanks to A. Abraham, M. Arandjelovic, B. Bradley, D. Lukas, K. Langergraber, H. Siedel and O. Thalmann for assistance with laboratory analyses and helpful discussion. B. Goossens and two anonymous reviewers made substantive comments which significantly improved the manuscript.

References

Abdelkrim J, Pascal M, Calmet C, Samadi S (2005) Importance of assessing population genetic structure before eradication of invasive species: examples from insular Norway rat populations. *Conservation Biology*, **19**, 1509–1518.

Aspi J, Roininen E, Ruokonen M, Kojola I, Vila C (2006) Genetic diversity, population structure, effective population size and demographic history of the Finnish wolf population. *Molecular Ecology*, **15**, 1561–1576.

Beaumont M, Barratt EM, Gottelli D *et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology*, **10**, 319–336.

Beaumont MA, Rannala B (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics*, **5**, 251–261.

Bergl RA, Oates JF, Fotso R (in press) Distribution and protected area coverage of endemic taxa in West Africa's Biafran forests and highlands. *Biological Conservation*.

Berry O, Tocher MD, Sarre SD (2004) Can assignment tests measure dispersal? *Molecular Ecology*, **13**, 551–561.

Bertorelle G, Bruford M, Chemini C, Cristiano V, Hauffe HC (2004) New, flexible Bayesian approaches to revolutionize conservation genetics. *Conservation Biology*, **18**, 1–2.

Bjilsma R, Bundgaard J, Boerema A (2000) Does inbreeding affect the extinction risk of small population? Predictions from *Drosophila*. *Journal of Evolutionary Biology*, **13**, 502–514.

Bourgain C, Abney M, Schneider D, Ober C, McPeek MS (2004) Testing for Hardy–Weinberg equilibrium in samples with related individuals. *Genetics*, **168**, 2349–2361.

Bradley BJ, Boesch C, Vigilant L (2000) Identification and redesign of human microsatellite markers for genotyping wild chimpanzee (*Pan troglodytes verus*) and gorilla (*Gorilla gorilla gorilla*) DNA from feces. *Conservation Genetics*, **1**, 289–292.

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

Bucknell D, Groves JL (2002) *Local Perception of the Population Size, Distribution and Ranging Behaviour of the Cross River Gorilla Within the Takamanda and Mone Forest Reserves and the Mbulu Forest*. Cameroon Wildlife Conservation Society Report, New York.

Carlsson J, McDowell JR, Diaz-Jaimes P *et al.* (2004) Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranean Sea. *Molecular Ecology*, **13**, 3345–3356.

Ciofi C, Beaumont MA, Swingland IR, Bruford MW (1999) Genetic divergence and units for conservation in the Komodo dragon *Varanus komodoensis*. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **266**, 2269–2274.

Ciofi C, Bruford MW (1999) Genetic structure and gene flow among Komodo dragon populations inferred by microsatellite loci analysis. *Molecular Ecology*, **8**, S17–S30.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989–2000.

Coulon A, Cosson JF, Angibault JM *et al.* (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular Ecology*, **13**, 2841–2850.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, **15**, 290–295.

Cunningham M, Moritz C (1998) Genetic effects of fragmentation on a rainforest restricted lizard (Scincidae: *Gnypetoscincus queenslandiae*). *Biological Conservation*, **83**, 19–30.

DiFiore A (2003) Molecular genetic approaches to the study of primate behavior, social organization, and reproduction. *Yearbook of Physical Anthropology*, **46**, 62–99.

Eggert LS, Mundy NI, Woodruff DS (2004) Population structure of loggerhead shrikes in the California Channel Islands. *Molecular Ecology*, **13**, 2121–2133.

Ernest HB, Boyce WM, Bleich VC *et al.* (2003) Genetic structure of mountain lion (*Puma concolor*) populations in California. *Conservation Genetics*, **4**, 353–366.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Fay JM (1987) Report on the participation of J. Michael Fay in the Takamanda Gorilla Survey Project (May 1–May 20 1987). World Wildlife Fund-US, Washington D.C.

Gilpin M, Soulé ME (1986) Minimum viable populations: processes of species extinctions. In: *Conservation Biology: the Science of Scarcity and Diversity* (ed. Soule ME). Sinauer Associates, Sunderland, Massachusetts.

Glaubitz JC (2004) CONVERT: a user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes*, **4**, 309–310.

Goossens B, Chikhi L, Ancrenaz M *et al.* (2005) Patterns of genetic diversity and migration in increasingly fragmented and declining orang-utan (*Pongo pygmaeus*) populations from Sabah, Malaysia. *Molecular Ecology*, **14**, 441–456.

Goossens B, Chikhi L, Ancrenaz M *et al.* (2006) Genetic signature of anthropogenic population collapse in orang-utans. *Plos Biology*, **4**, 285–291.

Goossens B, Funk SM, Vidal C *et al.* (2002) Measuring genetic diversity in translocation programmes: principles and application to a chimpanzee release project. *Animal Conservation*, **5**, 225–236.
Groves CP (2001) *Primate Taxonomy*. Smithsonian Institution Press, Washington, D.C.
Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics*, **48**, 361–372.
Hanski IA, Gilpin M (1997) *Metapopulation Biology: Ecology and Evolution*. Academic Press, Toronto.
IUCN (1998) *IUCN Guidelines for Re-Introductions*. IUCN. Gland, Switzerland.
IUCN (2005) *2005 IUCN Red List of Threatened Species*. IUCN. Gland, Switzerland.
Keller L, Arcese P, Smith J, Hochachka WM, Stearns SC (1994) Selection against inbred song sparrows during a natural population bottleneck. *Nature*, **372**, 356–357.
Keller LF, Waller DM (2002) Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, **17**, 230–241.
Kraaijeveld-Smit FJL, Beebee TJC, Griffiths RA, Moore RD, Schley L (2005) Low gene flow but high genetic diversity in the threatened Mallorcan midwife toad *Alytes muletensis*. *Molecular Ecology*, **14**, 3307–3315.
Kuehn R, Schroeder W, Pirchner F, Rottmann O (2003) Genetic diversity, gene flow and drift in Bavarian red deer populations (*Cervus elaphus*). *Conservation Genetics*, **4**, 157–166.
Lacy RC (1997) The importance of genetic variation to the viability of mammalian populations. *Journal of Mammalogy*, **78**, 320–335.
Lande R, Barrowclough G (1987) Effective population size, genetic variation and their use in population management. In: *Viable Populations for Conservation* (ed. Soulé ME), pp. 87–123, Cambridge University Press, New York.
Lecis R, Pierpaoli M, Biro ZS *et al.* (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology*, **15**, 119–132.
Louis EJ, Dempster ER (1987) An exact test for Hardy–Weinberg and multiple alleles. *Biometrics*, **43**, 805–811.
Lukas D, Bradley BJ, Nsubuga AM *et al.* (2004) Major histocompatibility complex and microsatellite variation in two populations of wild gorillas. *Molecular Ecology*, **13**, 3389–3404.
Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology & Evolution*, **20**, 136–142.
Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.
Mansfield A (1908) *Urwald Dokumente. Vier Jahre Unter Den Crossfluss-Negern Kameruns*. Reimer, Berlin.
Marshall TC, Slate J, Kruuk LE, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7**, 639–655.
Maudet C, Miller C, Bassano B *et al.* (2002) Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [Capra ibex (ibex)]. *Molecular Ecology*, **11**, 421–436.
McRae BH, Beier P, Dewald E, Huynh Y, Keim P (2005) Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American Puma. *Molecular Ecology*, **14**, 1965–1977.
Morin PA, Chambers KE, Boesch C, Vigilant L (2001) Quantitative PCR analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzes (*Pan troglodytes verus*). *Molecular Ecology*, **10**, 1835–1844.
Moritz C (1999) Conservation units and translocations: strategies for conserving evolutionary processes. *Hereditas*, **130**, 217–228.
Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences, USA*, **70**, 3321–3323.
Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
Neigel JE (2002) Is F_{ST} obsolete? *Conservation Genetics*, **3**, 167–173.
Nsubuga AM, Robbins MM, Roeder AM *et al.* (2004) Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. *Molecular Ecology*, **13**, 2089–2094.
Oates J, McFarland KL, Groves JL *et al.* (2003) The Cross River gorilla: natural history and status of a neglected and critically endangered subspecies. In: *Gorilla Biology* (eds Taylor A, Goldsmith ML), pp. 472–497. Cambridge University Press, Cambridge, UK.
Oates JF, Bergl RA, Linder JM (2004) *Africa's Gulf of Guinea Forests: Biodiversity Patterns and Conservation Priorities*. Conservation International Center for Applied Biodiversity Science, Washington D.C.
van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.
Paetkau D, Amstrup SC, Born EW *et al.* (1999) Genetic structure of the world's polar bear populations. *Molecular Ecology*, **8**, 1571–1584.
Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.
Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology*, **3**, 489–495.
Pearse DE, Crandall KA (2004) Beyond F_{ST}: Analysis of population genetic data for conservation. *Conservation Genetics*, **5**, 585–602.
Piggott MP, Banks SC, Taylor AC (2005) Population structure of brush-tailed rock wallaby (*Petrogale penicillata*) colonies inferred from analysis of faecal DNA. *Molecular Ecology*, **15**, 93–105.
Piry S, Alapetite A, Cornuet JM *et al.* (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, **95**, 536–539.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
Pritchard J, Wen W (2004) Documentation for STRUCTURE software: version 2. Department of Human Genetics, University of Chicago, Chicago.
Proctor M, McLellan BN, Strobeck C, Barclay MR (2005) Genetic analysis reveals demographic fragmentation of grizzly bears yielding vulnerably small populations. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **272**, 2409–2416.
Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences, USA*, **94**, 9197–9201.
Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
Rice WR (1989) Analyzing tables for statistical tests. *Evolution*, **43**, 223–225.
Rosenberg NA, Burke T, Elo K *et al.* (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics*, **159**, 699–713.

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

Sarmiento EE, Oates JF (2000) The Cross River gorilla: a distinct subspecies *Gorilla gorilla diehli* Matschie 1904. *American Museum Novitates*, **3304**, 1–55.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, version 2.0: a software for population genetics data analysis. Department of Anthropology and Ecology, University of Geneva, Geneva.

Sherwin WB, Moritz C (2000) Managing and monitoring genetic erosion. In: *Genetics, Demography and Viability of Fragmented Populations* (eds Young AJ, Clarke G). Cambridge University Press, New York.

Simberloff D (1988) The contribution of population and community biology to conservation science. *Annual Review of Ecology and Systematics*, **19**, 473–511.

Stumpf R, Polk J, Oates J *et al.* (2003) Patterns of diversity in gorilla cranial morphology. In: *Gorilla Biology: a Multidisciplinary Perspective* (eds Taylor A, Goldsmith M), pp. 35–61. Cambridge University Press, Cambridge, UK

Sunderland-Groves JL, Jaff B (2004) *Developing a Conservation Strategy for the Cross River Gorilla: Proceedings of the 2nd International Workshop and Conference on the Conservation of the Cross River Gorilla*. Wildlife Conservation Society, New York.

Sunderland-Groves JL, Maisels F, Ekinde A (2003) Surveys of the Cross River gorilla and chimpanzee populations in Takamanda Forest Reserve, Cameroon. In: *Takamanda, the Biodiversity of an African Rainforest*. (eds Comiskey JA, Sunderland TCH, Sunderland-Groves JL), pp. 129-140. Smithsonian Institution, Washington, D.C.

Taberlet P, Griffin S, Goossens B *et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, **24**, 3189–3194.

Taberlet P, Luikart G (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society*, **68**, 41–55.

Taylor BL, Dizon AE (1999) First policy then science: why a management unit based solely on genetic criteria cannot work. *Molecular Ecology*, **8**, S11–S16.

Vähä J, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, **15**, 63–72.

Wagner HW, Sefc KM (1999) Identity 1.0. Centre for Applied Genetics, University of Agricultural Sciences, Vienna.

Wahlund S (1928) Zusammensetzung von Populationen und Korrelationserscheinungen von Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, **11**, 65–106.

Waits LP, Luikart G, Taberlet P (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.

This paper is a part of Richard Bergl's PhD dissertation research, in which he uses a variety of field and laboratory-based approaches to examine the conservation status of the Cross River gorilla population. His primary interest is linking genetics and landscape ecology with the conservation of endangered populations. Genetic analysis for this project was conducted at the Max Planck Institute for Evolutionary Anthropology in the molecular primatology laboratory of Linda Vigilant. She is interested in the distribution of genetic variation in wild primate populations and the possible effects of kin relationships upon social behaviour.
