Kin Selection, Social Structure, Gene Flow, and the Evolution of Chimpanzees

Phillip A. Morin,* James J. Moore, Ranajit Chakraborty, Li Jin, Jane Goodall, David S. Woodruff

Hypotheses about chimpanzee social behavior, phylogeography, and evolution were evaluated by noninvasive genotyping of free-ranging individuals from 20 African sites. Degrees of relatedness among individuals in one community were inferred from allele-sharing at eight nuclear simple sequence repeat (SSR) loci. Males are related on the order of half-siblings, and homozygosity is significantly increased at several SSR loci compared to Hardy-Weinberg expectations. These data support the kin-selection hypothesis for the evolution of cooperation among males. Sequence variation patterns at two mitochondrial loci indicate historically high long-distance gene flow and clarify the relationships among three allopatric subspecies. The unexpectedly large genetic distance between the western subspecies, *Pan troglodytes verus*, and the other two subspecies suggests a divergence time of about 1.58 million years. This result, if confirmed at nuclear loci and supported by eco-behavioral data, implies that *P. t. verus* should be elevated to full species rank.

Almost nothing is known about genetic variation of the chimpanzee, *Pan troglodytes*, in nature. Hypotheses concerning chimpanzee sociobiology and evolution have gone untested for many years because of difficulties surrounding tissue acquisition for genetic analysis. We here demonstrate a method of noninvasive genotyping based on DNA amplified from hair shed in night nests or plucked from captive apes (*1*). Using eight hypervariable simple sequence repeat (SSR) nuclear loci and two mitochondrial (mtDNA) sequences, we demonstrate how multilocus data can be used to provide critical information concerning kin selection, mating structure, individual reproductive success, inclusive fitness, population structure, gene flow, phylogeography, and phylogenetic relationships. This technical advance also permits a reexamination of the relevance of chimpanzee behavioral ecology (*2, 3*) to the elucidation of the evolution of human behavioral and genetic patterns. Although this report focuses on one species of primate, these methods are immediately applicable to many others of concern to ecologists and evolutionary and conservation biologists.

Chimpanzees are geographically widespread and eco-behaviorally diverse. Despite up to 34 years of field study at a few sites, little is known about the phylogeny of the chimpanzee subspecies, their times of divergence, historical population sizes, or patterns of gene flow. Phylogenetic reconstructions are hampered by a complete lack of a fossil record, and data on long-distance dispersal and genetic variation. This study of genetic variation of chimpanzee populations across Africa describes phylogeographic patterns, begins to test hypotheses for the evolution of chimpanzees' social structure, and sets the foundation for comparative studies among communities within and between subspecies.

The Kasakela community in the Gombe National Park, Tanzania (Fig. 1, site L), is the longest studied wild chimpanzee population (*2*). At Gombe, we sampled all (N = 43) individuals in the Kasakela social community (*1, 4*). Adult chimpanzees make nests (in trees) of leaves and branches, and, with the exception of mother-infant pairs, typically sleep individually in a fresh nest each night. Hairs from Gombe were collected by following known individuals, watching them build nests, and then returning at dawn to search the abandoned nest. Genetic characterization of this community is the first step in understanding the diversity and similarities among chimpanzee communities, populations, and subspecies. Our study is necessarily limited by the logistical difficulties of intensive sample collection across Africa, but ongoing and planned studies at other sites will elucidate the constants and variables in chimpanzee population genetic structure.

DNA was amplified by the polymerase chain reaction (PCR) from hair samples collected from animals at 20 sites across Africa (Fig. 1). In all, we genotyped 67 individuals at eight variable nuclear SSR loci (*8*) and determined DNA sequence variation for many individuals at two informative mtDNA loci (*5*). The highly polymorphic di- and tri- and tetra-nucleotide SSR loci differ in the number of times the core sequence is repeated (typically 10 to 30 times) (*5–7*) and such simple sequence length polymorphism (SSLP) was scored on autoradiographs of polyacrylamide gels (*8*). Individual multilocus genotypes were used to establish pedigree relationships and characterize a population's relative genetic variability and behaviorally important substructuring. The mtDNA sequences were amplified and double-stranded products were directly sequenced to detect phylogenetically informative genetic variation. A 178-bp (base pair) segment of the cytochrome b (cyt b) region of 34 chimpanzees and one human was used to examine the deeper branches of chimpanzee phylogeny, and a 345-bp segment of the more variable control region of 66 chimpanzees, two bonobos (*Pan paniscus*), and one human was used to characterize phylogeo-

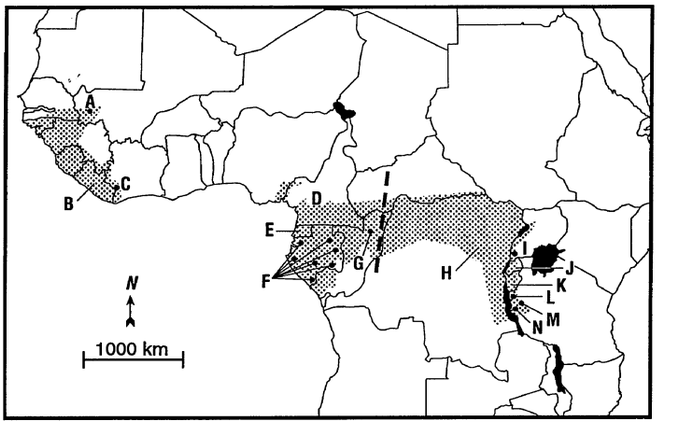

Fig. 1. Map of equatorial Africa, with sample collection sites or countries indicated by letters A to N, and approximate modern distributions of chimpanzees indicated by shading (*2, 12*). The approximate position of the poorly defined subspecies boundaries in central Africa is indicated by the dashed line.

P. A. Morin and D. S. Woodruff are with the Department of Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, CA 92093–0116, USA. J. J. Moore is with the Department of Anthropology, University of California, San Diego, La Jolla, CA 92093, USA. R. Chakraborty and L. Jin are at the Center for Demographic and Population Genetics, University of Texas Graduate School of Biomedical Sciences, P.O. Box 20334, Houston, TX 77225, USA. J. Goodall can be reached at Jane Goodall Institute–USA, P.O. Box 599, Ridgefield, CT 06877, USA.

*To whom correspondence should be addressed. The present address of P. A. Morin is Department of Anthropology, University of California, Davis, CA 95616–8522, USA.

graphic patterns, population genetic structure, and historical gene flow patterns both locally and regionally (9).

We present and discuss our results in a hierarchical fashion beginning with individuals and proceeding to populations, subspecies, and species. We have previously demonstrated the utility of hypervariable SSR variation for unambiguous paternity exclusion and the establishment of formal pedigree relationships among captive chimpanzees (5). In the free-ranging Kasakela community, with 19 matrilines in 1991, we have established probable paternity in the cases of *Faustino* and *Sherehe* (whose mothers conceived during episodes of promiscuous mating with nine or ten local males) (8). This noninvasive multilocus SSR genotyping method provides results that are genetically interpretable, replicable, and comparable across gels, samples, laboratories, and closely related taxa. With the addition of more loci and individuals sampled (such as tissues of deceased individuals), it should be possible to further reconstruct the Gombe pedigree and test hypotheses about chimpanzee behavior on the basis of more than 30 years of field observations.

Social Structure and Kin Selection

Hypotheses that kin selection has influenced the evolution of chimpanzee social structure predict that degrees of relatedness among males, who are philopatric, and females, who usually disperse at adolescence, should be significantly different (*10*). To test this, we compared the distributions of allele sharing at all scored loci between pairs of males and pairs of females from the Gombe community. The mean number of alleles per locus (± standard error) shared between males (0.854 ± 0.203) is significantly higher (*t* test, $P < 0.05$) than that between females (0.696 ± 0.240), indicating that males are indeed more related to one another than are females. To examine the degree of relatedness indicated by allele-sharing values (summed over all eight loci; Fig. 2) within each sex, we first performed a permutation test for resampling of allele-sharing values among males, females, known parent-offspring pairs, maternal sibs (assumed to be half-sibs), and maternally unrelated females. The results suggest that males are related at the level of half-sibs because the allele-sharing values fell below

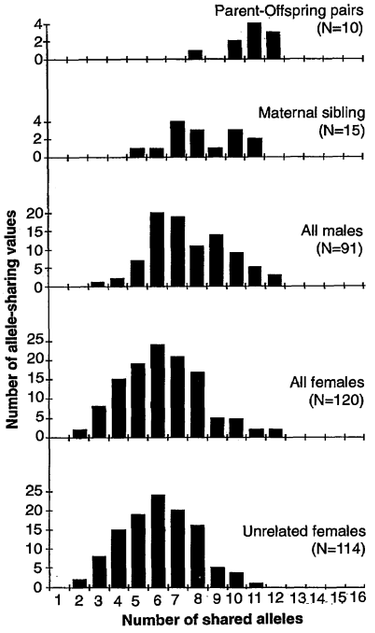

Fig. 2. Allele sharing distributions for all eight scored loci in subgroups of the Gombe Kasakela community. *N*, the number of pairwise allele-sharing values. The "unrelated female" group excludes maternal siblings and parent-offspring pairs. This analysis was used for the data on individuals of the Gombe community because information on kinship and gender were available only for this population.

Table 1. Allele frequencies (%) and standard errors for eight loci in the Gombe and non-Gombe sample sets. Allele frequencies were calculated for all individuals in the Gombe community (Total), as well as for all presumed unrelated individuals (Unrelated); (n), the number of chromosomes sampled. Two individuals (Mitumba 1 and 2, from the adjacent community at Gombe) were not included in the calculations because of their unknown genetic relationship to the Gombe Kasakela community. Since all loci are codominant, the allele frequencies were computed by the gene count method (*45*), and their standard errors were calculated as the square root of the variance of the corresponding binomial proportions (*46*).

Allele	Population			Allele	Population		
	Gombe		Non-Gombe		Gombe		Non-Gombe
	Unrelated	Total			Unrelated	Total	
	Mfd-3				*FABP*		
1			3.8 ± 2.7	1			3.7 ± 2.6
3			1.9 ± 1.9	3	22.5 ± 6.6	13.9 ± 4.1	14.8 ± 4.8
4			13.5 ± 4.7	4	50.0 ± 7.9	50.0 ± 5.9	33.3 ± 6.4
5	60.0 ± 7.7	54.2 ± 5.9	28.8 ± 6.3	5	7.5 ± 4.2	5.6 ± 2.7	16.7 ± 5.1
6	10.0 ± 4.7	6.9 ± 3.0	23.1 ± 5.8	6	7.5 ± 4.2	6.9 ± 3.0	20.4 ± 5.5
7			7.7 ± 3.7	7	12.5 ± 5.2	23.7 ± 5.0	11.1 ± 4.3
8		1.4 ± 1.4	3.8 ± 2.7	(n)	40	72	54
9	25.0 ± 6.8	34.7 ± 5.6	7.7 ± 3.7		*Mfd-32*		
10	5.0 ± 3.4	2.8 ± 1.9	1.9 ± 1.9	1		1.4 ± 1.4	2.0 ± 2.0
11			3.8 ± 2.7	4	5.3 ± 3.6	5.7 ± 2.8	8.0 ± 3.8
12			1.9 ± 1.9	5	21.1 ± 6.6	28.6 ± 5.4	2.0 ± 2.0
16			1.9 ± 1.9	6	18.4 ± 6.3	11.4 ± 3.8	4.0 ± 2.8
(n)	40	72	52	8			6.0 ± 3.4
	Mfd-18			9	55.3 ± 8.1	51.4 ± 6.0	12.0 ± 5.0
1	27.5 ± 7.1	25.0 ± 5.1	21.4 ± 5.5	10			24.0 ± 6.0
4	30.0 ± 7.2	30.3 ± 5.4	17.9 ± 5.1	11			12.0 ± 5.0
5	2.5 ± 2.5	1.4 ± 1.4		12		1.4 ± 1.4	18.0 ± 5.4
6	12.5 ± 5.3	12.5 ± 3.9	19.6 ± 5.3	13			6.0 ± 3.4
7	2.5 ± 2.5	1.4 ± 1.4	16.1 ± 4.9	15			4.0 ± 2.8
8			5.4 ± 3.0	16			2.0 ± 2.0
9			1.8 ± 1.8	(n)	38	70	50
10	12.5 ± 5.3	22.2 ± 4.9	8.9 ± 3.8		*Pla2a*		
11	12.5 ± 5.3	6.9 ± 3.0	5.4 ± 3.0	1			3.8 ± 2.7
12			3.6 ± 2.5	3			15.4 ± 5.0
(n)	40	72	56	4			23.1 ± 5.8
	Mfd-23			5	2.6 ± 2.6	2.9 ± 2.0	5.8 ± 3.2
1	16.7 ± 6.2	17.6 ± 4.6	18.5 ± 5.3	6	13.2 ± 5.5	12.9 ± 4.0	13.5 ± 4.7
3			3.7 ± 2.6	7	60.5 ± 7.9	64.3 ± 5.7	11.5 ± 4.4
4	16.7 ± 6.2	17.6 ± 4.6	1.9 ± 1.8	8	2.6 ± 2.6	1.4 ± 1.4	17.3 ± 5.2
5			5.6 ± 3.1	9	21.1 ± 6.6	18.6 ± 4.6	7.7 ± 3.7
7			5.6 ± 3.1	11			1.9 ± 1.9
9			20.4 ± 5.5	(n)	38	70	52
10			7.4 ± 3.6		*LL*		
12			7.4 ± 3.6	1			3.8 ± 2.7
13			7.4 ± 3.6	2			7.7 ± 3.7
14	2.8 ± 2.7	1.5 ± 1.5	7.4 ± 3.6	3	26.3 ± 7.1	32.9 ± 5.6	7.7 ± 3.7
15			3.7 ± 2.6	4	39.5 ± 7.9	32.9 ± 5.6	21.2 ± 5.7
19			5.6 ± 3.1	5			13.5 ± 4.7
20	5.6 ± 3.8	8.8 ± 3.4	3.7 ± 2.6	6		1.4 ± 1.4	5.8 ± 3.2
21	8.3 ± 4.6	5.9 ± 2.9	1.9 ± 1.8	7	21.1 ± 6.6	21.4 ± 4.9	9.6 ± 4.1
22	22.2 ± 6.9	26.5 ± 5.4		8	2.6 ± 2.6	1.4 ± 1.4	1.9 ± 1.9
24	22.2 ± 6.9	17.6 ± 4.6		9			11.5 ± 4.4
26	5.6 ± 3.8	4.4 ± 2.5		10			9.6 ± 4.1
(n)	36	68	54	11		2.9 ± 2.0	3.8 ± 2.7
	Rena4			12	5.3 ± 3.6	2.9 ± 2.0	1.9 ± 1.9
1			1.9 ± 1.8	13	5.3 ± 3.6	4.3 ± 2.4	1.9 ± 1.9
2	41.2 ± 8.4	39.1 ± 6.1	9.3 ± 3.9	(n)	38	70	52
3	58.8 ± 8.4	60.9 ± 6.1	87.0 ± 4.6				
4			1.9 ± 1.8				
(n)	34	64	54				

those of half-sibs only 50% of the time (expected if the distributions are the same), whereas female allele-sharing values were lower 65% of the time. Furthermore, male values were lower than female values only 26% of the time. Finally, a Kolmogorov-Smirnov one-tailed test for goodness of fit among distributions of allele-sharing values of non-sib, non–parent-offspring pairs of males compared to all maternal sib values was not significant, indicating that the distributions are not significantly different. In contrast, the corresponding test for female compared to maternal sib values was significant ($P \leq 0.05$).

These genetic data provide formal genetic evidence that the evolution of the sexual differences in social behavior in chimpanzees may be partially explained by kin selection theory. The hypothesis that the more closely related males form a kin group that cooperates to defend a territory, thereby increasing access to females and resources, can now be tested more rigorously and at other sites. The relation of allele-sharing to kinship coefficients can be determined more thoroughly as more data on pedigree relationships accrue, allowing quantitative application of kin selection theory to evolutionary studies of wild chimpanzees.

Allele frequencies and their standard errors were calculated for each of the eight loci in the Gombe Kasakela community (separately for unrelated individuals and for all individuals sampled) and the pan-African (non-Gombe) sample (Table 1). Kolmogorov-Smirnov tests for the significance of the difference between allele frequencies in two samples were performed on the Gombe and pan-African samples to determine whether allele frequencies were similar in the local population and continent-wide sample; significant differences ($P < 0.05$) were detected at two loci, *Mfd-23* and *Pla2a*. This indicates that the pan-African sample, at least at two loci, does not adequately represent the population of which the Gombe community is part. Subdivision of the species into three genetically distinct groups (subspecies) and population substructuring into social communities (see below) are most likely the reasons for the observed allele distribution differences. Comparisons between the data sets, however, reveal relative patterns that indicate the probable causes of substructuring.

If female dispersal was high enough to maintain a large and effectively panmictic population, we would expect pooled samples from across the species range and from individuals within a single social community to be in Hardy-Weinberg equilibrium (HWE). Deviations from the predictions of panmixia can be caused by several factors including nonrandom mating, distortions in the contributions of gametes to offspring, nonrandom sampling, Wahlund effect, strong selection, and allele mis-scoring [of which the last five are improbable in this case (*11*)]. Because these eight loci have not shown significant deviations from expectations of selective neutrality in humans (*7*), they are unlikely to cause distortions in the contributions of gametes to offspring or to deviate from expectations of panmixia because of strong selection in chimpanzees. Likewise, sampling methods were random with regard to these loci. Methods and accuracy of allele scoring have been described in (*8*). The assumptions of HWE in the Gombe sample are not satisfied according to the likelihood-ratio test at three loci (*Mfd3*, *FABP*, *LL*); the chi-square test corroborated these results in all but one case, that of the unrelated individuals at locus *Mfd3* (Table 2). Because the pan-African SSR samples comprise a collection of individuals from 15 sites across Africa, we would expect a larger number of significant departures from HWE in the pan-African sample if the departures were due only to population substructure (Wahlund's principle). The fact that there are two significant departures from HWE in the pan-African sample (loci *Mfd3* and *Mfd23*; likelihood-

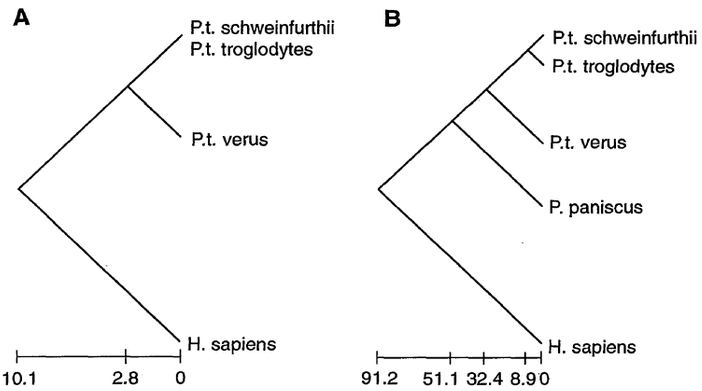

Fig. 3. Phylogenetic trees for chimpanzees and bonobos. (**A**) Based on weighted genetic distances corrected for intraspecific polymorphism for the cytochrome b locus (*23*). There were four synapomorphic transitions between *P. t. verus* and the other two subspecies at sites corresponding to 14995, 15004, 15052, and 15100 in the human reference sequence (*44*), and three polymorphic sites in each subspecies (one site polymorphic in both *P. t. schweinfurthii* and *P. t. troglodytes*). (**B**) Based on weighted genetic distances corrected for intraspecific polymorphism for the control region locus (*23*) (Table 4).

Table 2. Comparison of observed heterozygote frequencies (Obs.) with HWE expectations (Exp.) based on (i) total frequency of heterozygotes, in which all heterozygotes were used to calculate the expected allele frequencies for each locus, which were compared to the observed allele frequencies by chi-square analysis, and (ii) the likelihood ratio test criterion (−2 ln L), computed following (*47*), but the empirical significance was evaluated by permuting the allelic labels across individuals for each locus-population combination by the procedure described in (*48*). This permutation test is mathematically equivalent to the Monte Carlo method of exact test of HWE (*49*). All significance tests were conducted by permuting alleles across individuals and with 2000 replications. $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

Statistics	Population: Gombe Unrelated	Population: Gombe Total	Population: Non-Gombe
	Mfd3		
N	20	36	26
Obs.	8	15*	19
Exp. ± SE	11.6 ± 2.2	21.2 ± 3.0	21.9 ± 1.9
-2 ln L	21.0**	23.7**	48.5*
	Mfd18		
N	20	36	28
Obs.	14	27	23
Exp. ± SE	16.1 ± 1.8	28.3 ± 2.5	24.0 ± 1.8
-2 ln L	16.7	18.4	33.9
	Mfd23		
N	18	34	27
Obs.	11**	25	19**
Exp. ± SE	15.4 ± 1.5	28.4 ± 2.2	24.4 ± 1.5
-2 ln L	25.3	23.0	70.1**
	Mfd32		
N	19	35	25
Obs.	12	25	23
Exp. ± SE	12.0 ± 2.1	22.6 ± 2.8	22.0 ± 1.6
-2 ln L	3.2	9.0	40.2

Statistics	Population: Gombe Unrelated	Population: Gombe Total	Population: Non-Gombe
	FABP		
N	20	36	27
Obs.	5***	12***	22
Exp. ± SE	13.8 ± 2.1	24.4 ± 2.8	21.6 ± 2.1
-2 ln L	26.8***	31.2***	15.9
	Pla2a		
N	19	35	26
Obs.	10	16	21
Exp. ± SE	11.1 ± 2.1	19.0 ± 2.9	22.5 ± 1.7
-2 ln L	7.0	11.0	37.8
	Rena4		
N	17	32	27
Obs.	6	11	7
Exp. ± SE	8.5 ± 2.1	15.5 ± 2.8	6.4 ± 2.2
-2 ln L	1.3	2.5	1.0
	LL		
N	19	35	26
Obs.	9**	18**	22
Exp. ± SE	14.1 ± 1.9	26.1 ± 2.6	23.5 ± 1.5
-2 ln L	24.1**	34.7**	54.1

ratio test) compared to three in the Gombe sample suggests that they may be caused by local inbreeding or to a greater degree of relatedness in this social community, resulting from either single-sex dispersal among small communities or recent changes in dispersal behavior at Gombe, rather than prolonged genetic isolation of the population. In addition, all significant deviations in the Gombe sample are toward deficiencies in the number of heterozygotes, as expected if inbreeding or high male-male relatedness were the cause.

Nonrandom association of alleles from genetically unlinked loci (8) (deviations from gametic phase equilibrium) may be interpreted as evidence for either population substructure or selection. To examine this we used a test for independent segregation of alleles (Table 3). Because there were significant differences between observed and expected heterozygosities, the chosen test included both values for all loci. The analyses showed that observed variances of the number of heterozygous loci per individual were always within the 95% confidence limits, provided that there was independent segregation of alleles at different loci. Thus, despite the greater relatedness and homozygosity in the Kasakela community, significant cosegregation of alleles at these loci was not observed, and extensive inbreeding is not indicated.

The results showing patterns of deviation from the expectations of HWE, increased homozygosity, and no deviations from gametic phase equilibrium, together with the significant differences in the levels of relatedness among males compared to that among females, strongly support the hypothesis that premating single-sex dispersal is the primary source of genetic substructuring in this community, rather than recent changes in dispersal behavior of Gombe females or genetic isolation of the Gombe community leading to recent increases in inbreeding levels. The tests do not, however, exclude these recent changes as contributing factors, so that continued genetic monitoring of this community and comparative studies are recommended. Evidence of similar genetic patterns in other communities would further support the hypothesis that kin selection has been a strong force in the evolution of chimpanzee social structure.

Gene Flow and Phylogenetics

Systematists recognize two species in the genus *Pan*, the bonobo or pygmy chimpanzee (*P. paniscus*) of Zaire and the chimpanzee (*P. troglodytes*). The latter has an unusually wide distribution across equatorial Africa and in the past has been variously assigned to as many as 17 species and 34 subspecies. The three well-known subspecies are: western masked or pale-faced *P. t. verus*, central black-faced *P. t. troglodytes*, and eastern long-haired *P. t. schweinfurthii* (*12*). Vernacular names and minor craniometric variation notwithstanding, these geographically defined allopatric subspecies cannot be distinguished morphologically in captivity (*13*), so most previous genetic studies were limited because they involved captive animals of unknown geographic origin and taxonomic identity (*14, 15*). We therefore conducted a genetic survey of mitochondrial variation throughout the range of *P. troglodytes*, widely sampling within and among the geographic ranges of the three subspecies.

The 178-bp cyt b sequence was highly conserved within subspecies and no fixed differences were found between *P. t. troglodytes* and *P. t. schweinfurthii*. In contrast, a 2.8% genetic distance was found to clearly separate western *P. t. verus* from the other two subspecies (Fig. 3A). This short segment was chosen simply because we could look for differences among subspecies while using a fairly well-characterized gene locus. To look at patterns of intrasubspecific variation, however, and more clearly resolve the relationships between subspecies, we used sequences of a longer segment of the control region. The 345-bp control region sequence was far more variable both within and among subspecies, having 124 (35.5%) variable positions, including two single base deletions. Within *P. t. schweinfurthii*, there were 25 haplotypes (distinctive sequences; missing data were excluded) among 37 individuals, and 6 haplotypes

Table 3. Test of association of alleles among loci from the distribution of heterozygous loci. Only 57 individuals tested for all eight loci are included in this test. The expected distributions of the number of heterozygous loci and the theoretical mean, variance, and 95% confidence interval (CI) of variance are reported. In method (1) the observed proportions of heterozygotes for each locus (Table 2) were used for computations; in method (2) the unbiased estimates of heterozygosities (*50*) were used. Tests for nonrandom association of alleles among loci from the distribution of heterozygous loci were performed using the method described (*51*), following the theory of Brown *et al.* (*52*). An alternative method (*53*) provides equivalent results.

Number of heterozygous loci	Population								
	Gombe						Non-Gombe		
	Unrelated			Total					
	Obs.	Exp.		Obs.	Exp.		Obs.	Exp.	
		1	2		1	2		1	2
0	0	0.0	0.0	0	0.0	0.0	0	0.0	0.0
1	0	0.4	0.0	0	0.5	0.1	0	0.0	0.0
2	1	1.6	0.2	1	2.2	0.5	0	0.1	0.0
3	2	3.5	0.9	4	5.8	2.0	0	0.4	0.1
4	2	4.3	2.5	7	8.6	5.3	5	2.0	0.9
5	7	3.3	4.2	11	7.7	8.5	3	5.5	3.7
6	3	1.5	4.3	5	4.0	8.3	7	8.4	8.4
7	0	0.4	2.4	2	1.1	4.4	5	6.2	8.9
8	0	0.0	0.5	0	0.1	1.0	4	1.4	2.1
Total	15.0	15.0	15.0	30	30.0	30.1	24	24.0	24.1
Mean	4.600	3.946	5.390	4.700	4.265	5.302	6.000	5.904	6.295
Variance	1.400	1.835	1.654	1.459	1.775	1.684	1.913	1.254	0.989
95% CI of variance									
Lower		0.593	0.524		0.922	0.871		0.545	0.400
Upper		3.077	2.783		2.629	2.497		1.963	1.577

were shared among 2 to 6 individuals. Within *P. t. troglodytes*, there were 17 haplotypes among 18 individuals, and within *P. t. verus*, all 11 sequences were different from each other; no haplotypes were shared between subspecies. Several synapomorphies (shared derived characters) were fixed or nearly fixed in each subspecies (Fig. 4). Despite the large variation within subspecies, cladistic analysis shows that 10 fixed (and 2 slightly polymorphic) synapomorphies link the eastern and central subspecies to the exclusion of the western subspecies, which in turn exhibits 11 fixed and 1 polymorphic (in one individual) synapomorphies that bind all *P. t. verus* to the exclusion of the others (Fig. 4). Using the neighbor joining method, a phylogenetic tree was constructed from the pairwise distances (Fig. 5; summarized in Fig. 3B). Alternative parsimony and distance cluster analyses also yielded trees with the same subspecies clades and very similar subclusters within subspecies (*16, 17*).

Of 24 adults in the Kasakela social community at Gombe in July 1991, 19 individuals known or suspected to be maternally unrelated were sampled for sequencing; we found 15 different haplotypes in the control region. This surprisingly high within-community genetic variation is attributable to gene flow based on female dispersal behavior in this species (*18*). The sequence similarities of some individuals from Gombe to individuals from a geographic range covering over 600 km suggests that historical gene flow has been significant. In the last 50 years, however, the 32-km^2 Gombe National Park, now occupied by 150 to 160 chimpanzees, has become completely isolated from other habitats by deforestation (*19*). In the absence of gene flow with other populations, genetic erosion due to genetic drift and inbreeding may begin to significantly increase the probability of extirpation of this and other populations experiencing genetic isolation and habitat frag-

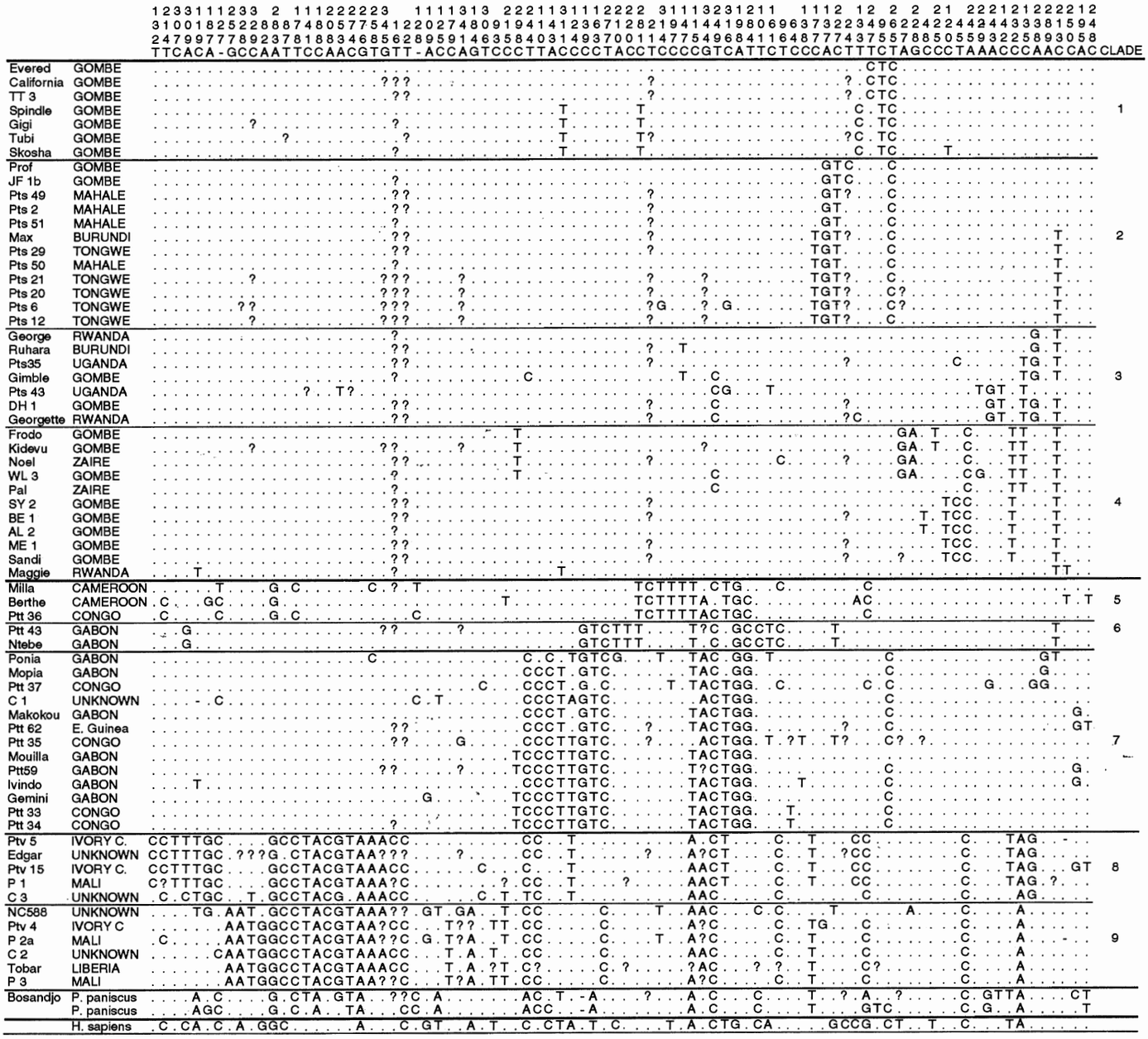

Fig. 4. Individual chimpanzee by control region nucleotide site matrix for nucleotide positions variable in two or more *P. troglodytes* individuals. Nucleotide sites have been rearranged to cluster markers for each clade (clades and subspecies separations match those in Fig. 5, described in text); sites 1 to 345 refer to the human reference sequence (*44*), with 1 corresponding to position 16047; ? = missing data; reference = chimpanzee consensus sequence.

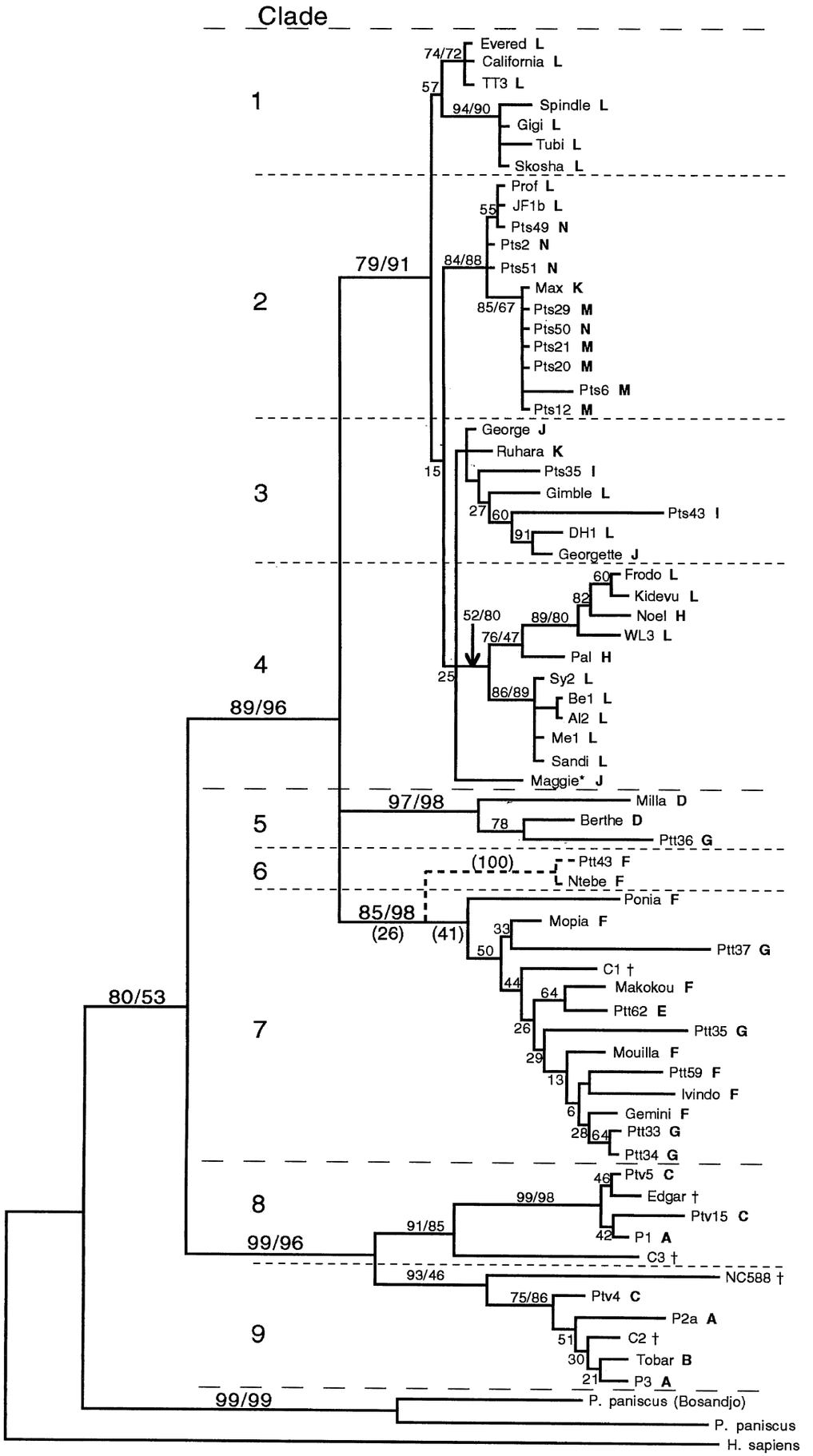

Fig. 5. Phylogenetic tree constructed using the neighbor joining method based on genetic distances estimated from control region sequences. Genetic distances were computed using the standard Kimura two-parameter model in MEGA software (*16*). Sample origins are indicated by letters A to N after sample identities, and correspond to letters on the map (Fig. 1). Clades referred to in the text are numbered 1 to 9; clades 1 to 4 are *P. t. schweinfurthii*; 5 to 7 are *P. t. troglodytes*; 8 and 9 are *P. t. verus*. The Ntebe-Ptt43 sequences are substantially different from other sequences of that subspecies, and were found in two Gabonese individuals separated by 400 km. Their branch position on the tree is indicated by a dashed line, and inclusion of these sequences in bootstrap analyses yielded the values in parentheses for clades 6 and 7 (others were not significantly affected). The reliability of each interior branch was tested by 500 bootstrap replications and a standard error test in the MEGA software package. Where single numbers occur on branches, they represent the bootstrap confidence levels (BCL); standard error confidence probabilities (CP) are shown on some branches in the format BCL/CP. Sites containing missing data and alignment gaps were not used in the bootstrap and standard error tests. *Maggie was included in clade 4 for simplicity because the branch was not significantly supported. †Samples of unknown geographic origin.

mentation (*20*). In that our data on local variation in nuclear genes (and field observations) provide no evidence for inbreeding depression in this community yet, we note that habitat loss and exposure to human-borne diseases are still more serious threats. The extent to which population isolation has or will affect other aspects of community structure and behavior remains to be investigated.

In the eastern subspecies, *P. t. schweinfurthii*, the samples include 32 individuals of precisely known geographic origin and 7 of known country of origin. Both parsimony and neighbor joining (Fig. 5) methods created trees with four primary clades (clades 1 to 4) within this subspecies, each with identical individual composition. There is some geographic differentiation; clade 3 contains primarily samples from the northernmost sites (Uganda, Rwanda, and Burundi) and clade 2 includes all nine apes from the two southernmost sites (Mahale and Tongwe). Widespread historical gene flow is evidenced, however, by inclusion of the geographically intermediate Gombe genotypes in all four clades. The absence of any of the southernmost samples in "northern" clade 3, and of northernmost samples in "southern" clade 2, indicates possible isolation by distance beyond about 600 km.

Control region variation provides evidence for similar high historical gene flow patterns in the other two subspecies. Within the central subspecies, *P. t. troglodytes*, we can show three clades (clades 5 to 7)

Table 4. Variation among individuals within and between subspecies for the control region sequence. On and below diagonal: mean weighted genetic distances corrected for intrasubspecific polymorphism (*23*) and mean number of transitions (ts) and transversions (tv), not corrected for missing data (ts/tv); above diagonal: time since divergence (myr + 95% CI).

Taxon	All Pt (%)	Pts (%)	Ptt (%)	Ptv (%)	Pp (%)	Human (%)
All *P. troglodytes*	20.19	NA	NA	NA	2.5 (2.0–3.0)†	4.7 (4.2–5.2)*
P. t. schweinfurthii	NA	4.0 (7/0.9)	0.44 (.21–1.64)	1.58 (.75–5.96)	NA	NA
P. t. troglodytes	NA	8.9 (17.5/3.2)	9.0 (13.4/2)	1.58 (.75–5.96)	NA	NA
P. t. verus	NA	32.4† (27/7)	32.4† (29/9.8)	6.3 (14.7/1.3)	NA	NA
P. paniscus	51.1 (29.8/14.7)	57.4 (NA)	62.3 (NA)	54.3 (NA)	26.6 (29.7/15)	NA
Human	91.2 (34.8/20.9)	99.3 (NA)	100.7 (NA)	94.3 (NA)	85.2 (38.5/20)	2.9‡ (NA)

*Data from (*26*). †Mean weighted distance to *P. t. schweinfurthii* and *P. t. troglodytes* combined, corrected for intrasubspecific polymorphism. ‡Data from (*24*).

involving 18 individuals from known localities in Gabon, Equatorial Guinea, Cameroon, and Congo (Fig. 5). Clade 7 comprises individuals from Gabon, Equatorial Guinea, and Congo which cluster together without any detectable geographic substructure over 800 km. Clade 5 was observed in the northern part of the subspecies range, but sympatric with members of clade 7. Clade 6 includes two individuals separated by approximately 400 km within Gabon. Additional variation in this subspecies is to be expected because our present geographic sampling in the north and east of its range is inadequate.

The third subspecies, *P. t. verus*, from western Africa is the most vulnerable today (*21*) and has the smallest geographic range. Samples from chimpanzees from Mali and Ivory Coast, separated by up to 900 km, shared nearly identical sequences from the two widely divergent mtDNA haplotype groups (clades 8 and 9) indicating, again, a historical pattern of long-distance gene flow. The trees based on these control region data (Figs 3B and 5) indicate that the earliest divergence was between the western African *P. troglodytes* clades and the lineage leading to the other two subspecies. The branch (and corrected genetic distances) (Table 4) connecting *P. t. verus* to the other two subspecies is relatively deep, suggesting a long separation between these two groups. In addition, the other two subspecies show substantially less divergence and less intrasubspecific polymorphism that has arisen since subspecies divergence; the relatively deep split within *P. t. troglodytes* may represent ancestral polymorphism (as shown in Fig. 5, clade 5), and accounts for the higher variation within *P. t. troglodytes* than within the western subspecies. The presence of this ancestral polymorphism, as well as relatively less variation within clades of *P. t. troglodytes* than within *P. t. verus*, supports a longer isolation of *P. t. verus* as a subspecies.

These patterns of substantial genetic variability indicate that populations have been sufficiently large since divergence of each subspecies to allow accumulation and maintenance of genetic variation. The observation that the variation is dispersed throughout the range of each subspecies further indicates that there has been substantial historical gene flow. These data are not consistent with a recent expansion of any subspecies populations, or with a rapid expansion followed by population isolation. Studies of larger samples from other populations within each subspecies range should allow quantitative evaluation of historical gene flow rates.

Comparisons of these cyt b and control region sequences provide insight into the evolutionary relations and ages of these taxa (*22*). At the cyt b locus, the corrected genetic distance (*23–25*) between *Homo* and *Pan* is 10.1%. A 4.7 ± 0.5 million year [myr ± 95% confidence interval (CI)] estimated divergence time between humans and chimpanzees (*26, 27*), yields a substitution rate of 2.15% myr^{-1} (95% CI range, 0 to 6.37% myr^{-1}). This calibration gives a divergence time based on cyt b between *P. t. verus* and the two other subspecies (corrected d = 2.8%) of about 1.3 myr (95% CI range: 0.44 to >2.5 myr) (*28*). Calibration by means of an earlier human-ape divergence [for example, 8 myr (*29*)], would yield a correspondingly earlier split within *P. troglodytes* (2.2 myr, roughly contemporaneous with the origin of the genus *Homo*).

To calibrate the control region data, we used estimates of divergence time between *P. paniscus* and *P. troglodytes* of 2.5 ± 0.5 myr ago, based on previous calibrations of this and other, more conserved, regions of the mitochondrial genome (*26, 30, 31*). According to Nei (*28, 32*), the substitution rate is then 20.4% (5.43 to 42.96%) myr^{-1} and the derived divergence times between the subspecies are given in Table 4. The calculated control region based divergence time between *P. t. verus* and the two other subspecies is 1.58 myr. *Pan paniscus* is seen to lie closer to the *Pan-Homo* split as postulated by Zihlman and others (*33*). Because gene trees may not reflect correct species trees and because standard errors associated with relatively short DNA sequences are large, further discussion of the bearing of these results on the Pliocene-Pleistocene speciation of these apes should await phylogenetic analysis of more loci (*34*).

These data indicate that at the sequences examined *P. t. verus* is more differentiated from the other two subspecies than are some other full species of mammals (*35*). The average genetic distance between the western subspecies and the other two is two-thirds the genetic distance between the two chimpanzee species *P. paniscus* and *P. troglodytes*. This degree of differentiation within *P. troglodytes* is greater than that reported for the cytochrome oxidase II sequence (*36*). The probable isolation of *P. t. verus* for about 1.6 myr is greater than the estimated mean duration of late Cenozoic mammal species (*37*). In our opinion, if this result is confirmed at three or more unlinked nuclear loci (*34*), and supported by eco-behavioral data, *P. t. verus* merits elevation to full species rank: *P. verus* (Schwartz) 1934.

Implications

Our demonstration that free-ranging chimpanzees can be genotyped noninvasively has resulted in four significant findings. First, our observation that, within a community, males are more closely related to one another than are females, and that this difference can account for genetic substructuring of the population, constitutes a substantial preliminary test of a sociobiological hypothesis concerning the evolution of cooperative behavior in males. Second, our discovery that dispersal, as evidenced by maternally transmitted genotypes, is detectable over distances of 600 to 900 km within each subspecies provides evidence that supports the hypothesis that high amounts of gene flow may account for the low amounts of morphological differentiation in chimpanzees (*38*). Third, our discovery that the West African *P. t. verus* is actually a well-differentiated and independently evolving taxon has important implications for both evolutionary and conservation biology. Although exact climatic or ecological (or both) patterns of change in equatorial Africa over the past 2 million years are unclear, repeated cycles of drying possibly associated with hypothermal periods occurred and periodically reduced forests to western, central, and eastern refugia. Extant chimpanzee taxa may have differentiated from

one another in such forest fragments (*39*). Reconstructing the history of the genus *Pan* and of our common ancestor will have to allow for the previously underappreciated variation within chimpanzees. This may help to resolve the gorilla-chimpanzee-human "trichotomy" question (*40*) by improving the resolution of synapomorphies over homoplasy (the phylogenetic "signal"). Fourth, our discovery of considerable genetic variation within chimpanzee populations across broad geographic ranges provides us with a way of quantifying genetic diversity in these endangered species and monitoring genetic erosion in both wild and captive populations.

Conservation efforts in the wild and in captivity need to account for the genetic diversity in the genus *Pan* and its partitioning into several well-differentiated evolutionarily significant units (*41*). Although captive chimpanzees have traditionally been sorted into and managed as two species, the three subspecies of *P. troglodytes* have, for example, been afforded no special significance under current NIH colony management policies. Our results suggest that treating the >2000 chimpanzees in biomedical facilities in the United States as genetically equivalent requires reassessment. Inappropriately matched individuals are undesirable in both experimental and breeding situations, and some previous results based on genetically dissimilar apes may require reinterpretation (*42*). However, chimpanzee social, emotional, and intellectual similarities to humans render management from a purely genetic standpoint unethical; the behavioral or social (or both) needs of captive individuals must also be considered. We do not advocate breaking up mismatched but long-bonded pairs.

Finally, management of increasingly fragmented populations in Africa may require occasional translocation of individuals for their own protection and to preserve naturally occurring variability in the wild. Such interventive management is fraught with hazards including disease introduction and the behavioral problems of translocated wild individuals and reintroduced captive-born or raised individuals (*43*). Our results suggest that the possibility of outbreeding depression must also be considered. Our appreciation of the natural diversity within these species may have important implications for their future evolution and for their use as models in biomedical, paleoanthropological, and evolutionary studies of humans.

REFERENCES AND NOTES

1. D. S. Woodruff, *J. Sci. Soc. Thailand* **16**, 117 (1990); *Primates* **34**, 333 (1993); P. A. Morin, J. Wallis, J. J. Moore, R. Chakraborty, D. S. Woodruff, *ibid*, p. 347.
2. J. Goodall, *The Chimpanzees of Gombe* (Harvard Univ. Press, Cambridge, MA, 1986).
3. T. Nishida, *The Chimpanzees of the Mahale Mountains: Sexual and Life History Strategies* (Univ. of Tokyo Press, Tokyo, 1990); T. Kano, *The Last Ape: Pygmy Chimpanzee Behavior and Ecology* (Stanford Univ. Press, Stanford, CA, 1991).
4. Locality data from P. A. Morin [thesis, University of California, San Diego (1992)] are generalized in Figs. 4 and 5. Samples from non-Gombe named individuals and Ptt-59, Ptt-62 are from captive animals with inferred country of origin and were obtained from the Chimfunshi Wildlife Orphanage, Zambia; the Centre International de Recherches Medicales de Franceville (CIRMF), Gabon; and the Jane Goodall Institute, Burundi. It is possible that some of these captives have been mislocated within the subspecies' range.
5. The hair sample collection, DNA extraction and amplification, electrophoresis, and autoradiography procedures have been described: P. A. Morin and D. S. Woodruff, in *Paternity in Primates: Genetic Tests and Theories*, R. D. Martin, A. F. Dixson, E. J. Wickings, Eds. (Karger, Basel, Switzerland, 1992), pp. 63–81.
6. J. T. Epplen *et al.*, in *DNA Fingerprinting: Approaches and Applications*, T. Burke, G. Dolf, A. J. Jeffreys, R. Wolff, Eds. (Birkhauser, Basel, Switzerland, 1991), pp. 50–69.
7. A. Edwards, A. Civitello, H. A. Hammond, C. T. Caskey, *Am. J. Hum. Genet.* **49**, 746 (1991).
8. P. A. Morin, J. J. Moore, J. Wallis, D. S. Woodruff, *Mol. Ecol.*, in press. The eight loci studied were (locus:human chromosome): (dinucleotide repeats): *Mfd3:1, Mfd18:8, Mfd23:16, Mfd32:18, LL:15* [J. L. Weber and P. E. May, *Am. J. Hum. Genet.* **44**, 388 (1989); *Nuclei Acids Res.* **18**, 6465 (1990); J. L. Weber, A. E. Kwitek, P. E. May, *ibid.*, p. 4034; J. L. Weber, D. Patterson, H. Drabkin, *ibid.*, p. 4038; M. Litt and J. A. Luty, *Am. J. Hum. Genet.* **44**, 397 (1989)]; (trinucleotide repeats): *Pla2a:12, FABP:4*, and (tetranucleotide repeats): *Rena4:1* (*7*). Chromosomal locations of loci in chimpanzees are unknown. Genotypes for all 67 individuals and eight loci are available from P. A. Morin.
9. The section of the cyt b gene amplified corresponds to bases 14816 through 15174 of the human mitochondrial sequence (*44*); the sequenced portion corresponds to human positions 14964 to 15141. The primers are identified by a coded name, followed by the human nucleotide reference number of the 3′ end and L or H to indicate light or heavy strand: *Cytb1* (L14841) 5′-CCATCCAACATCTCAGCATGATGAAA-3′ and *Cytb2* (H15149) 5′-GCCCCTCAGAATGATATTTGTCCTCA-3′ (*15*; without restriction sites). The section of the control region amplified corresponds to bases 16019 through 16428 of the human sequence (*44*), *D-88* (L16041) 5′-CTCTGTTCTTTCATGGGGAAGC-3′ and *D-441* (H16407) 5′-CGGGATATTGATTTCACGGAGG-3′ (K. Garner, Zool. Soc. San Diego, personal communication). Sequencing was done from both strands of the control region PCR product as described [J. C. Garza and D. S. Woodruff, *Mol. Phylogenet. Evol.* **1**, 202 (1992)]. Sequencing of the cyt b region was only performed with the heavy strand (*Cytb2*) primer for most samples. Cyt b sequences: GenBank L35346–L35379; control region sequences; GenBank L35380–L35442. Five additional published sequences were included in the analyses: three *P. troglodytes* (C1, C2, C3), one *P. paniscus* (plus another sequenced by us, GenBank L35443), and one human [D. R. Foran, J. E. Hixson, W. M. Brown, *Nucleic Acids Res.* **16**, 5841 (1988); T. D. Kocher and A. C. Wilson, in *Evolution of Life*, S. Osawa and T. Honjo, Eds. (Springer, Tokyo, 1991), pp. 391–413; A. Di Rienzo and A. C. Wilson, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1597 (1991)]. All sequences were aligned by hand. All chimpanzee control regions had a 1-bp deletion at human position 16077 (*44*). Chimpanzee sequences differed from one another by deletions of 1 or 2 bp at sites 127 and 128 (Fig. 4); two individuals also had deletions at site 250 (Fig. 4).
10. A. E. Pusey, in *The Great Apes*, D. A. Hamburg and E. R. McCown, Eds. (Benjamin-Cummings, Menlo Park, CA, 1979), pp. 465–480; R. W. Wrangham, *ibid.*, pp. 481–489; in *Current Problems in Sociobiology*, King's College Sociobiology Group, Eds. (Cambridge Univ. Press, Cambridge, UK, 1982), pp. 269–289; P. A. Morin, *Yearb. Phys. Anthropol.* **36**, 179 (1993).
11. E. B. Spiess, *Genes in Populations* (Wiley, New York, 1989); K. C. Staub, D. S. Woodruff, E. S. Upatham, V. Viyanant, *Am. Malacol. Bull.* **7**, 93 (1990).
12. W. C. O. Hill, in *The Chimpanzee*, G. H. Bourne, Ed. (Karger, Basel, Switzerland, 1969), pp. 22–49; C. P. Groves, *A Theory of Human and Primate Evolution* (Clarendon, Oxford, 1989); R. M. Nowak, *Walker's Mammals of the World* (Johns Hopkins Univ. Press, Baltimore, 1991), vol. 1; D. E. Wilson and D. M. Reeder, Eds., *Mammal Species of the World*, (Smithsonian Institution, Washington, DC, ed. 2, 1993)
13. C. P. Groves, C. Westwood, B. T. Shea, *J. Hum. Evol.* **22**, 327 (1992); B. T. Shea, S. R. Leigh, C. P. Groves, in *Species, Species Concepts, and Primate Evolution*, W. H. Kimbel and L. B. Martin, Eds. (Plenum, New York, 1993), pp. 265–296; U. S. Seal and N. R. Flessness, in *Primates, The Road to Self-Sustaining Populations*, K. Benirschke, Ed. (Springer, New York, 1986), pp. 47–55.
14. N. Saitou [*Am. J. Phys. Anthropol.* **84**, 75 (1991)] reviewed published interspecific and intergeneric comparisons involving DNA sequence data. Other reports involving chimpanzees of unknown identity include: B. B. Normark, A. R. McCune, R. G. Harrison, *Mol. Biol. Evol.* **8**, 819 (1991); M. Stoneking *et al.*, *Am. J. Hum. Genet.* **48**, 370 (1991); G. Ruano, J. Rogers, A. C. Ferguson-Smith, K. K. Kidd, *Mol. Biol. Evol.* **9**, 575 (1992); see also (*15, 30*).
15. T. D. Kocher *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6196 (1989).
16. Pairwise distances for neighbor-joining analysis [N. Saitou and M. Nei, *Mol. Biol. Evol.* **4**, 406 (1987)] were calculated by means of the Kimura two-parameter model [M. Kimura, *J. Mol. Evol.* **16**, 111 (1980)] in MEGA [S. Kumar, K. Tamura, M. Nei, "MEGA: Molecular Evolutionary Genetics Analysis Software for Microcomputers," *Comput. Appl. Biosci.* **10**, 189 (1994)].
17. Cladistic analysis was performed with the PAUP parsimony program [D. L. Swofford, PAUP: Phylogenetic Inference Using Parsimony (version 3.1.1) Illinois Nat. Hist. Survey, Champaign (1991)] with a heuristic search method, 20 replications with random addition of sequences, and 100 trees saved from each replicate (all other search variables were default settings; Consistency Index = 0.537, Rescaled Consistency Index = 0.440, Retention Index = 0.819, Homoplasy Index = 0.463, number of steps = 337). The transition/transversion ratio (*R*) of 15 (control region) was used.
18. T. Nishida, in *The Great Apes*, D. A. Hamburg and E. R. McCown, Eds. (Benjamin-Cummings, Menlo Park, CA, 1979), pp. 73–121; A. E. Pusey, in (*10*); *Anim. Behav.* **28**, 543 (1980); P. A. Morin *et al.*, in (*1*).
19. J. Goodall, in *Understanding Chimpanzees*, P. G. Heltne and L. A. Marquardt, Eds. (Harvard Univ. Press, Cambridge, MA, 1990), pp. 360–361.
20. M. Lynch and W. Gabriel, *Evolution* **44**, 1725 (1990); D. S. Woodruff, in *Proceedings of the International Conference on Conservation Tropical Biodiversity*, Y. S. Kheong and L. S. Win, Eds. (Malay Nature Soc., Kuala Lumpur, Malaysia, 1992), pp. 258–272; *Proceedings of the Second Princess Chulabhorn Congress* (Bangkok, in press).
21. Captive Breeding Specialist Group, IUCN, SSC, *Global Captive Action Plan for Primates with Supplement* (CBSG, Apple Valley, MN, 1991).
22. Genetic distances and divergence times can also be based on SSR allele trees if the loci studied are selected at random and homoplasy can be detected and excluded; neither assumption can be met here.
23. Genetic distances have been corrected for intrasubspecific polymorphism according to M. Nei [*Molecular Evolutionary Genetics* (Columbia Univ., New York, 1987), p. 276], and for multiple hits and transition bias according to the methods [W. M. Brown, E. M. Prager, A. Wang, A. C. Wilson, *J. Mol. Evol.* **18**, 225 (1982)] for the cyt b sequence, because of the lack of transversions within *P. troglodytes*, we used a transition to transversion ratio (*R*) of 8 (*25*), and the transversion method [R. Higuchi, B. Bowman, M. Freiberger, O. A. Ryder, A. C. Wilson, *Nature* **312**, 282 (1984); (*24, 28*)] with $R = 15$ for the control

region sequences (*24*). We calculated control region *R*'s from all of the chimpanzee data and from each subspecies separately by counting transversions and transitions on the minimum length trees found by parsimony analysis (*24*) (*R* ± 95% CI; all *P. troglodytes* 8.5% ± 3.52, *P. t. schweinfurthii* 8.6 ± 6.84, *P. t. troglodytes* 10.1 ± 7.06, *P. t. verus* 18.0 ± 21.36); we found that, although they are quite variable, these values are not significantly different from Vigilant *et al.*'s estimate. Divergence rates were determined (*24, 28*). Further, the estimate of *R* = 15 is probably low [J. Wakeley, *Mol. Biol. Evol.* **11**, 436 (1994); M. Ruvolo *et al.*, *ibid.* **10**, 1115 (1993); C. Wills, *Evolution*, in press] and results in a more recent estimate of divergence than if higher *R*'s were used in distance calculations.

24. L. Vigilant, M. Stoneking, H. Harpending, K. Hawkes, A. C. Wilson, *Science* **253**, 1503 (1991).
25. M. F. Smith and J. L. Patton, *Mol. Biol. Evol.* **8**, 85 (1991).
26. S. Horai *et al.*, *J. Mol. Evol.* **35**, 32 (1992).
27. A. C. Wilson and V. M. Sarich, *Science* **158**, 1200 (1967); M. Hasegawa, H. Kishino, K. Hayasaka, S. Horai, *J. Mol. Evol.* **31**, 113 (1990); A. C. Wilson, H. Ochman, E. M. Prager, *Trends Genet.* **3**, 241 (1987); A. C. Wilson *et al.*, *Biol. J. Linn. Soc.* **26**, 375 (1985).
28. M. Nei, *Mol. Biol. Evol.* **9**, 1176 (1992).
29. Y. Coppens, *Sci. Am.* **270**, 88 (May 1994).
30. S. D. Ferris *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 6319 (1981); A. C. Wilson *et al.*, *Biol. J. Linn. Soc.* **26**, 375 (1985).
31. G. Pesole, E. Sbis, G. Preparata, C. Saccone, *Mol. Biol. Evol.* **9**, 587 (1992).
32. K. Tamura and M. Nei, *ibid.* **10**, 512 (1993): an alternative approach based on three documented chimpanzee control region sequences (C1, C2, C3) (*15*) involves greater mathematical precision. Based on our having established (Fig. 5 [*42*)] that these published sequences represent one *P. t. troglodytes* and two *P. t. verus*, their data suggest that the divergence time of *verus* is 0.9 myr (0.57 to 2.72). Our higher estimate is a result of having studied 63 additional chimpanzee sequences.
33. A. L. Zihlman, J. E. Cronin, D. L. Cramer, V. M. Sarich, *Nature* **275**, 744 (1978).
34. C. I. Wu, *Genetics* **127**, 429 (1991).
35. Partially or fully comparable sequence data exist for very few other congeneric mammals. At the control region the genetic distances between *P. t. verus* and *P. t. schweinfurthii*/*P. t. troglodytes* are comparable to those found between some species of cetaceans [A. R. Hoelzel, J. M. Hancock, G. A. Dover, *Mol. Biol. Evol.* **8**, 475 (1991)]. At the cyt b sequence, the observed distance (2.8%) falls in the low range for different congeneric species of akodontine rodents (*25*) and bat species in the genus *Dermanura* [R. A. Van Den Bussche, R. J. Baker, H. A. Wichman, M. J. Hamilton, *Mol. Biol. Evol.* **10**, 944 (1994)], but greater than that reported between Arctic and kit foxes, which are currently placed in different genera [E. Geffen *et al.*, *J. Zool. London* **228**, 27 (1992); A. Mercure, K. Ralls, K. P. Koepfli, R. K. Wayne, *Evolution* **47**, 1313 (1993)]. Interspecific cyt b distances among *Phyllostomus* bats [R. A. Van den Bussche and R. J. Baker, *J. Mammal.* **74**, 793 (1993)] are 1.5 to 3 times higher than the observed *Pan* distance.
36. M. Ruvolo, Harvard University, personal communication; E. Culotta, *Science* **260**, 892 (1993).
37. S. M. Stanley, *Macroevolution* (Freeman, San Francisco, 1979).
38. B. T. Shea and N. J. Coolidge, *J. Hum. Evol.* **17**, 671 (1988).
39. A. C. Hamilton, in *A Primate Radiation: Evolutionary Biology of the African Guenons*, A. Gautier-Hion, F. Bourlière, J.-P. Gautier, J. Kingdon, Eds. (Cambridge Univ. Press, Cambridge, 1988), pp. 13–34.
40. J. Rogers, *J. Hum. Evol.* **25**, 201 (1993); but see M. Ruvolo, *Am. J. Phys. Anthropol.* **94**, 89 (1994).
41. D. S. Woodruff, in *Proceedings of the 1985 Interdisciplinary Workshop* on "The Preservation and Valuation of Biological Resources," G. H. Orians, G. M. Brown, W. E. Kunin, J. E. Swierzbinski, Eds. (Univ. of Washington Press, Seattle, 1990), pp. 119–132; O. Ryder, *Trends Ecol. Evol.* **1**, 9 (1986).
42. A noninvasive method of rapid subspecies identification of captive chimpanzees using ASO probes was developed by P. A. Morin, J. J. Moore, D. S. Woodruff [*Proc. R. Soc. London B.* **249**, 293 (1992)], and discussed by M. A. Moon [*NIH Res. Resources Rep.* **17**, 7 (1993)].
43. K. B. Strier, *Faces in the Forest* (Oxford Univ. Press, Oxford, 1992); Proceeding of the International Conference on Implications of Disease for Captive Propagation and Reintroduction of Threatened Species, *J. Zoo. Wildlife Med.* **24**, 229 (1993); M. H. Woodford and P. B. Rossiter, in *Creative Conservation: Interactive Management of Wild and Captive Animals*. P. J. S. Olney, G. M. Mace, A. T. C. Feistner, Eds. (Chapman and Hall, London, 1994), pp. 178–200.
44. S. Anderson *et al.*, *Nature* **290**, 457 (1981).
45. C. C. Li, *First Course in Population Genetics* (Boxwood, Pacific Grove, CA, 1976).
46. A. Edwards, H. A. Hammond, L. Jin, C. T. Caskey, R. Chakraborty, *Genomics* **12**, 241 (1992).
47. R. R. Sokal and J. F. Rohlf, *Biometry* (Freeman, New York, 1969).
48. R. Deka, R. Chakraborty, R. E. Ferrell, *Genomics* **11**, 83 (1991).
49. S. W. Guo and E. A. Thompson, *Biometrics* **48**, 361 (1992); R. Chakraborty and Y. Zhong, *Hum. Hered.* **44**, 1 (1994).
50. M. Nei, *Genetics* **89**, 583 (1978).
51. R. Chakraborty, *ibid.* **108**, 719 (1984).
52. A. H. D. Brown, M. W. Feldman, E. Nevo, *ibid.* **96**, 523 (1980).
53. Using the test for independent segregation of loci suggested by N. E. Morton [*Proc. Natl. Acad. Sci. U.S.A.* **89**, 2556 (1992)], we also obtained results of no significant deviation from independence at the 5% level.
54. We thank all who donated chimpanzee hair and DNA samples: C. and H. Boesch (Basel University), E. J. Wickings (CIRMF), J. Sept (Indiana University), K. Winters (Jane Goodall Institute, Burundi), C. Noon and S. Siddle (Chimfunshi Wildlife Orphanage), T. Nishida and H. Takasaki (Kyoto University), J. Wallis, A. Collins, H. Matama, H. Mkono, the rest of the Gombe Stream Research Centre staff, and F. Nutter. Research by these individuals was authorized by the relevant governments in Africa, and hair was imported under CITES permit. We thank G. Yamamoto, C. Garza, H. Boyd, and R. Pheneger for assistance with laboratory work; J. Weber (Marshfield Medical Research Foundation), R. Shaw (University of Wisconsin), G. Reinartz (Zoological Society of Milwaukee County), and J. Ziegle (Applied Biosystems, Inc.) for primers; R. D'Andrade for the idea and performance of marker clustering; D. Butcher for programming; S. Kumar for comments and bootstrap analysis; N. Lehman for assistance with the permutation test; and K. Garner, S. Hrdy, M. Lynch, E. Martins, M. Nei, C. Orrego, M. Slatkin, D. W. Smith, P. Smouse, A. R. Templeton, M. Turelli, M. Ruvolo, O. Ryder, B. Dornon, and two anonymous reviewers for useful suggestions. Supported by Human Genome training grant 1T32 HG00005-02 (P.A.M.); by grants from the National Science Foundation (D.S.W. and J.J.M.), and by grants from the National Institutes of Health (D.S.W.).