Divergence Population Genetics of Chimpanzees

Yong-Jin Won and Jody Hey

Department of Genetics, Rutgers the State University of New Jersey, Piscataway, New Jersey

The divergence of two subspecies of common chimpanzees (*Pan troglodytes troglodytes* and *P. t. verus*) and the bonobo (*P. paniscus*) was studied using a recently developed method for analyzing population divergence. Under the isolation with migration model, the posterior probability distributions of divergence time, migration rates, and effective population sizes were estimated for large multilocus DNA sequence data sets drawn from the literature. The bonobo and the common chimpanzee are estimated to have diverged approximately 0.86 to 0.89 MYA, and the divergence of the two common chimpanzee subspecies is estimated to have occurred 0.42 MYA. *P. t. troglodytes* appears to have had a larger effective population size (22,400 to 27,900) compared with *P. paniscus*, *P. t. verus*, and the ancestral populations of these species. No evidence of gene flow was found in the comparisons involving *P. paniscus*; however a clear signal of unidirectional gene flow was found from *P. t. verus* to *P. t. troglodytes* ($2Nm = 0.51$).

Introduction

Although biological diversity has many causes, the single most important factor is the physical separation of populations (Wagner 1889; Dobzhansky 1937; Mayr 1942). Organisms in one population compete and exchange genes with one another, and they share in processes of genetic drift and adaptation. When populations split, these evolutionary forces cease to be shared between populations, and the populations can thereafter diverge. For many closely related populations or for the many situations where different, yet similar, populations have been assigned some taxonomic status (i.e., subspecies of one common species or species of a common genus), evolutionary biologists would like to understand when and how the splitting events occurred.

The chimpanzees of Africa, including the subspecies of the common chimpanzee (*Pan troglodytes*) and their sister species, the bonobo or pigmy chimpanzee (*P. paniscus*), are of great interest because they are the closest living species to our own. In recent years, they have been the subject of many studies on polymorphism and divergence (Morin et al. 1994; Kaessmann, Wiebe, and Paabo 1999; Deinard and Kidd 2000; Stone et al. 2002; Fischer et al. 2004). Morphological and genetic data strongly support the distinct species status of the bonobo (Shea and Coolidge 1988; Ruvolo et al. 1994; Kaessmann, Wiebe, and Paabo 1999; Stone et al. 2002; Yu et al. 2003). In the case of the common chimpanzee, three geographically defined subspecies have been recognized: *Pan troglodytes verus* in western Africa, *P. t. troglodytes* in central Africa, and *P. t. schweinfurthii* in eastern Africa (Schwartz 1934; Hill 1969; Morin et al. 1994). Recently, additional populations between the lower Niger River and Cameroon were proposed as another separate subspecies, *P. t. vellerosus* (Gonder et al. 1997; Gonder 2000). The subspecies designations of *Pan troglodytes* also have some support from molecular data, as nuclear and mitochondrial loci reveal some divergence between the subspecies (Gagneux et al. 1999; Kaessmann, Wiebe, and Paabo 1999; Deinard and Kidd 2000). However, only in the case of the mtDNA of *P. t. verus* was a subspecies found to have a monophyletic gene tree estimate (Morin et al. 1994; Gagneux et al. 1999; Gonder 2000). A morphometric study, using cranial measurements, also found some divergence between the subspecies, but there was a large amount of overlap between the subspecies (Shea and Coolidge 1988).

Comparative DNA sequence data can be used to study divergence, but the relationship between DNA sequence differences and the timing of population splitting and the processes associated with population splitting can be complex. Even under the simplest models, in which an ancestral population splits into two descendant populations with no gene exchange thereafter, the amount of divergence in DNA sequences between the two populations is a complex function of the time since the split and the relative sizes of the three populations (the two descendants and the ancestral population) (Wakeley and Hey 1997; Wang, Wakeley, and Hey 1997). For histories that include gene flow between diverging populations, the situation is even more complex because gene flow can create the appearance of recent divergence even if the actual splitting events occurred long ago. Whether or not gene flow has been occurring among chimpanzee species and subspecies is a question of considerable interest (Morin et al. 1994; Gagneux et al. 1999; Gonder 2000).

Here, we adapt recently developed methods for fitting the "isolation with migration" (or IM) model to the question of how and when chimpanzee species and subspecies diverged. This Markov chain Monte Carlo method yields estimates of multiple demographic parameters (including divergence time, migration rate, effective population sizes of two current populations and an ancestral population) (Nielsen and Wakeley 2001). Hey and Nielsen (2004) enhanced this method so that a large number of unlinked loci can be studied jointly to yield a posterior probability distribution for each of the demographic parameters in the IM model. We have applied the IM model to multilocus comparative DNA sequence data from two of the subspecies of the common chimpanzee (*P. t. troglodytes* and *P. t. verus*) and the bonobo (*P. paniscus*). We used the large data set of Yu et al. (2003) for 50 autosomal loci together with four other independent DNA data sets from other regions of the genome.

Key words: Chimpanzee, Bonobo, Markov chain Monte Carlo, speciation, gene flow.

E-mail: hey@biology.rutgers.edu.

Mol. Biol. Evol. 22(2):297–307. 2005
doi:10.1093/molbev/msi017
Advance Access publication October 13, 2004

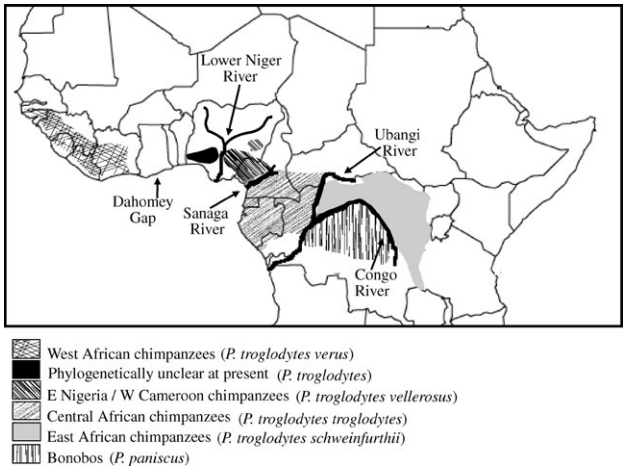

FIG. 1.—Distribution of common chimpanzee subspecies with some of their geographic boundaries (Schwartz 1934; Hill 1969; Gonder 2000; Gagneux et al. 2001; Bradley and Vigilant 2002). The subspecies status of chimpanzees from the west side of the Niger River is phylogenetically unclear at present (K. Gonder, personal communication).

Materials and Methods

Samples and Loci

The geographical distributions of chimpanzees and bonobos are represented in figure 1. The loci included in the study are listed in table 1. Yu et al.'s (2003) data includes 50 autosomal loci (GenBank accession numbers AY275957 to AY277244 and AY463943 to AY463951) sequenced in each of nine bonobos and 17 common chimpanzees (six *P. troglodytes verus*, five *P. t. troglodytes*, two *P. t. schweinfurthii*, and four individuals of unknown subspecies). We did not use the sequences of unknown origin or those from *P. t. schweinfurthii* because of the small sample size. We were left with three pairwise comparisons among the two common chimpanzees subspecies, *P. t. verus* and *P. t. troglodytes*, and the bonobos. The lengths of the 50 DNA segments range from 371 to 584 bp (average approximately 480 bp). Data for four other loci were also used: an intergenic region near the HoxB6 gene (~1 kb [AF116779 to AF116804]) (Deinard and Kidd 1999); a noncoding region of the X chromosome in Xq13.3 (10,154 bp [AJ270061 to AJ270095]) (Kaessmann, Wiebe, and Paabo 1999); cytochrome b gene (*cytb*) of the mtDNA (1,039 bp [AY585833 to AY585844]); and a portion of the nonrecombining Y chromosome (NRY) (2,784 bp [AF440112 to AF4401650]) (Stone et al. 2002).

The *cytb* gene sequences are previously unpublished data that were provided by Phillip Morin. These sequences from 12 chimpanzees were obtained using DNA from blood samples and by the PCR-sequencing method using flanking tRNA primers L14724 and H15915 (Irwin, Kocher, and Wilson 1991). Individuals were identified to subspecies on the basis of collection records. A check of mtDNA control region sequences in these samples revealed the same diagnostic subspecies differences previously found in larger samples (Morin, Moore, and Woodruff 1992; Morin et al. 1994).

The DNA sequences from the NRY, *cytb*, HoxB6, and Xq13.3 regions were collected in separate studies using mostly different individuals, spanning a range of sample sizes from five to 77 individuals per subspecies. Yu et al.'s (2003) DNA sequences for 50 loci were consistently obtained from the same set of individuals. The method of analysis assumes that populations are effectively panmictic. If this is approximately true, then the different sample sizes and the use of two sequences per individual in the study of Yu et al. (2003) should not introduce any biases.

The IM analysis requires sequence data from individual loci that show polymorphic variation within or between two populations and that do not show evidence of recombination (Nielsen and Wakeley 2001). Thus, we excluded those loci that showed no variation within or between the two taxa being compared in each pairwise analysis. Yu et al.'s (2003) 50 DNA segments reside in GenBank as diploid sequence (using IUPAC nucleotide ambiguity codes) without phase separation among heterozygous nucleotide sites. Because each region is short, it is unlikely that a given segment would show evidence of recombination had the data originally been obtained in haplotype form. Haplotypes were reconstructed using Clark's (1990) method, and then examined for evidence of recombination by the four-gamete test (Hudson and Kaplan 1985). Under Clark's method, haplotypes are first identified in homozygous individuals, after which those haplotypes already identified are subtracted from relevant heterozygous individuals to reveal remaining haplotypes. Approximately half of the loci in each of the analyses had no individuals who were heterozygous for more than one base position, in which case haplotype inference is straightforward. For a small number of loci (either two or three, depending on the species pair being considered), the reconstructed haplotypes showed evidence of recombination. These loci were excluded from the analysis. Among the remaining loci, there were two cases in which Clark's method leads to two configurations, one of which was consistent with recombination. In these cases, we used the haplotype configuration that did not show evidence of recombination.

The haplotype reconstruction protocol assumes that recombination over such short regions has been rare, and it also assumes that inferred haplotypes offer an unbiased view of history, relative to what would be found if true haplotype data had been available. For loci with several polymorphic sites, there can be multiple configurations of reconstructed haplotypes (Clark 1990). To check the effect of using alternative configuration, we also constructed data sets using alternative configurations for those loci that showed multiple possible haplotype configurations by Clark's method. Analyses with these alternative configurations were nearly identical to those for the primary data set and are available upon request. The similar results found for alternative haplotype configurations suggests that with these short loci, which have few polymorphic sites and little opportunity for recombination, the analyses are not highly sensitive to the method of haplotype inference. However, the larger question of how IM model analyses are affected by assumptions of low recombination, or of accurate haplotype inference, is an important one that has not been directly addressed by this comparison.

Table 1
DNA Segments Compared in Chimpanzees and Bonobos

Locus	DNA (bp)[a]	*P. paniscus* and *P.t. verus*	*P. paniscus* and *P.t. troglodytes*	*P.t. troglodytes* and *P.t. verus*	Reference[b]
T0151	421	—	✓	—	1
T10604	496	✓	✓	✓	1
T1251	493	✓	✓	✓	1
T1364	417	✓	✓	✓	1
T1412	553	✓	✓	✓	1
T1419	448	✓	—	✓	1
T1469	468	✓	—	✓	1
T1482	461	—	—	✓	1
T1506	428	✓	✓	—	1
T1568	424	—	✓	—	1
T1584	522	✓	✓	✓	1
T1636	505	✓	✓	✓	1
T2012	471	✓	✓	✓	1
T2018	518	✓	✓	✓	1
T2019	440	✓	✓	✓	1
T2020	456	✓	✓	✓	1
T2021	477	✓	✓	✓	1
T2041	522	—	—	✓	1
T2064	522	✓	✓	✓	1
T2085	509	✓	✓	✓	1
T2191	392	✓	✓	✓	1
T2265	451	✓	—	—	1
T2266	472	✓	✓	✓	1
T2294	710	✓	✓	✓	1
T2352	493	—	✓	✓	1
T2472	423	✓	✓	✓	1
T24894	434	✓	✓	✓	1
T2557	431	✓	✓	✓	1
T2558	427	✓	✓	✓	1
T2560	480	✓	✓	✓	1
T2563	523	✓	✓	✓	1
T2568	424	✓	—	✓	1
T2609	432	—	✓	—	1
T2659	452	✓	✓	✓	1
T2906	586	✓	✓	✓	1
T2920	470	✓	✓	✓	1
T2924	468	✓	✓	✓	1
T2963	541	✓	✓	✓	1
T2984	470	—	—	✓	1
T2986	493	✓	✓	✓	1
T2987	448	✓	✓	—	1
T2988	602	✓	✓	✓	1
T784	495	✓	✓	✓	1
T787	423	✓	✓	✓	1
T812	491	✓	✓	✓	1
T813	531	✓	✓	✓	1
T864	494	✓	✓	✓	1
T866	449	✓	✓	✓	1
T946	426	✓	✓	✓	1
T953	473	✓	✓	✓	1
CytB	1026	—	—	✓	2
HOXB6	993	✓	✓	✓	3
XQ13	5117	✓	✓	✓	4
NRY	2784	✓	✓	✓	5

NOTE.—Loci that could be included in each species contrast, because of available sequence and an absence of evidence for recombination, are indicated with a check mark. For the loci of Yu et al. (2003), the number of sequences are 18, 12, and 10 for *P. paniscus*, *P. t. verus*, and *P. t. troglodytes*, respectively. In that same order, the sample sizes for the following loci are 0, 5, and 7 (CytB); 38, 58, and 18 (HOXB6); 5, 16, and 12 (XQ13); and 7, 77, and 16 (NRY).

[a] The size of DNA fragments represents a maximum length chosen among the three comparisons.

[b] References: 1, Yu et al. (2003); 2, Morin et al. (1994); 3, Deinard and Kidd (1999); 4, Kaessmann, Wiebe, and Paabo (1999); 5, Stone et al. (2002).

The NRY data, consisting of 10 concatenated loci (2,784 bp in total consisting of sequence tagged site [STSs], in order, sY15, sY19, sY65, sY67, sY74, sY84, sY85, sY123, sY126, and sMCY) (Stone et al. 2002) showed no evidence of recombination by the four-gamete criterion. For the mtDNA data used in the analysis of *P. t. verus* and *P. t. troglodytes*, one polymorphic site at the 3′ end of the sequence was not congruent with the remainder

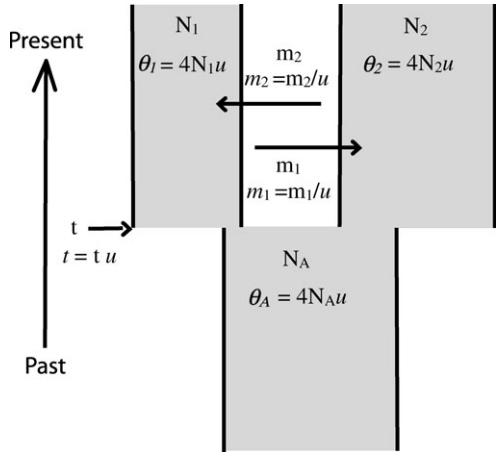

FIG. 2.—The isolation with migration (IM) model is depicted with two sets of parameters. The basic demographic parameters are constant effective population sizes (N_1, N_2, and N_A), gene-flow rates per gene copy per generation (m_1 and m_2), and the time of population splitting at t generations in the past. The parameters in the second set (in italics) are all scaled by the neutral mutation rate u, and it is these parameters that are actually used in the model fitting.

of the sequence, by the four-gamete criterion. This site was simply dropped from the analysis. For the HoxB6 and Xq13.3 data, applications of the four-gamete criterion revealed evidence of some recombination events. In each case, we selected the largest fragment of the data that showed no evidence of recombination. This action represents a tradeoff between the need to have a large number of loci, from different portions of the genome, and the concern that selecting large nonrecombinant blocks may bias the results. The reason for potential bias is that the analytical methods assume that loci have been sampled randomly with respect to their genealogical histories and that loci not having had recombination are expected to have shorter gene trees, on average, than other loci (i.e., those loci with shorter genealogical histories have had less time for recombination). This effect is probably quite subtle, particularly in these cases where the selected fragments of HoxB6 and Xq13.3 were not less polymorphic than those regions that were excluded. The data sets are summarized in table 1.

IM Model Computations

The posterior probability densities of the parameters of the IM model are generated by simulating a Markov chain having a stationary distribution that is proportional to that density. The basic procedure is to begin a simulation with a burn-in period (100,000 steps in our analyses), so that the state of the chain becomes independent of the starting point, and then to continue the simulation for a long time while measuring the parameter values repeatedly over the course of the run. Convergence by the Markov chain simulations, upon the true stationary distribution, is assessed by monitoring multiple independent chains started at different starting points and by assessing the autocorrelation of parameter values over the course of the run. We also used a procedure for swapping among multiple heated chains (Metropolis coupling) to further ensure that the distributions we obtained actually reflected the stationary distributions (Geyer 1992). Each locus was assigned an inheritance scalar, to adjust for its relative expected effective population size: 1.0 for autosomes, 0.25 for mtDNA and NRY, and 0.75 for X-linked loci. Individual simulations were run for 60 million updates or more. Metropolis-coupling runs used 10 coupled chains that varied over a range of heating values. The settings for the prior distributions were empirically obtained after preliminary running with larger parameter intervals. This approach violates the spirit of a Bayesian analysis (in which available prior information is included). However, we wished to exclude other information from these analyses and so selected uninformative prior distributions that would not contribute to the posterior distributions. When this is done, the posterior probability distributions are proportional to likelihood distributions, and the parameter values associated with the highest likelihoods are maximum-likelihood estimates (Nielsen and Wakeley 2001).

The IM model has six demographic parameters, each scaled by the overall neutral mutation rate (fig. 2). With multiple loci, each locus has a mutation rate scalar parameter such that the product of all mutation rate scalars is equal to 1. Thus, with multiple loci, the overall neutral mutation rate represented in the demographic parameters is the geometric mean of all of the individual locus mutation rates (Hey and Nielsen 2004). For each of the six demographic parameters in each analysis, we recorded the marginal density (as a histogram with 1,000 equally sized bins) over the course of multiple long simulations. The peaks of the resulting distributions were taken as estimates of the parameters (Nielsen and Wakeley 2001). For credibility intervals, we assessed for each parameter the 90% highest posterior density (HPD) interval, which are the boundaries of the shortest span that includes 90% of the probability density of a parameter.

To convert parameter estimates to more easily interpreted units, we used the divergence that has occurred over the approximately 6 Myr since the splitting between human and chimpanzee lineages (Chen and Li 2001; Brunet et al. 2002; Vignaud et al. 2002; Glazko and Nei 2003; Wildman et al. 2003). The geometric mean of the human-chimpanzee DNA sequence divergence of all loci was calculated and then used to convert the estimate of the time parameter, t, to divergence in years. We estimated the geometric mean divergence, between chimpanzees and humans, separately for each of the three chimpanzee comparisons because each analysis is based on a slightly different data set (table 1). The average divergence between humans and the two taxa examined with the IM model was estimated as 5.8801 for the *P. paniscus–P. t. verus* pair, 5.9945 for the *P. paniscus–P. t. troglodytes* pair, and 6.2463 for *P. t. troglodytes–P. t. verus* pair. Using the total human/chimpanzee divergence time of 12 Myr, these values correspond to 4.90008×10^{-7}, 4.99541×10^{-7}, and 5.205234×10^{-7} mutation events per locus per year, for the three different analyses, respectively. The differences between these values are primarily caused by some loci being used in only one or two of the three species comparisons. Using these values estimates of t (the

number of mutations since population splitting [see fig 2]) can be directly converted to estimates of the number of years since population splitting.

To convert the estimates of the population mutation rate parameters (θ_1, θ_2, and θ_A) to estimates of effective population size (N_1, N_2, and N_A, respectively), we need a measure of mutation rate on a scale of generations. We assumed 15 years per generation for the chimpanzees then multiplied the estimated mutation rate per year (based on human/chimpanzee divergence) by 15 years per generation. These calculations yielded values for the geometric mean number of mutations per generation per locus of 7.350×10^{-6} for the *P. paniscus–P. t. verus* pair, 7.493×10^{-6} for the *P. paniscus–P. t. troglodytes* pair, and 7.808×10^{-6} for the *P. t. troglodytes–P. t. verus* pair. These mutation rate values were then used to convert individual θ estimates to effective population size estimates (i.e., $\theta = 4Nu$ and $N = \theta/(4u)$).

Migration parameters in the model can be used to obtain population migration rate estimates (i.e., $M = 2Nm$, the product of the effective number of gene copies and the per gene copy migration rate) using an estimate of the population mutation rate ($\theta = 4Nu$). Thus, $M = \theta \times m/2 = (4Nu \times m/u)/2 = 2Nm$ (Hey and Nielsen 2004).

One of the benefits of a method that explicitly incorporates a changing genealogy for each locus over the course of the analysis is that the posterior densities of other quantities that are associated with the genealogy can also be recorded. We took this approach for migration events for those cases where the method reveals nonzero migration rate estimates. For each locus, we measured over the course of the simulation the distribution of the number of migration events and the distribution of the average time of migration.

Results

A total of 48 loci were used for the comparison between the central (*P. t. troglodytes*) and the western (*P. t. verus*) chimpanzees, and 46 loci were used for the two comparisons involving *P. paniscus* (table 1). Total lengths of the multiple aligned DNA sequences amounted to 29,626 bp, 27,169 bp, and 31,188 bp for the three comparisons in order as listed in table 1. The average sequence divergence per site between pairs of taxa, excluding indels, was 0.377% (*P. paniscus* versus *P. t. troglodytes*), 0.362% (*P. paniscus* versus *P. t. verus*), and 0.211% (*P. t. troglodytes* versus *P. t. verus*).

Repeated runs of the IM program revealed unambiguous marginal posterior probability distributions of the parameters for all three species comparisons. The peaks of the primary six parameter values were confined to fairly narrow ranges with corresponding credibility intervals illustrated in figure 3.

Pan paniscus and *P. troglodytes verus*

The maximum-likelihood effective population size estimates for *P. paniscus*, *P. t. verus*, and their ancestral population were 9,700 (90% HPD interval: 7,300 to 1,2500), 8,500 (6,300 to 11200), and 16,300 (30 to 29,600), respectively (table 2). The distribution of the ancestral population size parameter is broader and flatter than those for the current populations of *P. paniscus* and *P. t. verus* (fig. 3*a*), as expected if the ancestral population existed long ago. The marginal posterior probability distribution of the divergence time parameter, *t*, revealed a sharp peak at 0.42, with a narrow distribution (fig. 3*c*). When converted to a scale of years, the divergence time between the two taxa was estimated to be 0.859 MYA with 90% HPD interval of 0.589 to 1.31 MYA (table 2). The migration parameters revealed a peak at the lower limit of resolution in both directions (from *P. paniscus* to *P. t. verus* and from *P. t. verus* to *P. paniscus*) (fig. 3*b*). Although it is possible that a histogram record with a much finer resolution could reveal a nonzero peak that is further to the left of the smallest interval that was measured, we hereafter interpret the locations of these peaks as being at zero.

Pan paniscus and *P. troglodytes troglodytes*

In this comparison, the maximum-likelihood effective population size estimate of *P. paniscus* was 9,200 (90% HPD interval: 7,000 to 12,000). This estimate is similar to that obtained in the analyses between *P. paniscus* and *P. t. verus* comparison. The effective size of the *P. t. troglodytes* population was estimated to be 22,400 (90% HPD interval: 17,000 to 29,300), and the ancestral effective population size estimate of the two species was 15,300 (90% HPD interval: 1,900 to 26,200) (fig. 3*d* and table 2), which is similar to that found in the comparison between *P. paniscus* and *P. t. verus*. *P. t. troglodytes* appears to have had the largest effective population size among the three taxa examined. Unlike *P. paniscus* and *P. t. verus*, which were found to have similar or slightly smaller population sizes than their ancestral populations, the estimate for *P. t. troglodytes* is nearly twice that of the ancestral population size. Divergence time was estimated to be 0.89 MYA between *P. paniscus* and *P. t. troglodytes* (90% HPD interval: 0.638 to 1.33 MYA) (fig. 3*f* and table 2). This estimate is very close to that of the divergence time between *P. paniscus*s and *P. t. verus*, which suggests that *P. t. verus* and *P. t. troglodytes* descended from the very same ancestral population, diverging from the ancestor of *P. paniscus*. This finding is expected, given phylogenetic studies (Morin et al. 1994; Ruvolo et al. 1994; Gagneux et al. 1999; Stone et al. 2002). Both migration rate parameters revealed little evidence for gene flow (fig. 3*e* and table 2). Although the location of the highest probability for migration rate from *P. t. troglodytes* to *P. paniscus* (m_1 [table 2]) was not at the very lowest value, it is very close in position and height to the lowest migration value in the histogram (fig. 3*e*).

Pan troglodytes troglodytes and *P. t. verus*

The effective population sizes of *P. t. troglodytes*, *P. t. verus*, and their ancestral population were estimated to be 27,900 (90% HPD interval: 19,600 to 40,700), 7,600 (5,300 to 10,700), and 5,300 (200 to 11,300), respectively (table 2). These sizes are similar to the estimates from the comparisons with *P. paniscus*, with *P. t. troglodytes* having the largest effective population size (fig. 3*g*). The

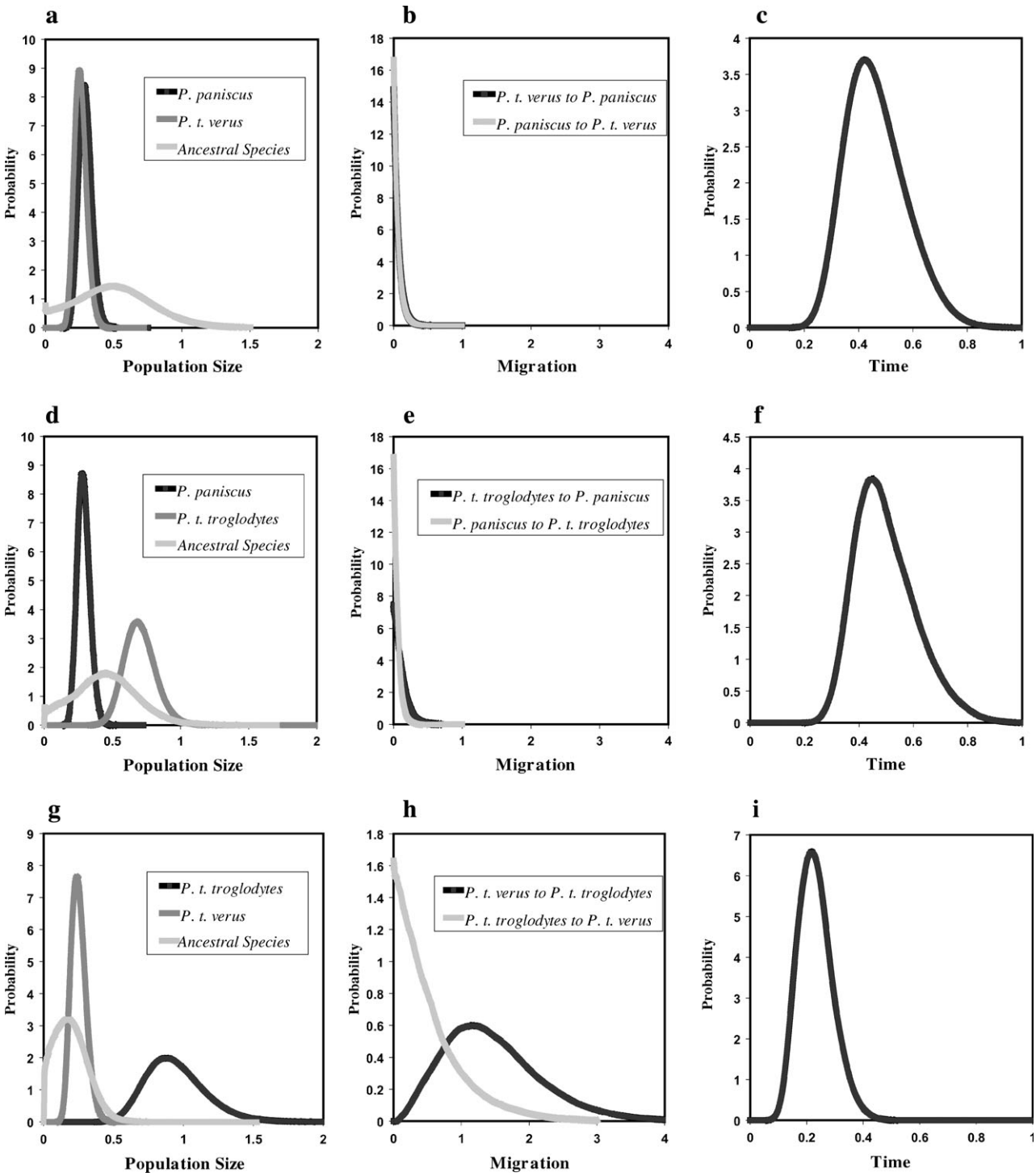

FIG. 3.—The marginal posterior probability distributions for model parameters (scaled by the neutral mutation rate). Curves are shown for the analysis with *P. paniscus* and *P. t. verus* (*a*), (*b*), and (*c*); for *P. paniscus* and *P.t. troglodytes* (*d*), (*e*), and (*f*); and for *P. t. troglodytes* and *P. t. verus* (*g*), (*h*), and (*i*).

divergence time was estimated to be 0.422 MYA (90% HPD interval: 0.255 to 0.629 MYA), with a sharply peaked marginal posterior distribution (fig. 3*i*). A comparisons of divergence times among the three taxa clearly suggests that the common chimpanzee populations (*P. t. troglodytes* and *P. t. verus*) were descended from an ancestor that had earlier separated from the lineage leading to *P. paniscus*, consistent with other phylogenetic studies (Morin et al. 1994; Ruvolo et al. 1994; Gagneux et al. 1999; Stone et al. 2002). Migration rate distributions suggest a moderate level of gene flow from *P. t. verus* to *P. t. troglodytes* ($2Nm = 0.514$ [table 2]). Also, because the estimate of the probability that the migration rate is zero is itself zero (fig. 3*h*), we can reject a gene-flow rate of zero

Table 2
Maximum-Likelihood Estimates[a] (MLE) and the 90% Highest Posterior Density (HPD) Intervals[b] of Demographic Parameters

Comparison	θ_1	θ_2	θ_A	m_1	m_2	t	N_1	N_2	N_A	$2N_1m_1$	$2N_2m_2$	t (years)
P. paniscus × *P. t. verus*												
MLE	0.284	0.249	0.479	0.0005	0.0005	0.421	9,700	8,500	16,300	0.00007	0.00006	859,000
Lower 90% HPD	0.213	0.184	0.0008	0.0005	0.0005	0.289	7,300	6,300	30			589,000
Upper 90% HPD	0.368	0.331	0.872	0.151	0.141	0.641	12,500	11,200	29,600			1,309.000
P. paniscus × *P. t. troglodytes*												
MLE	0.275	0.672	0.459	0.004	0.0005	0.459	9,200	22,400	15,300	0.00048	0.00017	890,000
Lower 90% HPD	0.210	0.510	0.057	0.0005	0.0005	0.319	7,000	17,000	1,900			638,000
Upper 90% HPD	0.360	0.877	0.786	0.245	0.125	0.665	12,000	29,300	26,200			1,332,000
P. t. troglodytes × *P. t. verus*												
MLE	0.872	0.238	0.164	1.179	0.002	0.220	27,900	7,600	5,300	0.514	0.00018	422,000
Lower 90% HPD	0.613	0.165	0.005	0.314	0.002	0.133	19,600	5,300	200			255,000
Upper 90% HPD	1.273	0.334	0.352	2.541	1.226	0.328	40,700	10,700	11,300			629,000

[a] MLE estimates are the locations of the peaks in the curves shown in figures 3, 4, and 5.

[b] The 90% HPD intervals are the shortest spans, along the *x* axes of figures 3, 4, and 5, that contain 90% of the area of those histograms. For basic parameters not scaled by the mutation rate including N_1, N_2, N_A, $2N_1m_1$, $2N_2m_2$, and t (see *Materials and Methods* and figure 2), these HPD intervals are not directly available. However in the case of time (t) and effective population sizes (N_1, N_2, and N_A), estimates of the 90% HPD intervals were made as the products of those for the corresponding scaled parameters and the conversion factors based on the human/chimpanzee divergence (which were taken to be correct without error, see *Materials and Methods*).

in the direction of *P. t. verus* to *P. t. troglodytes*. The estimate of migration in the opposite direction was estimated to be very near zero (fig. 3*h* and table 2), with a peak height that is very near to that at zero.

The finding of gene flow could be caused if one of the individuals identified as *P. t. troglodytes* in the study of Yu et al. (2003) was actually a hybrid or backcross hybrid between this subspecies and *P. t. verus*. To check this possibility, we examined the pairwise sequence divergence across all 50 loci within and between the subspecies of *P. t. verus* and *P. t. troglodytes*. None of the *P. t. troglodytes* individuals were appreciably closer than others to the *P. t. verus* individuals (results not shown), arguing against recent hybridization as the cause of the apparent gene flow.

To take a closer look at gene flow from *P. t. verus* to *P. t. troglodytes*, the distribution of the number and mean time of migration events were recorded over the course of the simulations, for each of the loci (table 3). The modal number of migration events per locus was one, and the mean time of migration rates was 0.098, which corresponds to 0.186 Myr, roughly half of the divergence time between *P. t. verus* and *P. t. troglodytes*. As shown in table 3, most loci (37 out of 48) had a modal number of migration events of one, with a few loci having a mode of zero (loci T2012, T2266, T2988, T812, T946, *cytb*, and NRY), two (loci T2019 and T2984), and three migrations (locus XQ13) during the simulations. It is interesting that both of the sex-limited loci (*cytb* of the mtDNA, and NRY) showed less evidence of gene flow than most other loci, although an absence of gene flow was expected for the mtDNA, given previous findings (Morin et al. 1994).

Discussion

In recent years, the timing and mode of divergence of the bonobo and the subspecies of common chimpanzees has received considerable attention, particularly with regard to whether or not different species and subspecies have been exchanging genes (Morin et al. 1994; Deinard and Kidd 1999; Kaessmann, Wiebe, and Paabo 1999; Gonder 2000; Gagneux and Varki 2001; Kaessmann et al. 2001; Stone et al. 2002; Fischer et al. 2004). We have conducted a detailed analysis of chimpanzee divergence using a new protocol that explicitly incorporates directional gene-flow parameters in a model of population splitting (Nielsen and Wakeley 2001; Hey and Nielsen 2004). In addition to estimates of population size and population splitting times, we find clear evidence of gene flow from *P. t. verus* to *P. t. troglodytes*.

Effective Population Size Estimates

Using data from 50 loci, Yu et al. (2003) estimated the effective population size of *P. paniscus*, *P. t. troglodytes*, and *P. t. verus* to be 12,400, 20,100, and 13,000, respectively. Our analyses, based primarily on these same data, also found a larger size for *P. t. troglodytes* and smaller sizes for *P. paniscus* and *P. t. verus* (table 2). However, our values for the latter two species are lower than estimated by Yu et al. (2003), who based their estimate on the average pairwise divergence between sequences. This estimator is known to have a large variance (Tajima 1983) and to be sensitive to the site frequency distribution of polymorphic sites (Tajima 1989). Our effective population size estimate of 27,900 for *P. t. troglodytes*, obtained in the analysis with *P. t. verus*, is larger than the values estimated by Yu et al. (2003), possibly because the IM method explicitly estimates the effective size since speciation. Given that the estimated size for *P. t. troglodytes* is considerably larger than the estimated size of the ancestral population for this and *P. t. verus* (5,300 [table 2]), it appears that the population size of *P. t. troglodytes* has grown since this divergence began. Interestingly, using nine unlinked loci representing 19,000 bp from 14 unrelated individuals, Fischer et al. (2004) reported that chimpanzees in central Africa have a larger effective population size (25,000 or 35,000 according to two different methods) than chimpanzees in western

Table 3
The Number of Migration Events and Mean Time of Migration Events from *P. t. verus* to *P. t. troglodytes*

Locus	T10604	T1251	T1364	T1412	T1419	T1469	T1482	T1584	T1636
Migration[a]	1	1	1	1	1	1	1	1	1
Time[b]	0.0922	0.0923	0.0947	0.0710	0.0950	0.1055	0.0915	0.1002	0.0951
Locus	T2012	T2018	T2019	T2020	T2021	T2041	T2064	T2085	T2191
Migration	0	1	2	1	1	1	1	1	1
Time	0.1351	0.0933	0.0678	0.0968	0.0879	0.0819	0.0974	0.0952	0.0883
Locus	T2266	T2294	T2352	T2472	T24894	T2568	T2557	T2558	T2560
Migration	0	1	1	1	1	1	1	1	1
Time	0.1169	0.0934	0.0930	0.0919	0.0923	0.0923	0.0933	0.0952	0.0925
Locus	T2563	T2659	T2906	T2920	T2924	T2963	T2984	T2986	T2988
Migration	1	1	1	1	1	1	2	1	0
Time	0.0941	0.0984	0.1016	0.0948	0.0944	0.0914	0.0589	0.0961	0.1037
Locus	T784	T812	T787	T813	T864	T866	T946	T953	cytb
Migration	1	0	1	1	1	1	0	1	0
Time	0.0980	0.1074	0.0962	0.0934	0.0968	0.0962	0.1354	0.1045	0.1689
Locus	HOXB6	XQ13	NRY	All Loci					
Migration	1	3	0	1[c]					
Time	0.1090	0.0514	0.1712	0.0980[d]					

[a] Modal number of migration events observed over the course of the Markov chain simulations.
[b] The mean value of the time at which mutation events occurred (scaled by the neutral mutation, as with *t*) for those genealogies that had at least one migration.
[c] Modal value across all loci.
[d] The mean of the times for all loci.

Africa. Also, a demographic change (population growth or fine-scale population structure) was inferred from the allele frequency spectrum (Fischer et al. 2004).

Divergence Times

In our study, the estimates of divergence time between *P. paniscus* and the two subspecies of common chimpanzees (*P. t. troglodytes* and *P. t. verus*) tend to be smaller than those of previous studies, which used different methods. The method applied here explicitly accounts for ancestral population size by assessing the divergence time parameter jointly with other demographic parameters. Like other studies, we used a calibration point based on estimates of *Homo-Pan* splitting time of 6 MYA (Chen and Li 2001; Brunet et al. 2002; Vignaud et al. 2002; Glazko and Nei 2003; Wildman et al. 2003).

According to the IM analysis, the most probable divergence time between *P. paniscus* and chimpanzees (*P. t. troglodytes* and *P. t. verus*) was estimated to be 0.862 to 0.896 MYA (table 2). In contrast, Yu et al. (2003) estimated the divergence time between chimpanzees and bonobos to be 1.8 MYA. These authors used the *Homo-Pan* split at 6 MYA as a calibration point, together with an average sequence divergence (1.22%) of the 50 loci between human and chimpanzee. However, their date is simply the estimated average time of the ancestral sequence for pairs of sampled sequences and takes no account of the variation between species that is caused by variation in the ancestral population. Similar analyses in studies of individual loci also lead to high values for divergence between bonobos and common chimpanzees. The study on the nonrecombining portion of the Y chromosome (NRY) had an estimated divergence time of 1.8 MYA (Stone et al. 2002). In the case of the mtDNA, the divergence time was estimated to be 2.5 MYA (Gagneux et al. 1999). For the X chromosome locus, Xq13.3, Kaessmann, Wiebe, and Paabo (1999) estimated the ancestor to be 0.9 MYA, which is similar to our estimate.

In the case of *P. t. troglodytes* and *P. t. verus*, the method of Yu et al. (2003) converts the observed average sequence divergence between *P. t. troglodytes* and *P. t. verus* (0.125%) to a divergence time estimate of 0.62 MYA. In other single-locus studies, the divergence time for these chimpanzee subspecies was estimated to be 0.61 MYA for the NRY (excluding the uncertain haplotypes that we also excluded in our analysis) (Stone et al. 2002); 0.6 MYA for the hypervariable mtDNA region (Gonder 2000), and 2.1 MYA for Xq13.3 (Kaessmann et al. 2001). In this last case, the divergence time between subspecies of chimpanzees was estimated to be older than that of the divergence time between bonobo and chimpanzees because of a large distance between some chimpanzee subspecies and the bonobo (Kaessmann, Wiebe, and Paabo 1999). In our analysis, the divergence time of the chimpanzees in western and central Africa (*P. t. verus* and *P. t. troglodytes*) was estimated to be 0.422 MYA (90% HPD interval: 0.255 to 0.629). Fischer et al. (2004) estimated divergence times using multilocus DNA sequences (including the data of Yu et al. [2003]) and the moment estimator method of Wakeley and Hey (1997). This approach assumes a four-parameter isolation model (the same model as in figure 1 but without migration) and, thus, explicitly accounts for that component of variation caused by common ancestry. When the authors calculated divergence times using only the Yu et al.'s (2003) data, their divergence time estimates were very similar to those reported here: 0.8 MYA for bonobos and chimpanzees and 0.43 MYA for chimpanzees in western and central Africa.

In summary, we estimate that the separation of common chimpanzees and bonobos began about 900,000 years ago and was followed later by a separation at about 400,000 years between the chimpanzees in western and central Africa. These events occurred during the Pleistocene (1.8 to 0.01 MYA) epoch, which has been

characterized by repeated glaciations accompanying climatic change and shift in geography of tropical forests in Africa (Hamilton 1988). Particularly, since the beginning of major glaciations around 2.4 MYA, climatic changes in Africa both before and after 1 MYA were considered to be more severe than before (Stein and Sarnthein 1984).

Isolation Since Population Splitting

A major benefit of being able to consider a full isolation with migration model is access to two-directional gene-flow rates (fig. 2). From our analyses, bonobos and common chimpanzees appear to have been isolated without gene flow since they began to diverge. Put another way, the divergence of these species appears consistent with a speciation model in which geographic isolation prevented gene flow during the separation of these species.

In contrast to the analyses involving *P. paniscus*, we found a clear signal of unidirectional gene flow from *P. t. verus* to *P. t. troglodytes* (fig. 3*h*). This finding is perhaps our most surprising, in part because of the current disjunct distributions of *P. t. verus* and *P. t. troglodytes* (fig. 1) (see figure 1 in Kortlandt [1983]). According to Kortlandt, an area (>1,000 km) from west of Ghana to the eastern side of the lower Niger River, except for one islandlike habitat region on the western side of the lower Niger River, appears to be nearly devoid of chimpanzees (the Dohomey Gap in figure 1), although some places were marked as exterminated habitats during recent years or since around 1940. Thus, evidence of gene flow suggests that the geographic distributions of these populations have changed over time.

A related puzzle is why migration between these populations might have occurred in only one direction. Although there may be many possible explanations, given the several hundred thousand years since these populations began to diverge, we consider the possibility that another population, that we have not sampled, may play a role in our findings. Recently, Gonder et al. (1997) studied a chimpanzee population that inhabits a region of Nigeria and Cameroon that lies between the lower Niger River and the Sanaga River. Mitochondrial DNA sequences from this region revealed a monophyletic lineage, having sister relationship to sequences from *P. t. verus* (Gonder et al. 1997; Bradley and Vigilant 2002). On this phylogeographic basis, these chimpanzees were recognized as a separate subspecies, *P. t. vellerosus* (Gray 1862; Gonder 2000). However, whereas this group is genetically close to *P. t. verus* (based on mtDNA), it is geographically close to *P. t. troglodytes* of central Africa, from which it is separated by the Sanaga River in Cameroon (Gagneux et al. 2001). Gonder investigated this border region with fine-scale sampling from both sides of the Sanaga River and found a few mtDNA haplotypes that clustered with samples from the southern side of the Sanaga River, suggesting that the river might not be a strict barrier for migration. Furthermore, microsatellite data revealed a high effective number of migrants per generation ($Nm = 11$) across the Sanaga River (Gonder 2000). Notwithstanding the possible inflation of this *Nm* estimate because of homoplasy in microsatellite markers, both these and the mtDNA suggest intermixing between these populations. In addition, cranial features of *P. t. vellerosus* more closely resemble those of *P. t. troglodytes* than *P. t. verus* (Groves 2001).

This information on the *P. t. vellerosus* population suggests an explanation for the observation of unidirectional gene flow from *P. t. verus* to *P. t. troglodytes*. The scenario begins with a separation of populations that today we recognize as *P. t. verus* and *P. t. troglodytes*. If, as suggested by mtDNA data, the population identified as *P. t. vellerosus* is indeed most closely related to *P. t. verus*, then it seems likely that this population formed sometime after the original split that gave rise to the central Africa and western Africa subspecies. However, today *P. t. vellerosus* occurs geographically near populations of *P. t. troglodytes* and appears to be exchanging genes with the central Africa subspecies. Thus, *P. t. vellerosus* may be a channel for genes into *P. t. troglodytes* that otherwise are closely related to genes of *P. t. verus*. It is possible that this is the mechanism whereby nonzero levels of gene flow appear to have occurred between *P. t. troglodytes* and *P. t. verus*, which today have quite geographically disjunct distributions. Clearly, testing this hypothesis would require inclusion of the *P. t. vellerosus* population in a multilocus study of the IM model. This scenario also reveals a limitation of the IM model as currently implemented. At present, we can only study populations in pairs, but in reality, many closely related populations occur in more complex geographic and demographic contexts.

The Population Status of Chimpanzee subspecies

Subspecies of the common chimpanzee have been designated on the basis of geographic ranges, with populations that are separated by large regions or by large rivers sometimes being assigned a subspecies designation (Groves 2001). These designations have found some support in phylogeographic studies of individual loci, at least in so far as some degree of genetic differentiation is repeatedly observed (see review Bradley and Vigilant [2002]). Their support on morphological grounds is less clear (Groves 2001). Certainly the current study, which includes just two subspecies, finds strong evidence of considerable (although not complete) isolation over several hundred thousands of years.

It might be argued that the designation of chimpanzee subspecies, which are based largely on current geographic distributions together with limited genetic data, are not well justified. In this study, we have taken the approach of treating previously identified taxa as hypotheses of the presence of distinct and isolated populations (Baum 1998; Hey et al. 2003). Notwithstanding the evidence of gene flow from *P. t. verus* to *P. t. troglodytes*, our analyses strongly affirm the use of these taxonomic designations, as the populations that include representatives of these taxa appear to have long been largely evolutionarily separated from one another.

It is also important to note that studies of gene flow within subspecies of common chimpanzees suggest that these taxa occur as intermingled populations. Goldberg and Ruvolo (1997) examined mtDNA migration rates for the eastern Africa chimpanzee *P. t. schweinfurthii*, with

extensive sampling at 19 locations encompassing most of the known range of the subspecies. They estimated that the population migration rate was 3.38 among the locations and that the maximal distance of any haplotype sharing was 583 km. Despite the high gene-flow rate, a pattern of long-distance differentiation was observed in that study. Similarly for the eastern Africa chimpanzee, Gagneux et al. (1999) reported a long distance (maximum 1,000 km) haplotype sharing within the western Africa chimpanzee.

Our findings of gene flow spanning subspecies ranges may be attributable to mating strategies that overcome local inbreeding effects. Chimpanzees are known to exchange genes among groups by female transfer and mating with noncommunity members (Goodall 1986; Boesch and Boesch-Achermann 2000). For example, genetic estimates of extragroup paternity (EGP) levels found a level of 1% in Taï in West Africa (Boesch and Boesch-Achermann 2000) and 7% in three contiguous communities in the same Taï National Park in West Africa (Vigilant et al. 2001). However, no evidence of EGP was detected in an eastern Africa community (Constable et al. 2001).

In contrast to the pattern of gene flow within and among common chimpanzee subspecies, the divergence of the bonobo and the common chimpanzee is consistent with a long history (approximately 900,000 years) without gene flow. One probable factor in this divergence is geographic separation caused by climatic changes that resulted in deforestation and the expansion of arid savanna. Historical fluctuation of forest and savanna ranges caused by climate changes during the Pleistocene could have split ancestral populations and confined them in multiple geographical groups (Grubb 1982). Also, rivers that are wide and that separate geographic regions have probably been major barriers to gene flow. The Congo River, which currently separates bonobos from common chimpanzees probably had a continuous existence since well before the Pleistocene (Beadle 1981), although it can not be ruled out that there were historical time periods in some locations along the river that would have been permissive of a dispersal corridor for chimpanzees.

Acknowledgments

We thank Phillip Morin for generously providing his unpublished *cytb* DNA sequences and for helpful comments on the manuscript. We thank Brenda Bradley for generously providing an electronic version of a map of the distributions of common chimpanzees and bonobos. We also thank Katy Gonder for a helpful comments on the phylogenetic status of the chimpanzees in Nigeria and Cameroon, as well as Makoto Shimada, Arjun Sivasundar, and Molly Przeworski for their helpful comments on the manuscript.

Literature Cited

Baum, D. A. 1998. Individuality and the existence of species in time. Syst. Biol. **47**:641–653.

Beadle, L. C. 1981. The inland waters of tropical Africa. Longman, London.

Boesch, C., and H. Boesch-Achermann. 2000. The Chimpanzees of the Taï forest. Oxford University Press, Oxford.

Bradley, B. J., and L. Vigilant. 2002. The evolutionary genetics and molecular ecology of chimpanzees and bonobos. Pp. 259–275 *in* C. Boesch, G. Hohmann, and L. F. Marchant, eds. Behavioural diversity in chimpanzees and bonobos. Cambridge University Press, Cambridge.

Brunet, M., F. Guy, D. Pilbeam et al. (38 co-authors). 2002. A new hominid from the Upper Miocene of Chad, Central Africa. Nature **418**:145–151.

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

Clark, A. G. 1990. Inference of haplotypes of PCR-amplified samples of diploid populations. Mol. Biol. Evol. **7**:111–122.

Constable, J. L., M. V. Ashley, J. Goodall, and A. E. Pusey. 2001. Noninvasive paternity assignment in Gombe chimpanzees. Mol. Ecol. **10**:1279–1300.

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

———. 2000. Identifying conservation units within captive chimpanzee populations. Am. J. Phys. Anthropol. **111**: 25–44.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

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

Gagneux, P., M. K. Gonder, T. L. Goldberg, and P. A. Morin. 2001. Gene flow in wild chimpanzee populations: what genetic data tell us about chimpanzee movement over space and time. Philos. Trans. R. Soc. Lond. B Biol. Sci. **356**: 889–897.

Gagneux, P., and A. Varki. 2001. Genetic differences between humans and great apes. Mol. Phylogenet. Evol. **18**:2–13.

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

Geyer, C. J. 1992. Practical Markov chain Monte Carlo. Stat. Sci. **7**:473–511.

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

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

Gonder, M. K. 2000. Evolutionary genetics of chimpanzees in Nigeria and Cameroon. Doctoral dissertation, City University of New York, New York.

Gonder, M. K., J. F. Oates, T. R. Disotell, M. R. Forstner, J. C. Morales, and D. J. Melnick. 1997. A new west African chimpanzee subspecies? Nature **388**:337.

Goodall, J. 1986. The chimpanzees of Gombe. Harvard University Press, Cambridge, Mass.

Groves, C. 2001. Primate taxonomy. Smithsonian Institution Press, Washington, DC.

Grubb, P. 1982. Refuges and dispersal in the speciation of African forest mammals. Pp. 537–553 *in* G. T. Prance, ed. Biological diversification in the tropics. Academic Press, New York.

Hamilton, A. C. 1988. Guenon evolution and forest history. Pp. 13–34 *in* A. Gautier-Hion, F. Bourlière, and J.-P. Gautier, eds.

A primate radiation: evolutionary biology of the African guenons. Cambridge University Press, Cambridge.
Hey, J., and R. Nielsen. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics **167**:747–760.
Hey, J., R. S. Waples, M. L. Arnold, R. K. Butlin, and R. G. Harrison. 2003. Understanding and confronting species uncertainty in biology and conservation. Trends Ecol. Evol. **18**:597–603.
Hill, W. C. O. 1969. The nomenclature, taxonomy and distribution of chimpanzees. Pp. 22–49 *in* G. H. Bourne, ed. The chimpanzee. Karger, New York.
Hudson, R. R., and N. L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics **111**:147–164.
Irwin, D. M., T. D. Kocher, and A. C. Wilson. 1991. Evolution of the cytochrome b gene of mammals. J. Mol. Evol. **32**: 128–144.
Kaessmann, H., V. Wiebe, and S. Paabo. 1999. Extensive nuclear DNA sequence diversity among chimpanzees. Science **286**:1159–1162.
Kaessmann, H., V. Wiebe, G. Weiss, and S. Paabo. 2001. Great ape DNA sequences reveal a reduced diversity and an expansion in humans. Nat. Genet. **27**:155–156.
Kortlandt, A. 1983. Marginal habitats of chimpanzees. J. Hum. Evol. **12**:231–278.
Mayr, E. 1942. Systematics and the origin of species. Columbia University Press, New York.
Morin, P. A., J. J. Moore, R. Chakraborty, L. Jin, J. Goodall, and D. S. Woodruff. 1994. Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science **265**: 1193–1201.
Morin, P. A., J. J. Moore, and D. S. Woodruff. 1992. Identification of chimpanzee subspecies with DNA from hair and allele-specific probes. Proc. R. Soc. Lond. B Biol. Sci. **249**: 293–297.
Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. Genetics **158**:885–896.
Ruvolo, M., D. Pan, S. Zehr, T. Goldberg, T. R. Disotell, and M. vonDornum. 1994. Gene trees and homioid phylogeny. Proc. Natl. Acad. Sci. USA **91**:8900–8904.
Schwartz, E. 1934. On the local races of the chimpanzee. Ann. Mag. Nat. Hist. Lond. **13**:576–583.
Shea, B. T., and H. J. Coolidge. 1988. Craniometric differentiation and systematics in the genus *Pan*. J. Hum. Evol. **17**: 671–685.
Stein, R., and M. Sarnthein. 1984. Late Neogene events of atmospheric and oceanic circulation offshore Northwest Africa: high resolution record from deep-sea sediments. Palaeoecol. Africa **16**:9–36.
Stone, A. C., R. C. Griffiths, S. L. Zegura, and M. F. Hammer. 2002. High levels of Y-chromosome nucleotide diversity in the genus *Pan*. Proc. Natl. Acad. Sci. USA **99**:43–48.
Tajima, F. 1983. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**:585–595.
———. 1989. Evolutionary relationships of DNA sequences in finite populations. Genetics **105**:437–460.
Vigilant, L., M. Hofreiter, H. Siedel, and C. Boesch. 2001. Paternity and relatedness in wild chimpanzee communities. Proc. Natl. Acad. Sci. USA **98**:12890–12895.
Vignaud, P., P. Duringer, H. T. Mackaye et al. (21 co-authors). 2002. Geology and palaeontology of the Upper Miocene Toros-Menalla hominid locality, Chad. Nature **418**:152–155.
Wagner, M. 1889. Die Entstehung der Arten durch räumliche Sonderung. Benno Schwalbe, Basel, Switzerland.
Wakeley, J., and J. Hey. 1997. Estimating ancestral population parameters. Genetics **145**:847–855.
Wang, R. L., J. Wakeley, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila* pseudoobscura and close relatives. Genetics **147**:1091–1106.
Wildman, D. E., M. Uddin, G. Liu, L. I. Grossman, and M. Goodman. 2003. Implications of natural selection in shaping 99.4% nonsynonymous DNA identity between humans and chimpanzees: enlarging genus *Homo*. Proc. Natl. Acad. Sci. **100**:7181–7188.
Yu, N., M. I. Jensen-Seaman, L. Chemnick, J. R. Kidd, A. S. Deinard, O. Ryder, K. K. Kidd, and W. H. Li. 2003. Low nucleotide diversity in chimpanzees and bonobos. Genetics **164**:1511–1518.

Michael Nachman, Associate Editor

Accepted October 4, 2004