Modern African Ape Populations as Genetic and Demographic Models of the Last Common Ancestor of Humans, Chimpanzees, and Gorillas

M. I. Jensen-Seaman, A. S. Deinard, and K. K. Kidd

In order to fully understand human evolutionary history through the use of molecular data, it is essential to include our closest relatives as a comparison. We provide here estimates of nucleotide diversity and effective population size of modern African ape species using data from several independent noncoding nuclear loci, and use these estimates to make predictions about the nature of the ancestral population that eventually gave rise to the living species of African apes, including humans. Chimpanzees, bonobos, and gorillas possess two to three times more nucleotide diversity than modern humans. We hypothesize that the last common ancestor (LCA) of these species had an effective population size more similar to modern apes than modern humans. In addition, estimated dates for the divergence of the Homo, Pan, and Gorilla lineages suggest that the LCA may have had stronger geographic structuring to its mtDNA than its nuclear DNA, perhaps indicative of strong female philopatry or a dispersal system analogous to gorillas, where females disperse only short distances from their natal group. Synthesizing different classes of data, and the inferences drawn from them, allows us to predict some of the genetic and demographic properties of the LCA of humans, chimpanzees, and gorillas.

The extant African apes and modern humans are the living representatives of a once more diverse radiation which contained at least several additional genera in the human lineage alone. Humans shared a common ancestor with chimpanzees and gorillas about 5–8 million years ago, but as of yet there is no fossil evidence of this ancestor. Today, two species of chimpanzees are recognized, the common chimpanzee (Pan troglodytes) and the pygmy chimpanzee, or bonobo (P. paniscus). Both species of Pan live in equatorial Africa, on either side of the Congo River. Gorillas have traditionally been considered a single species (Gorilla gorilla) (Coolidge 1929), although recently several authors have decided to split them into the western gorilla (G. gorilla) and the eastern gorilla (G. beringei) (Groves 2001). Like chimpanzees, gorillas live in equatorial Africa, but with a smaller geographic range. Many of the molecular studies involving these species over the last decade have been for the purposes of generic-level phylogeny, resulting in the generally accepted conclusion that chimpanzees and humans are closer to each other than either one is to gorillas (Satta et al. 2000). Going "beyond the phylogeny," molecular data can now be used to explore more interesting questions in our evolutionary history, and that of our closest relatives.

It has become obvious by now that examining the patterns of genetic diversity within our own species has been crucial to gaining a full understanding of the evolution of modern humans (Harpending et al. 1998). Most of the data point to an origin of our species in Africa, with an expansion into the rest of the world during the last 200,000 years, replacing the then indigenous populations (Jorde et al. 2000; Tishkoff et al. 1996; but see Hawks 2001). Although modern humans are geographically widespread and number in the billions, we show reduced genetic variation compared to the more geographically restricted African apes (Deinard and Kidd 1995, 1999, 2000; Gagneux et al. 1999a; Kaesmann et al. 2001) and have an effective population size of only about 10^4. That humans are unusual among African apes in this respect also suggests that the last common ancestor (LCA) of Homo, Pan, and Gorilla was probably much more similar to the extant apes than to modern humans. However, our understanding of the genetic structuring of modern humans far exceeds the current understanding of...
the genetic structuring of chimpanzees and gorillas, despite these species’ significantly greater genetic diversity. Here we examine the distribution of mitochondrial and nuclear autosomal genetic variation in the African apes and explore the advantages of using extant African ape populations as a model for the LCA of Homo, Pan, and Gorilla.

Materials and Methods

Nucleotide sequences of apes were obtained by amplifying the homologous fragment with polymerase chain reaction (PCR) and sequencing on an automated sequencer using primers designed from published human sequence, using procedures described elsewhere (Deinard 1997; Jensen-Seeman 2000), or were taken directly from previously published studies (Deinard and Kidd 2000; Jensen-Seeman et al., in press). Sequences were aligned manually, introducing gaps as needed. Genetic distances were estimated under the Jukes–Cantor model (Jukes and Cantor 1969) with the assistance of MEGA software (version 1.0; Kumar et al. 1993). Nucleotide positions with gaps in one or more sequences were excluded from all analyses. Measures of nucleotide diversity (π) and the parameter theta (θ) were estimated with the assistance of the Arlequin software (version 2000; Schneider et al. 2000). Effective population size (N_e) was estimated with the formula N_e = 4N_μθ, with N_μ in units of per site per year, where N_μ = (d/2t)g_L, where d is the Jukes–Cantor distance, t is the time since divergence, g is an assumed generation length of 15 years, and L is the length of the sequence. Two estimates of the mutation rate (N_μ) were obtained by calibrating sequence divergence with a Homo-Pan divergence of 6 million years ago and with a divergence between Pongo and the African ape/human clade at 15 million years ago; the average of these two estimates was used in all calculations.

Results and Discussion

Genetic Variation

At the five autosomal loci examined here, among African apes, common chimpanzees carry the greatest amount of nucleotide diversity (π), with bonobos and gorillas possessing somewhat less variation (Table 1). Numerous estimates of π in global samples of modern humans have now been generated (reviewed in Przeworski et al. 2000; Wall and Przeworski 2000; Yu et al. 2001), with consistent estimates falling near the early value of 0.11% as estimated by Li and Sadler (1991). Although far fewer loci have been examined in the great apes, the data presented here-in suggest that modern African apes possess two to three times more nucleotide diversity than do modern humans. There now seem to be sufficient DNA sequence data from both the mitochondrial and nuclear genomes (Deinard and Kidd 1995, 1999, 2000; Gagneux et al. 1999a; Kaessmann et al. 2001) to state that this greater amount of diversity is a general feature of the African apes, and that earlier reports to the contrary based on blood protein polymorphisms or microsatellites can most easily be explained as a result of ascertainment bias (Lucotte 1983; Wise et al. 1997). If we accept that humans are unusual in their low level of nucleotide diversity; and that the LCA of humans, chimpanzees, and gorillas was undoubtedly much more like modern apes than modern humans in terms of its morphology, ecology, population structure, and demography, then it is reasonable to assume that this LCA had levels of nucleotide diversity more on a par with modern apes than modern humans.

Correspondingly, the effective population size (N_e) of modern African apes as estimated from the relationship θ = 4N_μθ is several times greater than that of modern humans (Table 2). Our estimates are similar to that of Kaessmann et al. (1999), who estimated the N_e of common chimpanzees to be about 35,000. In comparison, multiple studies of modern human genetic variation place the N_e of humans at about 10^4 (e.g., Yu et al. 2001; Zhao et al. 2000). Consequently our estimates for the N_e of modern African apes are, in general, two to five times higher than for modern humans, with most values closer to two or three times greater. Again, assuming that the LCA of African apes and humans was more similar to modern apes than to modern humans in most respects, the effective population size of this LCA was also likely to be several times greater than the estimate calculated for modern humans. Such a finding is intriguing, given that modern humans currently have a census population size of more than 6 billion and growing rapidly, while the African apes have shrinking populations several orders of magnitude smaller. The paradoxically small N_e in humans is consistent with our species having recently greatly expanded from a much smaller population (Harpending et al. 1998).

Our results also show that at most of the loci examined, common chimpanzees have greater π and larger N_e than bonobos or gorillas. Relative to common chimpanzees, bonobos live in a more restricted geographic range and are less subdivided than chimpanzees, who are divided into at least three well-recognized subspecies (Fleagle 1999). Bonobos have been isolated south of the Congo River at least since its inception about 1.5 million years ago (Beadle 1981). The area of suitable habitat available to bonobos may always have been smaller than that of their congeners to the north, especially during periodic contractions of the tropical forests of equatorial Africa, when much of the forest cover south of the Congo River was substantially reduced (Hamilton 1981). Perhaps during these times the bonobo population passed through bottlenecks, resulting in the smaller N_e we see today in these apes. Long-term effective population size can be strongly affected by intermittent bottlenecks, even if the population is large more often than not (Li 1997). There seem to be little if any social or reproductive factors that could give chimpanzees a larger N_e than bonobos: bonobos have a more equal adult sex ratio in their com-

---

Table 1. Nucleotide diversity (π; as a percentage) among loci and species⁴⁻⁶

<table>
<thead>
<tr>
<th>Locus</th>
<th>ADH1 (575–606 bp)</th>
<th>APOB (1060–1064 bp)</th>
<th>DRD2 (301 bp)</th>
<th>DRD4 (1109–1134 bp)</th>
<th>HOXB (833–843 bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>0.0968</td>
<td>0.0622</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chimpanzee</td>
<td>0.2468</td>
<td>0.321</td>
<td>0.2039</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonobo</td>
<td>0.1133</td>
<td>0.2123</td>
<td>0.2115</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gorilla</td>
<td>0.1403</td>
<td>0.2069</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

⁴ Loci are either intronic (ADH1, APOB, DRD2) or intergenic (DRD4, HOXB).
⁵ Length of sequence is given in parentheses below each locus symbol as a range when length varied among species.
⁶ Sample size (minimum number of independent chromosomes, e.g., exact number unknown for half sibs) is given in parentheses below each value of π.

—

476 The Journal of Heredity 2001:92(6)
munities (Kano 1996) and chimpanzee females are probably not more promiscuous than bonobo females (Takahata et al. 1996; but see Kano 1992).

In contrast, there are many social and ecological differences between chimpanzees and gorillas that may result in the lower $N_e$ of gorillas. Gorillas live in smaller social groups, which typically contain one or two adult males who monopolize access to the several females in the group (Robbins 1995; Watts 1996), whereas chimpanzee females mate promiscuously within the social group (Takahata et al. 1996). Species with strong polygyny, where one male may mate with many females, leaving many other males reproductively unsuccessful, will have a reduced male effective population size, and therefore a lower overall $N_e$. In addition, chimpanzees and gorillas differ ecologically in that gorillas are more restricted to forest habitats, while chimpanzees can live in a wider range of habitats including forest, open woodland, and savanna (Kortland, 1983; Yamagiwa 1999). The more limited ecological tolerance of gorillas may have made them more susceptible to past environmental changes, sending them through several population bottlenecks, resulting in a smaller long-term effective population size. Recently it has been suggested that East African gorilla populations have gone through large population bottlenecks associated with forest reduction during the last glacial maximum (Jensen-Seaman and Kidd, 2001), although it is unclear if such population bottlenecks have occurred periodically throughout the Pliocene and Pleistocene.

Although our estimates of the $N_e$ of African apes are several times larger than that of modern humans, they are several times smaller than Chen and Li’s (2001) $N_e$ estimates for the ancestral population of hominids that eventually gave rise to humans, chimpanzees, and bonobos. Chen and Li (2001) used the amount of sequence divergence between Homo, Pan, and Gorilla, and the percentage of loci favoring one or the other of the three alternative phylogenies of these genera, to estimate the $N_e$ of the ancestral human-chimpanzee population under a model developed by Wu (1991). They estimated the $N_e$ of the human-chimpanzee ancestor to be between 52,000 and 96,000, or between 5 and 10 times that of modern humans (Chen and Li 2001). Therefore, if these values—as well as our values of $N_e$ of modern African ape populations—are even approximately accurate, then the ancestral human-chimpanzee population may have had an effective population size several times larger than even modern chimpanzees. Of course, all methods of estimating $N_e$ are sensitive to several assumptions, including mutation rate and generation length, so it is possible that either the values of Chen and Li (2001) or ours are inaccurate. Nonetheless, if we accept for the moment that the human-chimpanzee LCA had a larger effective population size than modern chimps, we can suggest several scenarios which could produce such a result. One possibility is that modern ape populations may have gone through several more population bottlenecks in the last few million years than did the ancestral hominid population prior to its splitting into the Homo and Pan lineages, perhaps due to a reduction in suitable habitat. During the climatic vicissitudes of the Pleistocene, the extent of African tropical forest cover was dramatically reduced, concomitant with the global cooling and drying associated with glaciations in the temperate regions (Hamilton 1981). Another scenario under which the human-chimpanzee ancestral population may have had a larger $N_e$ than any of the present African ape species is that it may have had a more equal ratio of effective population sizes of males and females, which can occur under nonpolygynous mating systems, such as monogamy or high female promiscuity.

### Table 2. Estimates of $\mu$, $\theta$, and $N_e$

<table>
<thead>
<tr>
<th></th>
<th>ADH</th>
<th>APOB</th>
<th>DRD2</th>
<th>DRD4</th>
<th>HOXB</th>
<th>Average*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^a$ (per site per year)</td>
<td>$1.1 \times 10^{-4}$</td>
<td>$1.52 \times 10^{-4}$</td>
<td>$1.07 \times 10^{-4}$</td>
<td>$1.70 \times 10^{-4}$</td>
<td>$7.6 \times 10^{-5}$</td>
<td>$1.23 \times 10^{-5}$</td>
</tr>
<tr>
<td>$\mu^b$ (per sequence per generation)</td>
<td>$9.30 \times 10^{-3}$</td>
<td>$2.41 \times 10^{-2}$</td>
<td>$4.81 \times 10^{-2}$</td>
<td>$2.83 \times 10^{-2}$</td>
<td>$9.55 \times 10^{-3}$</td>
<td>$n/a$</td>
</tr>
</tbody>
</table>

* Average values are calculated using only nonzero values.

* $\mu$ is the average of two estimates: one assumes a Homo-Pan divergence of 6 million years ago, the other assumes a Pongo-Hominid split of 15 million years ago.

* $\theta$ is Watterson’s (1975) estimate of the mutational parameter $\theta$.

* $\theta$ is Tajima’s (1983) estimate of $\theta$.

* $N_e$ is estimated by the formula $N_e = 4\mu L$, with $\mu$ in units of per site per year, where $d$ is the Jukes–Cantor distance, $t$ is the time since divergence, $g$ is the average of two estimates: one assumes a Homo-Pan divergence of 6 million years ago, the other assumes a Pongo-Hominid split of 15 million years ago.

### Human

<table>
<thead>
<tr>
<th>$\theta_e$</th>
<th>$\theta_t$</th>
<th>$N_e$ ($\theta_e$)</th>
<th>$N_e$ ($\theta_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>18510</td>
<td>15914</td>
</tr>
</tbody>
</table>

### Bonobo

<table>
<thead>
<tr>
<th>$\theta_e$</th>
<th>$\theta_t$</th>
<th>$N_e$ ($\theta_e$)</th>
<th>$N_e$ ($\theta_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>16837</td>
<td>15388</td>
</tr>
</tbody>
</table>

### Gorilla

<table>
<thead>
<tr>
<th>$\theta_e$</th>
<th>$\theta_t$</th>
<th>$N_e$ ($\theta_e$)</th>
<th>$N_e$ ($\theta_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1696</td>
<td>14664</td>
</tr>
</tbody>
</table>

Jensen-Seaman et al. • Ape Models of Human Ancestors 477
Mitochondrial versus Nuclear DNA

After much debate and collection of substantial data, a near-universal consensus has been reached among molecular systematists that of the three genera of extant African apes, humans and chimpanzees are each other's closest relatives, to the exclusion of gorillas. However, several aspects of these divergences remain unclear. For example, the date of the split of gorillas from the human-chimpanzee ancestor has been placed at anywhere from 18 (Arnason et al. 1996) to 2.7 million years ago (Hasegawa et al. 1985); both of these dates come from mtDNA sequence data. Another uncertainty concerns the amount of time between the divergence of gorillas and the subsequent split between the lineages leading to humans and chimpanzees. Several estimates of this “internode” distance were assembled from published data, and are expressed as a percentage of the time of the human-chimpanzee divergence from the same study (Figure 1a). This internode distance has been estimated from mtDNA to be 33–64% (Arnason et al. 1996; Gagneux et al. 1999a; Hasegawa et al. 1985; Horai et al. 1992). Recently this percentage was estimated from 53 non-coding intergenic nuclear DNA regions to be about 35% (Chen and Li 2001), and as low as 16% from the synonymous substitutions at a large number of protein-coding autosomal loci (Satta et al. 2000). Obviously there is substantial variation among different studies, even using the same genomic system, but in general, estimates from mtDNA give much larger estimates of the time between the gorilla split and the later human-chimpanzee split.

An analogous situation exists among the gorillas, where the date estimated for the divergence between the western gorilla (G. gorilla) and the eastern gorilla (G. beringei) differ substantially, depending upon whether they were generated from mtDNA or independent noncoding nuclear DNA loci (Figure 1b; Jensen-Seaman et al., in press). Whereas several studies utilizing mtDNA have shown the amount of genetic distance between the eastern and western gorillas to be as great or greater than that between chimpanzees and bonobos (Garner and Ryder 1996; Jensen-Seaman et al., in press; Ruvolo et al. 1994), results from the nuclear loci summarized here reveal that the distance between the species of Pan is nearly twice as great as that between chimpanzees and bonobos (Garner and Ryder 1996; Jensen-Seaman et al., in press). An independent study of an X-chromosomal locus confirms the much smaller divergence between the nuclear genomes of the Gorilla species when compared to Pan species (Kaessmann et al. 2001). The significance of this observation can be appreciated if one were to imagine a future molecular systematist sampling the descendants of these species 5 million years from now: if no further gene flow were to occur between these species, the systematist would be forced to estimate very different divergence times for the gorilla species based on mitochondrial compared to nuclear DNA data, analogous to the situation seen in estimating the time between the Gorilla split from Homo-Pan and the subsequent split between these latter two genera.

Perhaps if we can explain the causes of this pattern seen in modern gorillas, we can infer what may have produced the analogous situation in the last common ancestor of the extant African apes and humans. There are several possible explanations for this discrepancy between mtDNA and nuclear DNA in the estimates of gorilla divergence. First, not all loci are expected to give the same results, both in terms of tree topology, and in divergence dates because of allelic sorting of ancestral polymorphisms and the stochastic nature of the molecular clock (Lanyon 1988; Pamilo and Nei 1988; Rogers 1994). Therefore the most accurate estimates will come from examining a large number of independent loci; however, this is not pos-

![Figure 1](image-url)
sible for mtDNA since it is inherited as a single nonrecombining locus. While the higher mutation rate for mtDNA should reduce the stochastic aspects of the clock, the fourfold smaller effective population size will increase the stochastic aspects of drift.

Another explanation is that sex-biased dispersal and migration resulted in an ancestral gorilla population with strongly geographically structured mtDNA diversity but essentially homogenous panmictic structure to the nuclear DNA. Then, when this population split in two, the resulting mitochondrial lineages had already diverged well before the nuclear genomes. In general, this discordance between mtDNA and nuclear DNA will most likely be found in species with extremely strong female philopatry, like some Old World monkeys, where females always remain in their natal troops and males always disperse (Melnick and Hoelzer 1992; Melnick et al. 1993). In contrast, all great ape species exhibit female transfer (Watts 1996). However, there is variation among species of extant African apes in this regard. Chimpanzee females uniformly disperse from their natal group, while males remain—all though the possibility of substantial extra-group paternity in chimpanzees (Gagneux et al. 1997; 1999b) could impact strongly on the resulting patterns of diversity, since this has the same genetic effect as male dispersal; however, the extent to which this actually happens may be quite small or nonexistent (Constable et al. 2001; Vigilant and Boesch 2001). In contrast, both male and female gorillas may leave their natal group, with the difference being that when females disperse they tend to travel very short distances to a neighboring group, while males may wander long distances for several years before establishing their own social group (Harcourt 1978), although of course some males do remain in their natal group (Watts 2000). This difference between the sexes in gorillas could lead to the differential structuring of mitochondrial and nuclear DNA variation discussed above.

Therefore we may suggest a scenario for the successive splits of the last common ancestral population into the three extant genera of African apes we see today. First, we propose that the last common ancestral population of humans, chimpanzees, and gorillas was geographically widespread, with strong subdivision of its mtDNA variation, while exhibiting much more panmictic distribution of its nuclear DNA variation. This was possibly the result of sex-biased dispersal, with either strong female philopatry and male dispersal like many Old World monkeys, or a gorilla-like dispersal of both sexes but with short-distance female dispersal and potentially long-distance male dispersal. Following the split of gorillas, the ancestral Homo-Pan population may have shifted to a more chimpanzee-like dispersal pattern, with widespread long-distance female gene flow, which would eliminate the previous differences between mtDNA and nuclear DNA. Modern chimpanzee populations have been shown to share mtDNA D-loop haplotypes across distances exceeding 900 km (Goldberg and Ruvolo 1997; Morin et al. 1994). Perhaps this demographic shift in the ancestral population of humans and chimpanzees was concomitant with a socioecological shift, such as larger group size or an increased reliance on vertebrate meat, that modern humans and chimpanzees share. Finally, when this population split into the lineages leading to modern humans and modern chimpanzees, the distributions of mtDNA and nuclear DNA were similar, and so today we do not observe any large differences in the estimates of this split between mitochondrial and nuclear DNA. If such a shift toward a more “chimp-like” social structure and mating behavior occurred in the human-chimpanzee ancestor, it could also help explain the large $N_e$ of this population as inferred by Chen and Li (2001).

**Conclusion**

Relative to the living African apes, modern humans possess unusually small amounts of nucleotide diversity and have an effective population size several times smaller than that of our closest relatives. Of the extant African apes, common chimpanzees carry the largest amount of nuclear genetic diversity. The LCA of humans, chimpanzees, and gorillas likely also had a much larger $N_e$ than modern humans. Other studies suggest that the LCA of humans and chimpanzees may have been several times larger then even modern chimpanzees. At the least, this suggests that this human-chimpanzee ancestral population may have shared some features with modern chimpanzees that give them a larger $N_e$ compared to gorillas. For example, the larger group size, the more equal male to female effective population size, and the increased tolerance to ecological variability of chimpanzees relative to gorillas may be partly responsible for their larger $N_e$ than gorillas. If the LCA population of humans and chimpanzees had an $N_e$ several times larger than modern chimpanzees, then perhaps this population was unusually large and widespread, and had a nonpolygonous mating system, or the ancestral populations of modern chimpanzees and gorillas may have experienced intermittent population bottlenecks during the last 5 million years.

The difference in estimates of the length of the internode between the splitting of gorillas from the human-chimpanzee ancestor, and the splitting of these latter two species, whether using mtDNA or nuclear DNA remains an uncertainty in understanding late Miocene hominid evolution. An interesting parallel is seen in gorillas, where it is hypothesized that the different patterns seen from mtDNA versus nuclear DNA with regard to the timing of the split between eastern and western gorillas is due to the different geographical structuring of the genetic diversity in the LCA of these gorilla species, due in turn to the differences in male versus female dispersal patterns. Therefore we suggest that the LCA of humans, chimpanzees, and gorillas may also have had strongly sex-biased dispersal with female philopatry or female gene flow restricted to short distances. Following the split from gorillas, the LCA of humans and chimpanzees may have shifted toward increased female vagility, which may have coincided with the shift to a mating system with equal male and female effective population sizes.

Many of these hypothetical scenarios and descriptions of the LCA of humans and the extant African apes, including its social structure, dispersal patterns, and mating system, are highly speculative. Such speculation is important in these early stages of developing a comprehensive and comparative view of the amounts and patterns of genetic diversity in all African ape species. Understanding these aspects of our extant relatives is essential to a full understanding of our own evolution. Further examination of the diversity within the extant apes, both morphologically and genetically, will enable us to better understand the LCA not as a single fossil or single nucleotide sequence, but rather as a diverse and dynamic population.

**References**


Received November 20, 2001

Corresponding Editor: Oliver A. Ryder