VIVAMUS

Biogeographic Predictors of Genetic Diversity in Populations of Eastern African Chimpanzees (*Pan troglodytes schweinfurthi*)

Tony L. Goldberg[1,2]

I collected mitochondrial DNA sequences—hypervariable region 1 of the control region—from 281 eastern chimpanzees in 19 geographically defined populations and calculated genetic diversity measures to test the hypothesis that populations inhabiting the reconstructed locations of Pleistocene forest refugia harbor higher genetic diversities than those of other populations. The hypothesis is only weakly supported. Population genetic diversity is not significantly correlated with geographic proximity to refugia, with the area of forest that the populations currently occupy, or with the degree of geographic isolation of the populations. However, the two populations displaying the consistently highest genetic diversities are located in refuge areas: Uganda's Rwenzori Mountains and the eastern Democratic Republic of Congo's Ituri Forest. These results, in combination with previously findings, imply that chimpanzees may have lived both in and out of refugia during periods when tropical forests were restricted to refugia. This interpretation is consistent with the notion of chimpanzees as an extraordinarily vagile species, capable of maintaining gene flow across habitat mosaics of forest, woodland, and savannah.

[1]Department of Anthropology, Harvard University, Cambridge, Massachusetts 02138.
[2]Current address Department of Veterinary Pathobiology, University of Illinois, College of Veterinary Medicine, 2001 South Lincoln Avenue, Urbana, Illinois 61801. e-mail: tlgoldbe@uiuc.edu.

INTRODUCTION

The population history of the easternmost subspecies of chimpanzees, *Pan troglodytes schweinfurthi*, is intimately linked to the history of the forests which it principally occupies. Although chimpanzees also inhabit marginal, nonforest environments (McGrew *et al.*, 1981; Kortlandt, 1983; Moore, 1992), they depend for survival on forest foods and standing water (Kortlandt, 1983; Wrangham *et al.*, 1993). Chimpanzees do not range far beyond the limits of the African rain forest.

The distribution and extent of African forest have, however, varied considerably throughout the Pleistocene. Specifically, paleoclimatic cycles of global warming and cooling have caused forests to expand and contract correspondingly (Kendall, 1969; van Zinderen Bakker and Coetzee, 1972; Moeyersons and Roche, 1982). During cold, dry periods, forests were restricted to areas where local conditions allowed their persistence despite general climatic aridity. Forest taxa presumably dispersed out of these refugia to their current distributions when the climate became warmer and wetter (Hamilton, 1976; Grubb, 1982; Hamilton, 1988).

Since eastern chimpanzees diverged from central African chimpanzees (*Pan troglodytes troglodytes*) approximately 440 Ka (1000 years ago) [confidence interval = 0.21 to 1.64 Ma (million years ago) (Morin *et al.*, 1994)], the eastern subspecies must have experienced several periods of minimal forest cover, the most recent of which occurred approximately 18 ka (Hamilton, 1988). At such times, montane refugia existed at high-altitude locations on the western edge of the African Rift (Coetzee, 1964; Osmaston, 1967; Taylor, 1993). Lowland forest refugia were present in the northeastern Democratic Republic of Congo (D.R. Congo, formerly Zaïre; Hamilton, 1976; Grubb, 1982) and probably consisted of a series of interconnected fluvial refugia (Colyn *et al.*, 1991). Primate species diversity and endemism decline with distance from these areas today (Struhsaker, 1981; Rodgers *et al.*, 1982).

To the extent that chimpanzees have been restricted to forest habitats, their historical distribution should have tracked changes in the distribution of forest cover. During warm, wet phases, chimpanzees should have been widespread, distributed approximately throughout their current range. The onset of glacial conditions should have led to the vicariance of chimpanzee habitat. During periods of maximal aridity, chimpanzees should have been restricted to areas in and close to forest refugia. As forests reexpanded during postglacial climatic amelioration, chimpanzees should have dispersed out of these refugia into increasingly more distant areas.

This study uses genetic data to test these hypotheses. Pleistocene refuge theory, as outlined above, predicts that localized populations of eastern chimpanzees should have experienced different degrees and kinds of bottleneck effects. Bottlenecks reduce the genetic variability of populations that experience them (Nei *et al.*, 1975; Chakraborty and Nei, 1977; Watterson, 1984). Furthermore, the effects of a bottleneck persist over many generations following population reexpansion (Nei *et al.*, 1975). Low levels of genetic variability across a range of species have been attributed to bottlenecks caused specifically by climatically induced population size contractions (Nei and Graur, 1984; Sage and Wolff, 1986).

Eastern chimpanzee populations inhabiting refugia should have been buffered from drastic size reductions. Populations currently inhabiting locations corresponding to Pleistocene refugia should therefore harbor greater genetic diversities than should nonrefuge populations. Moreover, the genetic diversity of a population should be a negative function of its geographic distance from a refuge. This is because founder effects (Mayr, 1954), leading to potential reductions in genetic diversity, would have accompanied the colonization of nonrefuge forests. Since the most geographically distant forests should have been colonized though a stepping stone series of sequential founder events, they should harbor the lowest genetic diversities. An incremental decline of genetic variability should thus occur with increasing geographic distance from refugia.

Also, the genetic diversity of a chimpanzee population should be positively correlated with the area of the forest that it occupies, assuming that the carrying capacities of all forests are similar. This is because a population's genetic diversity is largely a function of its long-term effective size (Wright, 1931). Finally, because of the diversifying effects of migration, a positive correlation should exist between genetic diversity and geographic proximity to other populations.

Genetic studies of other taxa have revealed such island effects (Baker *et al.*, 1990; Gilbert *et al.*, 1990; Inoue and Kawahara, 1990; Edwards, 1993; Seutin *et al.*, 1993, 1994; Westerbergh and Saura, 1994). I investigated whether similar trends are evident in populations of *Pan troglodytes schweinfurthi* inhabiting insular forests separated by open habitat.

MATERIALS AND METHODS

I collected hair from chimpanzee sleeping nests encountered during ground searches in 19 natural populations (Fig. 1). I also sampled 26 chimpanzees of unknown provenance from a captive population housed at Entebbe Zoo, Uganda. Names of sampling locations are in Table I, along
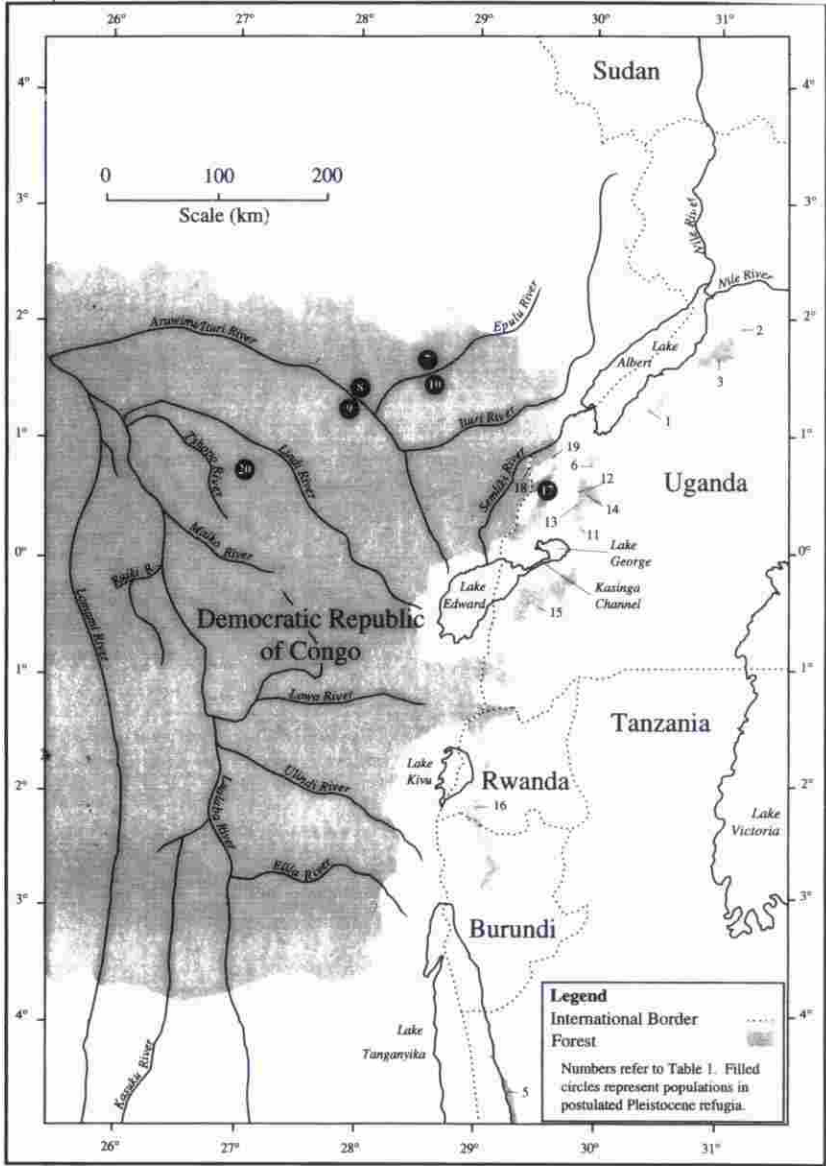
Fig. 1. Map of sampling locations.

with the biogeographic variables used in the study. I use abbreviated location names (Table I) throughout this study. On the basis of independent

Table I. Description of Sampling Locations[a]

	Abbrev.	Forest	Location	Area (km^2)[b]	Distance to the "source" (km)[c]	Isolation (km)[d]
1.	BAMA	Bugoma	Mwela Sawmill	365	130	233
2.	BOPI	Budongo	Pabidi	793	238	319
3.	BOSO	Budongo	Sonso Sawmill	793	209	293
4.	EEZO		Entebbe Zoo			
5.	GEKA	Gombe	Kasekela	32	583	617
6.	IARA	Itwara	Rwebikuya	87	45	184
7.	IIAA	Ituri	Afarama	470,000	0	257
8.	IIAE	Ituri	Avakubi NE	470,000	0	312
9.	IIAW	Ituri	Avakubi SW	470,000	0	313
10.	IILA	Ituri	Lenda	470,000	0	242
11.	KEDN	Kibale	Dura Station	560	44	189
12.	KEKA	Kibale	Kanyawara	560	28	178
13.	KEKU	Kibale	Kanyanchu	560	26	179
14.	KENO	Kibale	Ngogo	560	36	181
15.	KUSL	Kalinzu	Kalinzu Sawmill	580	111	227
16.	NESN	Nyungwe	Nyungwe Station	1140	335	389
17.	RIKA	Rwenzori	Katebwa	996	0	176
18.	SIMU	Semliki	Mbume-Busaru	212	23	177
19.	SINI	Semliki	Ntandi	212	31	178
20.	TOBA	Tshopo	Bafwabalinga	470,000	0	353

[a]See Fig. 1, where sampling locations are shown by number.
[b]Estimates of present-day areas of Ugandan, Tanzanian, and Rwandan forests are from Weber (1989) and Howard (1991); area estimate of D.R. Congo forest (470,000) is from Teleki (1989).
[c]RIKA was defined as the source for all non-refuge forests because it is the refuge location closest to these forests.
[d]Isolation is measured as the mean distance of a population from all other populations in the study.

biogeographic evidence, populations considered to be refuge populations are all D.R. Congo populations [Nos. 7–10 and 20 (Hamilton, 1976; Grubb, 1982; Colyn, 1991)] and Rwenzori/Semliki Forest populations in Uganda [Nos. 17–19 (Struhsaker, 1981; Rodgers et al., 1982; Howard, 1991)]. The forests that they inhabit contain exceptionally high specific diversities and high numbers of endemic taxa.

I examined a 368-bp segment of the mitochondrial control region corresponding to Anderson reference sequence coordinates 16042–16410, which includes the first hypervariable region (Anderson et al., 1981), which is the most quickly evolving region in primate mitochondrial genomes (Kocher and Wilson, 1991). I extracted DNA from hair follicles and PCR amplified and sequenced it. Specific laboratory methods are described by Goldberg (1996).

RESULTS

I generated 255 complete DNA sequences. They and 19 sequences from Morin *et al.* (1994) yielded 123 unique haplotypes, with a total of 90 variable nucleotide positions. Of 123 haplotypes, 99 are population-specific, and 24 are shared among populations. The newly generated sequences are available through GenBank (accession numbers U77181–U77293).

Because individual subjects were not observed constructing nests, repeated sampling of individuals within sampling locations is possible. However, a separate analysis of the spatial distribution of identical haplotypes within sampling locations indicates that double sampling was not, in fact, a major problem (Goldberg and Wrangham, 1997). The analyses which follow operate under this assumption.

Genetic Diversity

I calculated five genetic diversity measures for each population: the Shannon–Wiener Diversity Index (Lewontin, 1972), Nei's (1973) gene diversity, Nei's (1987) nucleotide diversity (mean pairwise sequence difference between haplotypes), modal pairwise sequence difference between haplotypes, and maximum pairwise sequence difference between haplotypes. The first two measures assume equidistance among haplotypes; the last three incorporate information about sequence-level genetic distances. Results are in Table II.

Genetic diversities are no greater for western forest populations (i.e., those in the D.R. Congo; Nos. 7–10 and 20 in Table I) than for eastern forest populations by any measure (independent t tests with 17 df; $p >$ 0.05). I obtained identical results when comparing refuge populations— western locations plus Rwenzori and Semliki Forests in Uganda—and nonrefuge populations. However, these results should be considered with caution. Sample sizes within populations are small, making sampling error a problem of some concern. Furthermore, the underlying distributions of the diversity statistics are not fully known. The results did not change when I used nonparametric statistical tests (Wilcoxon two-sample tests) in place of t tests.

Endemism

Biogeographically, a taxon is endemic if it exists exclusively within a restricted geographic range (Major, 1988). Taxa with cosmopolitan distributions are eurytopic. I applied the concept of endemism to genetic data

Table II. Measures of Genetic Diversity for Individual Sampling Locations

Location	n^a	Diversity measure[b]				
		SW	GD	ND	MD	XD
1. BAMA	13	1.29	0.64	0.85	0.00	2.65
2. BOPI	12	1.94	0.91	1.61	1.78	3.56
3. BOSO	15	1.97	0.89	1.55	2.04	2.62
4. EEZO	26	3.04	1.02	2.04	1.70	4.80
5. GEKA	19	2.31	0.94	2.02	2.29	3.73
6. IARA	12	1.94	0.91	1.55	1.19	3.85
7. IIAA	13	2.20	0.95	2.97	3.24	5.00
8. IIAE	12	1.23	0.67	1.21	0.00	2.96
9. IIAW	13	1.99	0.92	2.25	2.36	3.83
10. IILA	17	2.28	0.94	1.87	1.88	3.75
11. KEDN	14	1.30	0.73	1.97	2.93	3.51
12. KEKA	15	1.84	0.89	0.97	1.16	1.75
13. KEKU	14	1.77	0.82	2.22	3.22	3.80
14. KENO	13	1.48	0.81	2.06	0.88	4.42
15. KUSL	13	2.35	0.97	1.59	1.47	4.12
16. NESN	13	1.88	0.91	1.87	2.06	2.94
17. RIKA	13	2.35	0.97	2.70	3.24	5.30
18. SIMU	13	1.26	0.69	1.99	0.00	4.12
19. SINI	11	1.85	0.90	2.27	2.09	3.89
20. TOBA	10	1.83	0.91	1.84	1.81	3.32
Eastern forests[c]	171	3.96	0.98	2.04	1.91	5.47
Western forests[c]	65	3.54	0.98	2.16	1.93	4.97
Total	281	4.43	0.98	2.08	1.91	5.45

[a]Sample sizes of DNA sequences.
[b]SW, Shannon–Wiener index; GD, gene diversity; ND, nucelotide diversity (%); MD, modal pairwise sequence difference (%) for all pairs of individuals; XD, maximum pairwise sequence difference (%) for all pairs of individuals.
[c]The eastern forest sample was created by combining haplotypes from Uganda, Rwanda, and Tanzania (locations 1–3, 5, 6, and 11–19). The western forest sample was created by combining haplotypes from locations in the D.R. Congo (7–10 and 20). EEZO and GEKA haplotypes are excluded from the analysis. EEZO animals are of unknown geographic origin, and GEKA haplotypes contain missing data which would have biased diversity estimates.

at the allelic level: a haplotype is endemic if it appeared exclusively in one sampling location. If a haplotype appeared in more than one sampling location, it is eurytopic. Endemic alleles probably occur at low frequencies in populations in which they were not sampled. Similarly, eurytopic alleles are probably shared among more sampling locations than the data suggest. Endemism—the proportion of endemic haplotypes in a population—is therefore a sample size-dependent diversity measure which emphasizes low-frequency alleles.

Table III lists numbers of endemic and eurytopic haplotypes for each sampling location and the frequency distributions of endemic haplotypes.

Table III. Numbers of Endemic and Eurytopic Haplotypes Identified in Individual Sampling Locations

Location	Sample size (n)	Number of haplotypes	Number of endemic haplotypes[a]	Number of eurytopic haplotypes[b]	Frequency distribution of endemic haplotypes[c]
1. BAMA	13	6	2	4	1,1
2. BOPI	12	8	4	4	1,1,1,1
3. BOSO	15	9	4	5	1,1,1,1
4. EEZO	26	22	11	11	1,1,1,1,1,1,1,1,1,1,1
5. GEKA	19	11	10	1	3,3,2,2,2,2,1,1,1,1
6. IARA	12	8	4	4	1,1,1,1
7. IIAA	13	10	10	0	3,2,1,1,1,1,1,1,1,1
8. IIAE	12	5	5	0	7,2,1,1,1
9. IIAW	13	7	6	1	2,2,2,1,1,1
10. IILA	17	11	10	1	3,3,2,1,1,1,1,1,1,1
11. KEDN	14	5	2	3	5,1
12. KEKA	15	6	1	5	1
13. KEKU	14	8	0	8	0
14. KENO	13	5	1	4	3
15. KUSL	13	11	8	3	1,1,1,1,1,1,1,1
16. NESN	13	7	7	0	3,2,2,2,2,1,1
17. RIKA	13	11	3	8	2,1,1
18. SIMU	13	5	3	2	1,1,1
19. SINI	11	7	3	4	2,1,1
20. TOBA	10	7	6	1	3,2,1,1,1,1
Eastern forests[d]	190	75	55	20	
Western forests[d]	65	41	41	0	
Total	281	123	99	24	

[a]Endemic haplotypes appear exclusively in one sampling location.
[b]Eurytopic haplotypes appear in more than one sampling location.
[c]Frequency distributions show absolute numbers of each distinct endemic haplotype in each sampling location.
[d]Numbers of endemic and eurytopic haplotypes were calculated separately among populations within each region. Haplotypes from EEZO population are excluded from the analysis.

Of the 123 haplotypes 99 (80%) are endemic. Endemism is higher in western locations than in eastern locations. Within western locations, all haplotypes are endemic. In other words, no haplotype is shared among locations within the D.R. Congo. In contrast, 73% of eastern haplotypes are shared among eastern locations. Four haplotypes (3.3%) are shared between western and eastern locations.

I quantified endemism for each location as the percentage of haplotypes that was found only there. Endemism ranged from 20 to 100% for the eastern forest sample and from 86 to 100% for the western sample. Endemism within western forests is significantly greater than endemism

within eastern forests (independent t test with 17 df; $p = 0.003$). Within the five western locations, the number of endemic haplotypes is greater than the number of eurytopic haplotypes (paired t test with 4 df; $p = 0.004$). Within the 14 eastern populations, however, the number of endemic haplotypes is not statistically different from the number of eurytopic haplotypes (paired t test with 13 df; $p = 0.87$).

The number of endemic haplotypes is consistently low for locations within Kibale Forest (Nos. 11–14). This is probably an artifact of uneven geographic sampling, since Kibale was sampled on a finer geographic scale than other forests were. To correct this bias, I collapsed the Kibale sample into a single Kibale Forest population. The number of endemic haplotypes within eastern forests is still not significantly greater than the number of eurytopic haplotypes (paired t test with 10 df; $p = 0.72$). Endemism within western forests is still higher than endemism within eastern forests (independent t test with 14 df; $p = 0.004$). The results described above are identical when I used nonparametric statistical tests in place of t tests.

Biogeographic Predictors of Diversity and Endemism

I obtained areas of eastern forests (Table I) from Howard (1991) and Weber (1987). Teleki (1989) estimated the area of forest in the D.R. Congo suitable for eastern chimpanzees as 470,000 km². For the purposes of this analysis, I defined the area of all western locations as equivalent to the area of the largest eastern forest (Nyungwe; 1140 km²) to preserve approximate normality in the distribution of areal measurements. Using Teleki's area estimate directly did not change the results described below.

For each sampling location, I calculated a distance to the nearest source (refuge) population. Distances from refuge populations to the source are defined as zero. I measured distances from nonrefuge eastern populations to the source as great-circle distances to the Rwenzori population (RIKA), based on an estimated diameter of the Earth of 12,756 km (Skinner and Parker, 1987). I chose RIKA because it is the closest refugium to all nonrefuge forests (Hamilton, 1976; Struhsaker, 1981; Rodgers *et al.*, 1982). Populations also varied with respect to their geographic centrality within the study area. Therefore, I quantified isolation as the mean great-circle distance of a population to all other populations in the study (Table I).

I calculated multiple linear regressions for the relationship between Shannon–Wiener diversity and forest area, distance to the source, and isolation. Shannon–Wiener diversity shows the predicted positive relationship to each of the geographic variables. However, in no case is this relationship

statistically significant (area $r^2 = 0.04$, $p = 0.43$; distance $r^2 = 0.07$, $p = 0.29$; isolation $r^2 = 0.08$, $p = 0.23$). I calculated similar regressions for genetic endemism. The relationship between endemism and area is positive but not significant ($r^2 = 0.21$, $p = 0.07$). Contrary to the expectations of Pleistocene refuge theory, endemism correlates positively with both distance to the source and isolation. Although the relationship between endemism and distance is not significant ($r^2 = 0.03$, $p = 0.52$), the relationship between endemism and isolation is significant ($r^2 = 0.34$, $p = 0.02$).

I calculated multiple linear regressions between the four remaining genetic diversity measures and forest area, distance to the source, and isolation. The relationship between area and all four measures is in the predicted positive direction but is not significant in any case (gene diversity $r^2 = 0.05$, $p = 0.35$; nucleotide diversity $r^2 = 0.05$, $p = 0.38$; modal sequence difference $r^2 = 0.07$, $p = 0.27$; maximum sequence difference $r^2 = 0.00$, $p = 0.86$). There is no consistent trend between any genetic diversity measure and geographic distance (gene diversity $r^2 = 0.05$, $p = 0.36$; nucleotide diversity $r^2 = 0.00$, $p = 0.90$; modal sequence difference $r^2 = 0.00$, $p = 0.81$; maximum sequence difference $r^2 = 0.03$, $p = 0.45$), or between any diversity measure and geographic isolation (gene diversity $r^2 = 0.06$, $p = 0.33$; nucleotide diversity $r^2 = 0.00$, $p = 0.94$; modal sequence difference $r^2 = 0.01$, $p = 0.71$; maximum sequence difference $r^2 = 0.03$, $p = 0.50$).

The general lack of significant association between genetic diversity measures and biogeographic variables may reflect inaccuracy in the biogeographic model tested. If so, then the distance from the actual chimpanzee refuge might still be a valid predictor of genetic diversity and endemism. Therefore, I asked whether any single location within the bounds of the study area has properties that conform to the expectations of a Pleistocene refuge. Such an area would presumably be maximally genetically diverse and would contain a large proportion of endemic alleles. Both diversity and endemism would decline with distance from this area in all directions.

To locate such an area in space, I constructed a rectangular grid of geographic coordinates between longitudes 26.0 and 32.0° and between latitudes −5.0 and 3.0°. Points are regularly spaced every 0.2° (longitudinal and latitudinal) within these boundaries. I measured great-circle distances between each of the resulting 1200 locations and the 19 forest locations from which I have chimpanzee DNA sequences. Then I calculated the Pearson's correlation coefficient for the association between genetic diversity and the geographic distance of each forest location from that point. The magnitude of the resulting correlation coefficient indicates the strength of the association; its sign indicates the direction of the association. By the predefined

criteria of this analysis, refugia should correspond to locations from which the correlation is maximally negative.

I present the results in Figs. 2 and 3 as a series of three-dimensional surfaces, with z axes representing the value of the Pearson correlation coefficient calculated from each point defined by an x (longitudinal) and a y
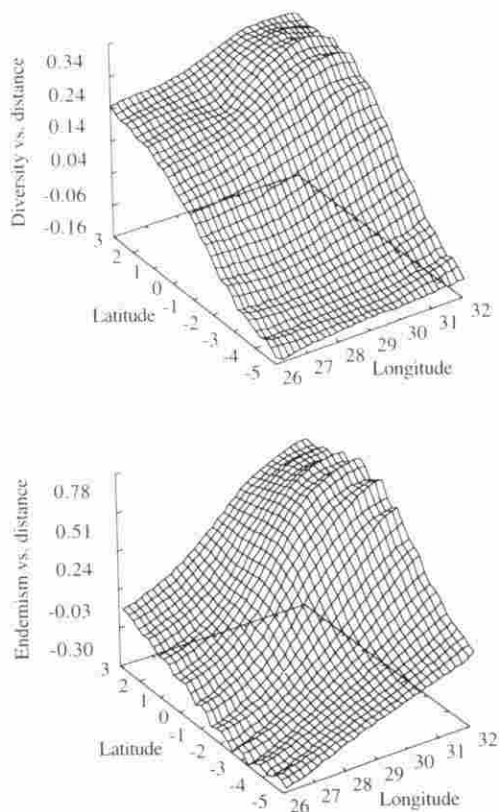
Fig. 2. Topographic representation of the strength of association between geographic distance and genetic diversity (Shannon–Wiener index; top) and between geographic distance and genetic endemism (percentage endemic haplotypes; bottom) as a function of location. For each point on a 0.2° latitudinal by 0.2° longitudinal grid within the study area, a Pearson's correlation coefficient was calculated for the relationship between genetic diversity/endemism and geographic distance from that point for 19 forest locations. The z axis displays the resulting correlation coefficients, unsquared to preserve negativity.

a

b

c

d
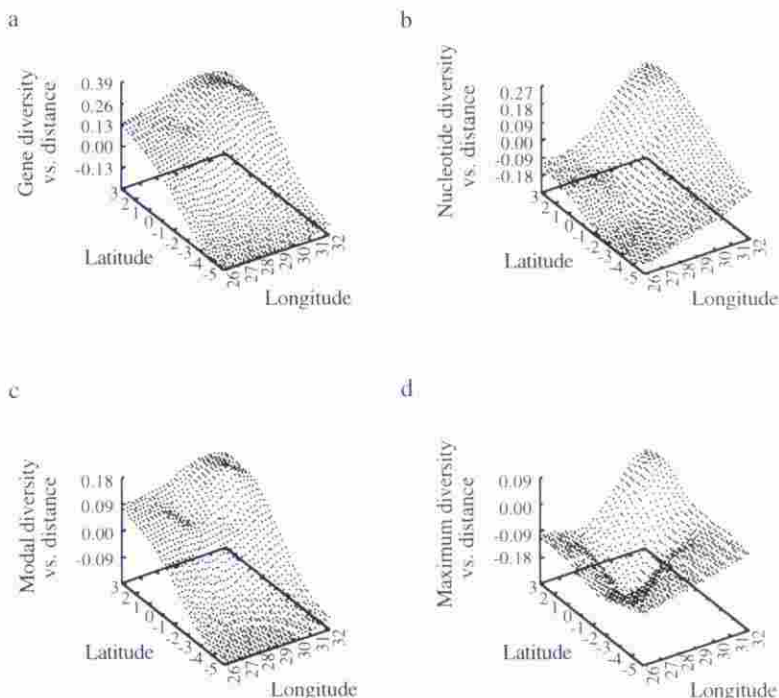
Fig. 3. Topographic representations of the strength of association between geographic distance and four measures of genetic diversity as a function of location. Diversity measures are (a) genetic diversity, (b) nucleotide diversity (mean pairwise sequence difference between haplotypes), (c) modal pairwise sequence difference between haplotypes, and (d) maximum pairwise sequence difference between haplotypes. The z axes are Pearson's correlation coefficients, unsquared to preserve negativity.

(latitudinal) coordinate. Figure 2 shows a clear southwest–northeast cline in the strength of association between Shannon–Wiener diversity and distance, with a maximally negative r (r_{-max}) of -0.26 at longitude $29.0°$ and latitude $-5.0°$. The strength of association between endemism and distance follows a very similar cline, with $r_{-max} = -0.61$ at longitude $26.0°$ and latitude $-2.8°$. Figure 3 presents results of the same analysis for the association between distance and gene diversity ($r_{-max} = -0.26$; longitude $27.8°$, latitude $-5.0°$), nucleotide diversity ($r_{-max} = -0.27$; longitude $27.4°$, latitude $-1.6°$), modal sequence difference ($r_{-max} = -0.17$; longitude $28.0°$, latitude $-4.8°$), and maximum sequence difference ($r_{-max} = -0.27$; longitude $28.8°$, latitude $-0.2°$).

The latitudinal and longitudinal concordance of r_{-max} for all measures is intriguing, as is its general decrease in strength and negativity with dis-

tance from points in the southwest of the study area. The geometric center of mass of r_{-max} for all five diversity measures and endemism lies at longitude 27.8°, latitude –3.2°, which is located in the D.R. Congo, approximately 100 km due west of the northern shore of Lake Tanganiyka. These results are, however, entirely dependent on the geometric arrangement of sampling locations. For example, a southwest–northeast cline of correlational strength would never be observed for forests situated on a perpendicular northwest–southeast archipelago. The precise location of r_{-max} would be unbiased only if sampling locations were evenly (or randomly) distributed across the landscape.

More importantly, no value of r_{-max} is notably high. Distance from any location explains a maximum of only 37% of the variation in any diversity measure ($r_{-max}2 = 0.37$ for the relationship between distance and endemism). Distance from any single source population is therefore a poor predictor of genetic diversity, regardless of where that population is inferred to be.

DISCUSSION

In general, the model of postglacial dispersal from forest refugia does not have much explanatory power for eastern chimpanzee genetic diversity. Forest area is a poor predictor of genetic diversity, accounting for a maximum of only 30% of the variation in genetic diversity among populations by any measure. Geographic distance from the source is an equally poor predictor. Within the study area, distance from any geographic point explains only a maximum of 37% of the variation in genetic diversity among populations in the expected direction.

If eastern chimpanzee populations retain a genetic signature of dispersal from any single refugium within the study area, the evidence has been largely obscured. This could, in part, be a statistical phenomenon. Sample sizes of chimpanzees within populations were both small relative to actual population numbers and heterogeneous. Point estimates of population genetic diversities were therefore of limited and variable accuracy, reducing the power of the statistical tests used. More intensive sampling either of individuals or of independently assorting genetic loci could alleviate this problem.

Alternatively, the biogeographic model tested—dispersal from a single forest refugium—may be an oversimplification. Chimpanzee population history may have been characterized by multiple refugia and complex dispersal routes (Kingdon, 1981; Colyn, 1991), which would significantly confound the prediction of population genetic diversity. In addition, even if the initial

diversity of chimpanzee populations were perfectly predicted by area and geographic distance, subsequent gene flow could have overwritten this pattern. Phylogeographic analyses of these data indicate that migration has indeed been extensive in the subspecies, with an average of between three and four migrants moving between populations per generation (Goldberg and Ruvolo, 1997b).

Despite general lack of support for the genetic trends predicted, two lines of evidence are nevertheless suggestive in light of the refuge model of chimpanzee dispersal. First, haplotype endemism is higher in western forests than in eastern forests, as would be expected if the eastern D.R. Congo was indeed the location of an important Pleistocene refuge. However, the western populations sampled tend to be separated by large rivers, which are barriers to dispersal for forest primates (Colyn, 1991). Rivers may have impeded gene flow between these populations, thereby elevating estimates of endemism.

The second line of evidence supporting the refuge hypothesis is the observation that the two populations with the consistently highest genetic diversities by all measures are RIKA and IIAA (Table II). RIKA, in Uganda's Rwenzori Mountains, is an acknowledged montane forest refuge, unusually high in specific diversity and endemism (Struhsaker, 1981; Rodgers et al., 1982; Colyn, 1991). Similarly, IIAA in the D.R. Congo's Ituri Forest is located in the heart of a region of exceptional biotic diversity, often reconstructed as a refuge (Hamilton, 1976; Grubb, 1982; Colyn, 1991). These two areas are unique in being the only such locations that I sampled.

The evidence that *Pan troglodytes schweinfurthi* was once restricted to a large western forest refugium from which it subsequently dispersed is therefore equivocal. However, the combined results of these analyses suggest an alternative biogeographic model that both accounts for the data and is consistent with behavioral observations of modern chimpanzees. Chimpanzees are extraordinarily vagile, showing impressive ranging abilities and ecological flexibility (Kortlandt, 1983; Goodall, 1986). Accordingly, it is possible that chimpanzees were able to live outside of refugia, as well as in them, during periods of minimal forest cover. Chimpanzees may have travelled among dispersed patches of suitable habitat in an ecological mosaic of savannah, woodland, and gallery forest during times when other, more obligately forest taxa were restricted to refugia. If so, then extensive and uninterrupted gene flow could have occurred throughout the history of the subspecies.

There is genetic evidence that the eastern chimpanzee population underwent at least one population expansion during the recent Pleistocene (Goldberg and Ruvolo, 1997b). The vicissitudes of the tropical climate in

Africa make it likely that this expansion was not an isolated event but was probably the most recent in a series of cyclic population reductions and expansions, each corresponding to a climatically driven deforestation/reforestation cycle. My findings suggest that such events may have been contractions and expansions in chimpanzee population numbers only, without any appreciable reduction in geographic range. During cold, arid periods, chimpanzee population numbers may have decreased as the carrying capacity of the environment decreased. Refuge populations such as RIKA and IIAA may have nevertheless remained large, accounting for their high genetic diversities today. When the climate became more favorable, eastern chimpanzee populations across the subspecific range may have reexpanded *in situ*, maintaining continual gene flow. Populations would not therefore have expanded out of refugia, as the traditional forest refuge model predicts. Refuge populations would simply have been populations that were relatively buffered from climatic change and which therefore maintained large long-term effective sizes.

This model is consistent with the interpretations of Kortlandt (1983), who emphasizes the historical importance of marginal environments to chimpanzees. In fact, rather than being marginal, nonforest environments may have typified much of chimpanzee evolutionary history. If so, we should view the chimpanzee not as a forest species occasionally forced to occupy suboptimal habitats but, instead, as a historical ecological generalist. Genetic data from other great ape subspecies, and from sympatric forest mammals, may someday indicate whether chimpanzees were unique in their ability to range throughout the radically changing environments of the Pleistocene. Such data would be most useful, however, if coupled with direct observations of chimpanzees currently living in open habitats between the large forest patches in which we have most intensely studied them.

REFERENCES

Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., and Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465.

Baker, A. J., Denison, M. D., Lynch, A., and Grand, G. L. (1990). Genetic divergence in peripherally isolated populations of chaffinches in the Atlantic islands. *Evolution* 44: 981-999.

Chakraborty, R., and Nei, M. (1977). Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution* 31: 347-356.

Coetzee, J. A. (1964). Evidence for a considerable depression of the vegetation belts during the upper Pleistocene on the East African mountains. *Nature* 204: 564-566.

Colyn, M. (1991). L'Importance zoogéographique du bassin du fleuve Zaïre pour la spéciation: Le cas des primates simiens. *Ann. Mus. Roy. l'Afr. Centrale Tervuren Belg.* 264: 1-249.

Colyn, M. M., Gautier-Hion, A., and Verheyen, W. (1991). A re-appraisal of paleoenvironmental history in Central Africa: Evidence for a major fluvial refuge in the Zaire Basin. *J. Biogeog.* 18: 403-407.

Edwards, S. V. (1993). Mitochondrial gene genealogy and gene flow among island and mainland populations of a sedentary songbird, the grey-crowned babbler (*Pomatostomus temporalis*). *Evolution* 47: 1118-1137.

Gilbert, D. A., Lehman, N., O'Brien, S. J., and Wayne, R. K. (1990). Genetic fingerprinting reflects population differentiation in the California Channel Island fox. *Nature* 344: 764-766.

Goldberg, T. (1996). *Genetics and Biogeography of East African Chimpanzees, Pan troglodytes schweinfurthi*, Ph.D. thesis, Harvard University, Cambridge, MA.

Goldberg, T., and Ruvolo, M. (1997a). Molecular phylogenetics and historical biogeography of East African chimpanzees. *Biol. J. Linn. Soc.* 61: 301-324.

Goldberg, T., and Ruvolo, M. (1997b). The geographic apportionment of mitochondrial genetic diversity in East African chimpanzees, *Pan troglodytes schweinfurthi. Mol. Biol. Evol.* 14: 976-984.

Goldberg, T., and Wrangham, R. (1997). Genetic correlates of social behavior in wild chimpanzees: Evidence from mitochondrial DNA. *Anim. Behav.* 54: 559-570.

Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*, Harvard University Press, Cambridge, MA.

Grubb, P. (1982). Refuges and dispersal in the speciation of African forest mammals. In Prance, G. T. (ed.), *Biological Diversification in the Tropics*, Columbia University Press, New York.

Hamilton, A. (1976). The significance of patterns of distribution shown by forest plants and animals in tropical Africa for the reconstruction of upper Pleistocene paleoenvironments: A review. *Palaeoecol. Afr.* 9: 63-97.

Hamilton, A. C. (1988). Guenon evolution and forest history. In Gautier-Hion, A., Bourliere, F., Gautier, J. P., and Kingdon, J. (eds.), *A Primate Radiation: Evolutionary Biology of the African Guenons*, Cambridge University Press, Cambridge.

Howard, P. C. (1991). *Nature Conservation in Uganda's Tropical Forest Reserves*, IUCN, Gland, Switzerland, and Cambridge, UK.

Inoue, K., and Kawahara, T. (1990). Allozyme differentiation and genetic structure in island and mainland Japanese populations of *Campanula punctata* (Campanulaceae). *Am. J. Bot.* 77: 1440-1448.

Kendall, R. L. (1969). An ecological history of the Lake Victoria basin. *Ecol. Monogr.* 39: 121-176.

Kingdon, J. (1981). Where have the colonists come from? A zoogeographical examination of some mammalian isolates in eastern Africa. *Afr. J. Ecol.* 19: 115-124.

Kocher, T. D., and Wilson, A. C. (1991). Sequence evolution of mitochondrial DNA in humans and chimpanzees. In Osawa, S., and Honjo, T. (eds.), *Evolution of Life: Fossils, Molecules and Culture*, Springer, Tokyo.

Kortlandt, A. (1983). Marginal habitats of chimpanzees. *J. Hum. Evol.* 12: 231-278.

Lewontin, R. C. (1972). The apportionment of human diversity. *Evol. Biol.* 6: 381-398.

Major, J. (1988). Endemism: A botanical perspective. In Myers, A. A., and Giller, P. S. (eds.), *Analytical Biogeography*, Chapman & Hall, London.

Mayr, E. (1954). Changes of genetic environment and evolution. In Huxley, J., Hardy, A. C., and Ford, E. B. (ed.), *Evolution as a Process*, Allen and Unwin, London.

McGrew, W. C., Baldwin, P. J., and Tutin, C. E. G. (1981). Chimpanzees in a hot, dry and open habitat: Mt. Assirik, Senegal, West Africa. *J. Hum. Evol.* 10: 227-244.

Moeyersons, J., and Roche, E. (1982). Past and present environment. In Van Noten, F. (ed.), *The Archaeology of Central Africa*, Akademische Druck u. Verlagsanstalt, Graz/Austria.

Moore, J. (1992). "Savanna" chimpanzees. In Nishida, T., McGrew, W. C., Marler, P., Pickford, M., and deWaal, F. B. M. (eds.), *Topics in Primatology: Human Origins*, University of Tokyo Press, Tokyo.

Morin, P. A., Moore, J. J., Chakraborty, R., Jin, L., Goodall, J. & Woodruff, D. S. (1994). Kin selection, social structure, gene flow and the evolution of chimpanzees. *Science*, 265: 1193-1201.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70: 3321-3323.

Nei, M. (1987). *Molecular Evolutionary Genetics*, Columbia University Press, New York.

Nei, M., and Graur, D. (1984). Extent of protein polymorphism and the neutral mutation theory. *Evol. Biol.* 17: 73-118.

Nei, M., Maruyama, T., and Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* 29: 1-10.

Osmaston, H. A. (1967). The sequence of glaciations in the Ruwenzori and their correlation with glaciations of other mountains in East Africa and Ethiopia. *Palaeoecol. Afr.* 2: 26-28.

Rodgers, W. A., Owen, C. F., and Homewood, K. M. (1982). Biogeography of East African forest mammals. *J. Biogeog.* 9: 41-54.

Sage, R. D., and Wolff, J. O. (1986). Pleistocene glaciations, fluctuating ranges, and low genetic variability in a large mammal (*Ovis dalli*). *Evolution* 40: 1092-1095.

Seutin, G., Brawn, J., Ricklefs, R. E., and Bermingham, E. (1993). Genetic divergence among populations of a tropical passerine, the streaked saltator (*Saltator albicollis*). *Auk* 110: 117-126.

Seutin, G., Klein, N. K., Ricklefs, R. E., and Bermingham, E. (1994). Historical biogeography of the bananaquit (*Coereba flaveola*) in the Caribbean region: A mitochondrial DNA assessment. *Evolution* 48: 1041-1061.

Skinner, B. J., and Parker, S. C. (1987). *Physical Geology*, John Wiley and Sons, New York.

Struhsaker, T. T. (1981). Forest and primate conservation in East Africa. *Afr. J. Ecol.* 19: 99-114.

Taylor, D. M. (1993). Environmental change in montane southwest Uganda: A pollen record for the Holocene from Ahakagyezi Swamp. *The Holocene* 3: 324-332.

Teleki, G. (1989). Population status of wild chimpanzees (*Pan troglodytes*) and threats to survival. In Heltne, P. G., and Marquardt, L. A. (eds.), *Understanding Chimpanzees*, Harvard University Press, Cambridge, MA.

van Zinderen Bakker, E. M., and Coetzee, J. A. (1972). A re-appraisal of late-Quaternary climatic evidence from tropical Africa. *Palaeoecol. Afr.* 7: 151-181.

Watterson, G. A. (1984). Allele frequencies after a bottleneck. *Theor. Pop. Biol.* 26: 387-407.

Weber, A. W. (1987). Socioecologic factors in the conservation of Afromontane forest reserves. In Mittermeier, R. (ed.), *Primate Conservation in the Tropical Rain Forest*, Alan R. Liss, New York.

Westerbergh, A., and Saura, A. (1994). Genetic differentiation in endemic Silene (Caryophyllaceae) on the Hawaiian Islands. *Am. J. Bot.* 81: 1487-1493.

Wrangham, R. W., Conklin, N. L., Etot, G., Obua, J., Hunt, K. D., Hauser, M. D., and Clark, A. P. (1993). The value of figs to chimpanzees. *Int. J. Primatol.* 14: 243-256.

Wright, S. (1931). Evolution in Mendelian populations. *Genetics* 16: 97-159.