Biological Journal of the Linnean Society (1997), *61:* 301–324, with 9 figures

Molecular phylogenetics and historical biogeography of east African chimpanzees

TONY L. GOLDBERG AND MARYELLEN RUVOLO

Department of Anthropology, Harvard University, Peabody Museum, 11 Divinity Avenue, Cambridge, MA 02138, U.S.A.

Received 6 August 1996; accepted for publication 5 December 1996

Two hundred and sixty eight DNA sequences (hypervariable region 1 of the mitochondrial control region) were obtained from chimpanzees (*Pan troglodytes*) in 19 natural populations within the range of the easternmost subspecies, *P. t. schweinfurthii*. Methods of phylogenetic reconstruction were applied at both the haplotype and population levels. Chimpanzee haplotypes do not sort into location-specific clades on any haplotype trees, indicating that the subspecies is free of major phylogeographic subdivisioning. Trees of populations in which geographic structure was imposed on the data lacked phylogenetic resolution in that interpopulational relationships were poorly supported statistically. These results indicate either a near simultaneous origin for the chimpanzee populations sampled, or an obscuring of interpopulational phylogenetic relationships by gene flow. In contrast, area cladograms of the forests from which chimpanzees were sampled (constructed using lists of endemic taxa) were robust and statistically well-supported. Chimpanzee population history is apparently decoupled from the history of the forests which the populations inhabit. Eastern chimpanzee data are also used to draw phylogenetic and molecular evolutionary comparisons to humans.

ADDITIONAL KEY WORDS:—*Pan troglodytes schweinfurthii* – phylogeography – molecular evolution – mitochondrial DNA.

CONTENTS

Correspondence to T.L. Goldberg, current address: Department of Pathobiology, College of Veterinary Medicine, University of Illinois, 2001 South Lincoln Avenue, Urbana, IL 61801, U.S.A. email: tlgoldbe@students.uiuc.edu

0024–4066/97/070301+24 $25.00/0/bj960132

INTRODUCTION

The argument that abiotic change drives evolutionary divergence among taxa is the benchmark of 'phylogenetic biogeography' (Brundin, 1988; Croizat, 1958, 1964). 'Cladistic biogeography' (Humphries *et al.*, 1988), the melding of phylogenetic biogeography with the methodology of Hennig (1966), endeavours specifically to reconstruct historical relationships among geographic areas from phylogenies of indigenous taxa, built using character-based methods (Nelson, 1981; Platnick & Nelson, 1984; Rosen, 1988; Morrone & Carpenter, 1994). The recent availability of molecular data (especially from the mitochondrial genome) has added power to such methods by increasing the resolution of reconstructed phylogenies (Avise *et al.*, 1987; Avise, 1989, 1994).

This paper applies a phylogenetic biogeographic perspective to the analysis of a large data set of mitochondrial DNA sequences in order to reconstruct historical relationships among populations of chimpanzees in eastern Africa. The current geographic distribution of chimpanzees corresponds closely to the distribution of forested habitat in Africa (Kortlandt, 1972, 1983; Goodall, 1986; Teleki, 1989). This observation implies that chimpanzee distribution is somewhat restricted by the geographic limits of forest habitat, although chimpanzees can and do exist outside of forests (McGrew, Baldwin & Tutin, 1981; Kortlandt, 1983; Moore, 1992). To the extent that chimpanzees have historically been restricted to forested habitats, evolutionary relationships among extant chimpanzee populations should parallel historical relationships among the forests which they inhabit.

Forest history is, in turn, largely driven by climate. Dramatic shifts in global temperature have occurred throughout the Pleistocene (Emliani, 1955; Taylor *et al.*, 1993), and have caused corresponding expansions and contractions of forest in eastern Africa (Kendall, 1969; van Zinderen Bakker & Coetzee, 1972; Moeyersons & Roche, 1982). The centers of origin for many extant African mammals are thought to be Pleistocene forest 'refugia', where forests persisted during cold, dry periods in the earth's climate (Endler, 1982; Haffer, 1982; Mayr & O'Hara, 1986). Forest taxa presumably dispersed to reach their present distributions during the re-expansion of forests following post-glacial climatic amelioration (Struhsaker, 1981; Grubb, 1982; Hamilton, 1988; Colyn, 1991).

The hypothesis that forest taxa in eastern Africa have retracted into and expanded out of Pleistocene refugia implies that present-day forests in this region are related by a unique, if complicated, history. Under the refuge model, forest biota would have differentiated due to: (1) the dispersal of forest taxa into previously unforested areas during post-glacial forest expansion, and (2) the fragmentation and reduction (vicariance) of forests and their indigenous biota during the onset of glacial conditions. Applied to chimpanzee population history, refuge theory predicts that populations of eastern chimpanzees inhabiting different forests should be related by a unique, discernible history. It also predicts that this history should be concordant with the historical relationships among the forests themselves. The goal of this paper is to test these two hypotheses using genetic data from wild chimpanzee populations.

MATERIAL AND METHODS

Genetic material was sampled non-invasively (in the form of shed hairs collected from sleeping nests) from chimpanzees in 19 geographically defined populations

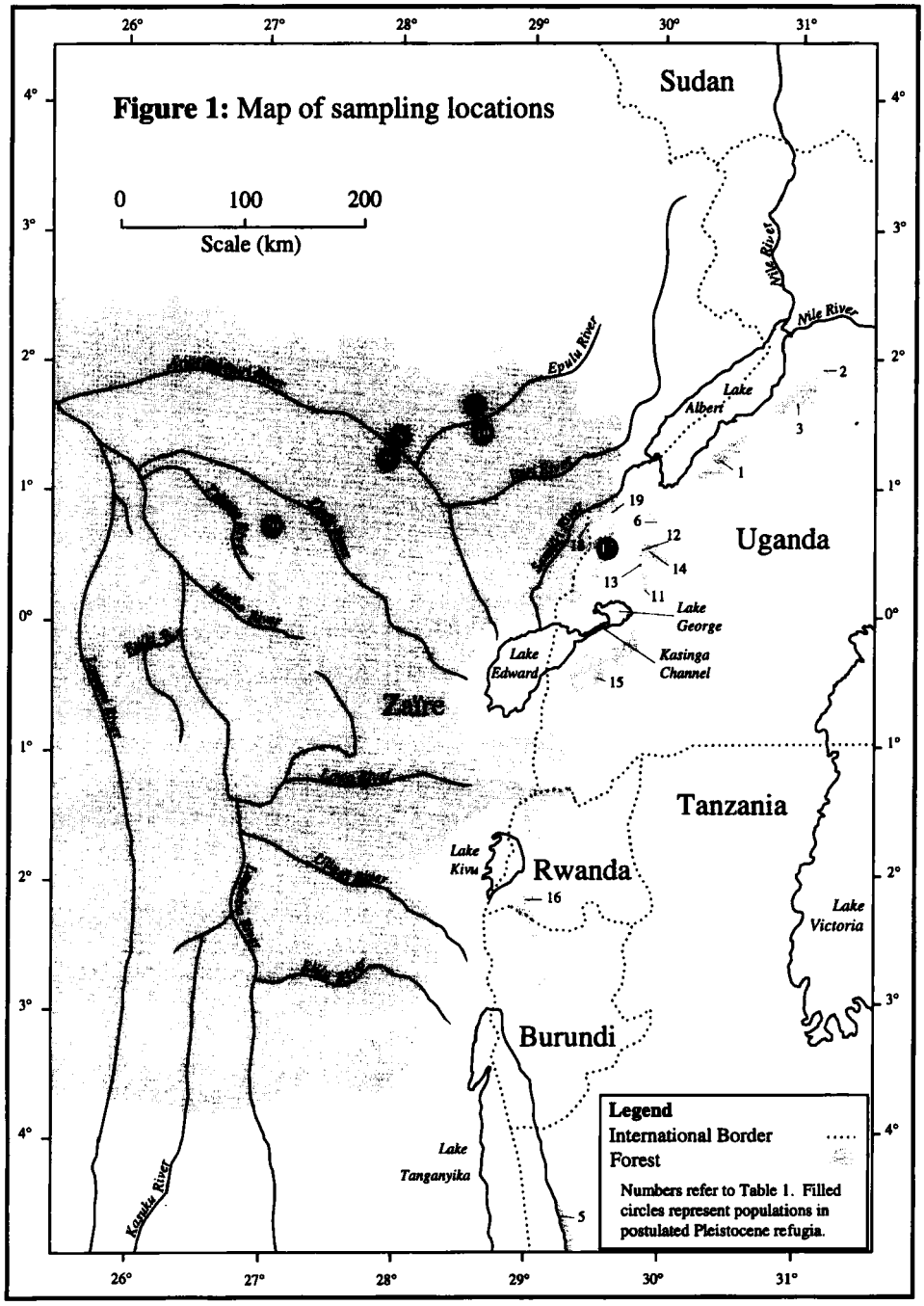

Figure 1. Map of sampling locations.

within the range of *P. t. schweinfurthii* (Fig. 1). A captive population of eastern chimpanzees was also sampled (population #4, not shown in Fig. 1), but is not included in the analyses below because of the uncertain geographic origins of these

TABLE 1. Description of sampling locations

Abbrev.	*Forest*	*Location**	*Longitude***	*Latitude***	*n*†	
1	BAMA	Bugoma	Mwela Sawmill	30°59′E	1°18′N	13
2	BOPI	Budongo	Pabidi	31°44′E	1°55′N	12
3	BOSO	Budongo	Sonso Sawmill	31°33′E	1°44′N	15
4	EEZO		Entebbe Zoo			26
5	GEKA	Gombe	Kasekela	29°36′E	4°41′S	19
6	IARA	Itwara	Rwebikuya	30°25′E	0°47′N	12
7	IIAA	Ituri	Afarama	28°32′E	1°33′N	13
8	IIAE	Ituri	Avakubi NE	27°34′E	1°20′N	12
9	IIAW	Ituri	Avakubi SW	27°33′E	1°19′N	13
10	IILA	Ituri	Lenda	28°38′E	1°19′N	17
11	KEDN	Kibale	Dura Station	30°19′E	0°12′N	14
12	KEKA	Kibale	Kanyawara	30°21′E	0°34′N	15
13	KEKU	Kibale	Kanyanchu	30°20′E	0°28′N	14
14	KENO	Kibale	Ngogo	30°25′E	0°30′N	13
15	KUSL	Kalinzu	Kalinzu Sawmill	30°03′E	0°28′S	13
16	NESN	Nyungwe	Nyungwe Station	29°30′E	2°25′S	13
17	RIKA	Rwenzori	Katebwa	30°06′E	0°32′N	13
18	SIMU	Semliki	Mbume-Busaru	30°02′E	0°43′N	13
19	SINI	Semliki	Ntandi	30°09′E	0°48′N	11
20	TOBA	Tshopo	Bafwabalinga	27°04′E	0°51′N	10

* See Figure 1; sampling locations are shown by number on this map.
** Latitudes and longitudes are measured to the nearest minute.
† Sample sizes of DNA sequences obtained.

animals (Goldberg, 1996). Laboratory methods for the extraction, amplification and sequencing of DNA are described elsewhere (Goldberg, 1996).

A 368 bp segment of the mitochondrial d-loop region (corresponding to Anderson reference sequence coordinates 16042–16410), which includes the first hypervariable region, was chosen for study. Hypervariable region 1 was selected for its especially high rate of evolution (Kocher & Wilson, 1991), because of the anticipated close evolutionary relationships of the taxa being studied. Also, hypervariable region 1 has been well-characterized in humans and chimpanzees. and has been used to discern patterns of chimpanzee phylogeography across Africa (Morin *et al.*, 1994).

RESULTS

Two hundred and sixty two DNA sequences were generated. Table 1 presents the names and geographic coordinates of each sampling location, as well as the sample sizes of DNA sequences obtained. Abbreviated location names listed in Table 1 will be used throughout the study. Nineteen geographically-localized DNA sequences from Morin *et al.* (1994) were also incorporated into the study. These sequences correspond to population #5 (GEKA).

The geographic arrangement of locations sampled in the present study, and their proximity to reconstructed Pleistocene refugia, predict a specific pattern of historical relationship for these forested areas. This pattern is presented in Figure 2 as a

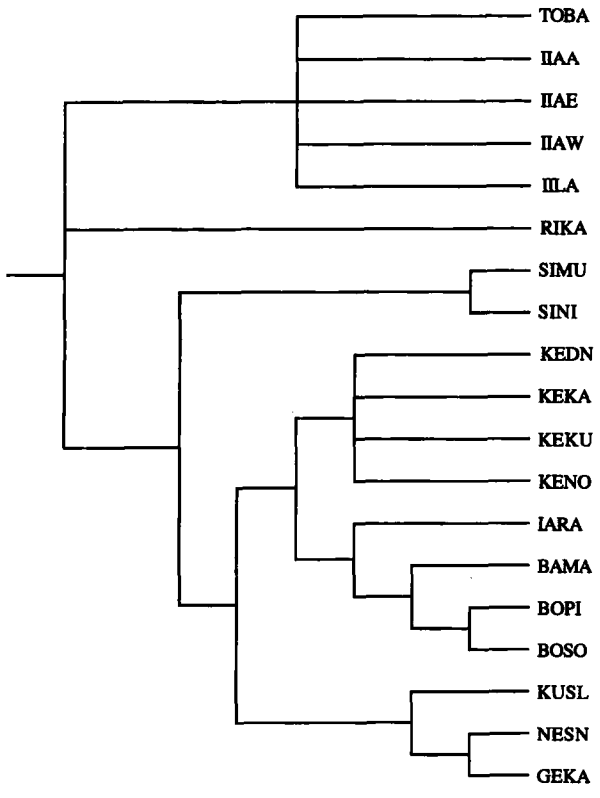

Figure 2. Hypothetical phylogeny of sampling locations based on reconstructed locations of Pleistocene forest refugia and the current geographic arrangement of forests (see text for full explanation).

phylogenetic tree (constructed manually). Refuge populations (RIKA and the Zaïrian populations) appear as outgroups relative to the other locations under the assumption that refugia are 'source' areas from which the indigenous biota of other peripheral (non-refuge) forests have dispersed (Hamilton, 1976, 1981; Kingdon, 1981; Howard, 1991). Peripheral forests diverge in the order of their geographic proximity to refugia, under the assumption that increasingly-peripheral locations were colonized by forest taxa sequentially during post-glacial forest expansion. This same topology would also be predicted under an alternate vicariant model in which the most peripheral forests were the first to become isolated from a shrinking 'mainland' of forest during the onset of glacial conditions. Figure 2 also reflects the hypothesis that, within eastern forests, southern forests (KUSL, NESN, GEKA) should cluster distinctly from northern forests due to the possible role of lakes Edward and George and the Kasinga Channel as a geographic barrier to dispersal (see Figure 1).

Haplotype trees

One hundred and twenty three distinct chimpanzee haplotypes were identified in the present study. The raw DNA sequences are available through GenBank (accession numbers U77181–U77293) and in Morin *et al.* (1994). Of the 368 bp

sequenced, 90 positions were variable. The number of unique, bifurcating trees that could be constructed for a data set of this size is astronomical (Felsenstein, 1978), and precludes the use of exact search algorithms which employ maximum-parsimony optimality criteria (Farris, 1986; Felsenstein, 1993; Swofford, 1993). Initial computational attempts using simultaneous runs of PAUP (Swofford, 1993) on separate microcomputers suggest that a statistically meaningful searching of 'tree space', even using heuristic approximation, would take on the order of months to years. The problem is analogous to that described for the human mitochondrial gene tree (Vigilant *et al.*, 1991), from which topological inferences about modern human origins have been problematic due to large numbers of equally-parsimonious cladograms (Maddison, Ruvolo & Swofford, 1992). Furthermore, the number of taxa in the present sample (123) is greater than the number of variable characters (90), rendering parsimony-based cladograms intrinsically weak (Swofford, 1993; Rogers & Jorde, 1995). For these reasons, the haplotype trees described below are distance trees. The analysis which follows acknowledges that alternate, more parsimonious trees probably do exist, but assumes that distance methods provide sufficient phylogenetic resolution to address the hypotheses outlined above. This assumption is tested below.

Figure 3 presents a phylogeny of eastern chimpanzee haplotypes reconstructed from a matrix of distances generated using the DNAdist module of the computer program PHYLIP with a maximum-likelihood distance correction and a transition: transversion ratio of 10:1. Transition-transversion ratios as low as 1:1 and as high as 30:1 were also used, but did not alter the results of this or the following analyses. The algorithm used to generate the tree is the neighbour-joining algorithm (Saitou & Nei, 1987). The tree is rooted using outgroup sequences from three central African chimpanzees (*P. t. troglodytes*), three west African chimpanzees (*P. t. verus*), two bonobos (*P. paniscus*) and two humans (*Homo sapiens*). One west African chimpanzee sequence and both bonobo sequences were taken from the literature (Horai *et al.*, 1992; Morin *et al.*, 1994), as were the two human sequences (Vigilant *et al.*, 1991), which were specifically chosen to be maximally-divergent within Vigilant's 135 haplotype world human sample. Haplotypes are numbered sequentially from the top to the bottom of the tree and are described by number in Table 2.

Diversity within eastern chimpanzees (measured as the length of the branches connecting the most divergent eastern chimpanzee sequences) is lower than that within humans. The two human sequences differ from one another by a total of 22 substitutions (20 transitions and 2 transversions), while the most divergent clade of eastern chimpanzee haplotypes (sequences 11–14) differs from the other eastern chimpanzee haplotypes by a maximum of 20 substitutions (18 transitions and 2 transversions). These estimates did not change significantly when the data were corrected for multiple substitutions using a maximum likelihood distance correction and a transition/transversion ratio of 10:1. The approximate coalescent time for eastern chimpanzees would therefore be (using corrected estimates) 20/22, or 91% of the human estimate. A human estimate of 298 kya (Ruvolo *et al.*, 1993) would therefore imply a coalescence time of 272 kya for the last common ancestor of all eastern chimpanzees.

Diversity within eastern chimpanzees is also lower than that within western African chimpanzees (28 transitions, 2 transversions) and bonobos (29 transitions, 2 transversions), but is higher than that within the central African chimpanzees sampled (14 transitions, 1 transversion). Diversity within bonobos and west African

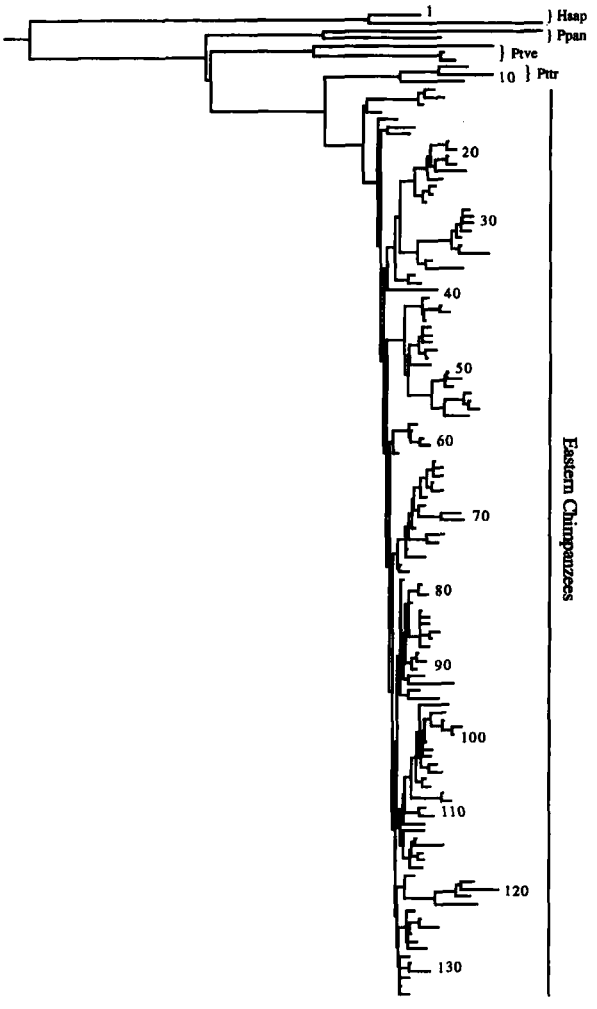

Figure 3. Neighbour-joining tree of 123 eastern chimpanzee haplotypes, rooted using Human (Hsap), Bonobo (Ppan), *P. t. verus* (Ptve) and *P. t. troglodytes* (Pttr) as outgroups. Branch lengths are proportional to genetic distances between haplotypes. Taxa are numbered sequentially from top to bottom, and are described by number in Table 2.

chimpanzees exceeds that within humans, despite minimal geographic sampling within these subspecies. Indeed, maximum sequence difference within *P. t. verus* and *P. paniscus* is greater even than the genetic distance between *P. t. schweinfurthii* and *P. t. troglodytes* (25 transitions, 2 transversions). Again, correction for multiple substitutions did not appreciably change any of these estimates.

On the tree in Figure 3, Zaïrian locations do not sort to the exclusion of eastern forest locations. However, the most divergent clade (from the others) within eastern chimpanzees consists exclusively of sequences from Zaïre (Figure 3, sequences 11–14). This topology is consistent with a Zaïrian location for the basal node linking eastern chimpanzees, as would be predicted if a major Pleistocene forest refuge existed in northeast Zaïre (Struhsaker, 1981; Grubb, 1982; Hamilton, 1988; Colyn,

TABLE 2. Description of taxa on the phylogenetic tree in Figure 3

Taxon[1]	Found in populations[2]	Number of individuals with haplotype[3]	Taxon[1]	Found in populations[2]	Number of individuals with haplotype[3]
1 to 10	Outgroup taxa[4]		67	RIKA	1
11	IIAA	2	68	SINI	2
12	EEZO, IIAA	1,1	69	IARA	1
13	IIAA	3	70	KEKA	1
14	IILA	1	71	IARA	1
15	KEDN	1	72	EEZO, IIAW, KEDN, KEKU, KENO, RIKA	1,1,1,1,2,1
16	NESN	3			
17	GEKA	2	73	BOPI	1
18	IILA	2	74	TOBA	1
19	IILA	1	75	BOPI, KEKU, RIKA	3,1,1
20	KEKU, SINI	1,3	76	EEZO	1
21	SINI	1	77	BOPI	1
22	BOPI	1	78	GEKA	3
23	IILA	1	79	RIKA	1
24	KUSL	1	80	RIKA	2
25	GEKA	2	81	IIAE	2
26	BAMA	1	82	IIAA	1
27	EEZO	1	83	GEKA	1
28	IIAA	1	84	IIAW	2
29	IILA	3	85	IILA	1
30	IILA	3	86	IIAW	1
31	BOPI	1	87	IIAA	1
32	KENO	3	88	EEZO	1
33	IIAA	1	89	KUSL	1
34	EEZO, GEKA, RIKA	2,1,2	90	KUSL	1
35	IIAA	1	91	KUSL	1
36	IARA	1	92	KUSL	1
37	KEDN	5	93	KUSL	1
38	IIAW	2	94	EEZO, IIAW, KEKU, SIMU, SINI	1,3,1,7,3
39	BAMA, BOPI, BOSO	1,1,4			
40	NESN	2	95	IIAE	1
41	GEKA	1	96	GEKA	3
42	BOSO, IARA	1,1	97	KUSL	1
43	GEKA	1	98	IIAE	7
44	BOSO	1	99	NESN	2
45	BOSO	1	100	NESN	2
46	EEZO, KENO	1,1	101	RIKA, TOBA	1,1
47	KEKU, SINI	2,1	102	EEZO	1
48	EEZO, RIKA	1,1	103	IIAA	1
49	NESN	2	104	IARA	1
50	KUSL	1	105	EEZO, IARA, KEKA	2,3,2
51	GEKA	2	106	IIAE	1
52	RIKA, SIMU	1,3	107	NESN	1
53	GEKA	2	108	EEZO	1
54	SIMU	1	109	EEZO	1
55	SIMU	1	110	TOBA	1
56	IILA	1	111	EEZO	1
57	IIAW	2	112	TOBA	2
58	EEZO	1	113	IILA	1
59	BOPI, BOSO, IARA	3,1,1	114	BOPI, BOSO, EEZO	1,4,2
60	KEKA, KENO, RIKA	4,2,1	115	TOBA	1
61	EEZO, KEDN, KEKA, KUSL	1,6,2,2	116	BAMA, KEKA	1,1
62	IIAW	1	117	KEDN, KEKA, KUSL	6,2,2
63	IIAW	1	118	BAMA, BOSO, EEZO, KUSL	8,1,2,2
64	EEZO	1	119	NESN	1
65	BOSO	1	120	SIMU	1
66	BOSO	1	121	BAMA	1

TABLE 2. Description of taxa on the phylogenetic tree in Figure 3—*continued*

Taxon[1]	Found in populations[2]	Number of individuals with haplotype[3]	Taxon[1]	Found in populations[2]	Number of individuals with haplotype[3]
122	IILA	1	128	TOBA	1
123	GEKA	1	129	EEZO	1
124	IIAE	1	130	BAMA, IARA, KEKU, SINI	1,3,2,1
125	IIAA	1	131	EEZO, IILA, KUSL, RIKA	2,1,1
126	TOBA	3	132	SINI	1
127	KEKA, KEKU	3,1	133	EEZO	1

[1] Taxa (haplotypes) are numbered sequentially from top to bottom of the tree in Figure 3.
[2] See Figure 1 and Table 1 for description of populations.
[3] Numbers refer to the number of individuals posessing the haplotype in each population.
[4] 1–2 = human; 3–4 = bonobo; 5–7 = western chimpanzee; 8–10 = central chimpanzee

1991). Individual sampling locations (shown in Figure 1) do not form exclusive clades on this tree.

Two other tree reconstruction algorithms were applied to these data. Both the UPGMA algorithm (Sokal & Michener, 1958) and the Fitch-Margoliash algorithm (Fitch & Margoliash, 1967) produced trees with properties very similar to the neighbour-joining tree, and are not, therefore, shown. The compositions of major clades within eastern chimpanzees varied somewhat between the UPGMA and neighbour-joining trees, although both preserved the pattern of low overall diversity and lack of geographic substructuring. The topology of the Fitch-Margoliash tree was also different in its lack of a maximally-divergent clade of exclusively Zaïrian sequences within eastern chimpanzees, and in its (almost certainly incorrect) clustering of *P. t. verus* with *P. paniscus*.

Haplotype phylogenies were used to investigate molecular evolutionary processes within hypervariable region 1 of the mitochondrial control region (Kocher & Wilson, 1991). Character-state changes (in this case, nucleotide base substitutions) were traced along UPGMA, neighbour-joining and Fitch-Margoliash trees of haplotypes for both the eastern chimpanzee sample ($n = 123$) and Vigilant *et al.*'s (1991) world human sample ($n = 135$) using the computer program MacClade (Maddison & Maddison, 1992). Results are presented in Figure 4 for the neighbour-joining tree only, since results for the UPGMA and Fitch-Margoliash trees were indistinguishable. Results obtained by non-tree-based methods (i.e. simple counts of nucleotide variation along the sequence) also yielded a pattern statistically indistinguishable from the one shown in Figure 4. Inferred numbers of transitional changes (open bars) and transversional changes (closed bars) are plotted at ten-nucleotide intervals along the gene region, which corresponds to bases 16041–16413 of human reference sequence (Anderson *et al.*, 1981).

Patterns of nucleotide substitution are similar between humans and eastern chimpanzees. Numbers of transitional changes are greatest in the middle of the region sequenced, as predicted from interspecific comparisons (Kocher & Wilson, 1991). The transitional bias observed in the eastern chimpanzee data (approximately 19:1) is similar to that observed for the human data (approximately 17:1). In neither taxon were the numbers of transitional or transversional changes normally distributed (Kolmogorov-Smirnov test using standard normal distribution; $P<0.0001$ in all

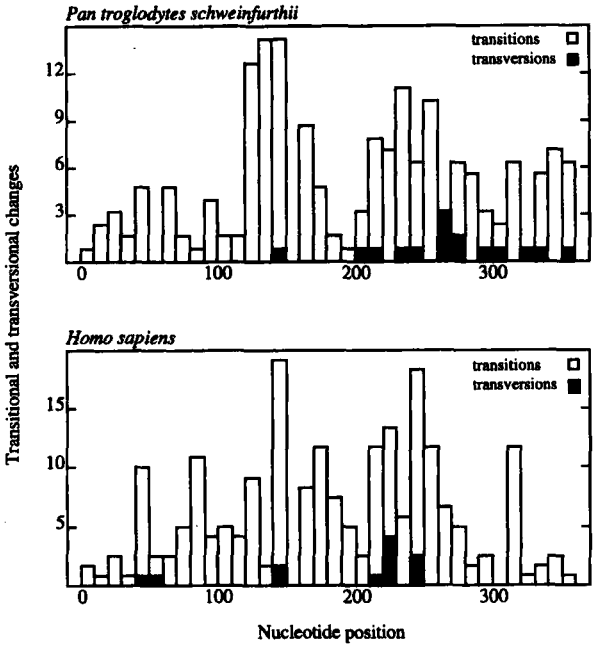

Figure 4. Comparative molecular evolution in eastern chimpanzee ($n=123$) and human ($n=135$) mitochondrial d-loop hypervariable region 1. Minimum numbers of transitional and transversional changes were calculated at 10-nucleotide intervals along the gene region sequenced using MacClade (Maddison & Maddison, 1992). Changes were traced along a phylogenetic tree of haplotypes constructed for each taxon with the program PHYLIP (Felsenstein, 1993) using the neighbour-joining algorithm (Saitou & Nei, 1987). Results were indistinguishable when trees constructed using other methods were analysed. Human data were taken from Vigilant *et al.* (1991).

cases), nor could the distribution of changes be made to fit a normal distribution by standardizing the mean and standard deviation to zero and one, respectively (Lillefors test, $P<0.0001$ in all cases; Lillefors, 1967). Therefore, the statistical comparisons presented below are nonparametric. Tests were run using the program SYSTAT (Wilkinson, Hill & Vang, 1992).

In eastern chimpanzees, areas of high transitional variation did not tend to cluster statistically along the length of the sequence (Wald-Wolfowitz runs test; $P=0.41$). Similarly, transversional changes were not clustered along the gene region in eastern chimpanzees (Wald-Wolfowitz runs test; $P=0.71$). In humans, however, transitional changes were clustered (Wald-Wolfowitz runs test; $P=0.01$), although transversional changes were not (Wald-Wolfowitz runs test; $P=0.15$). Direct comparisons of human and eastern chimpanzee sequences detected no statistical difference, however, either in the case of transitions (Wilcoxon signed rank test; $P=0.40$), or in the case of transversions (Wilcoxon signed rank test; $P=0.97$). No evidence therefore exists that the molecular evolutionary dynamics of this segment of the mitochondrial control region have changed since the divergence of humans and chimpanzees.

The analyses presented above operate under the previously-stated assumption that distance methods of tree reconstruction yield phylogenetically-meaningful topologies. To the extent that parsimony is the most appropriate optimality criterion for reconstructing phylogeny (Farris, 1983; Sober, 1988), this assumption may be

tested empirically. Using MacClade (Maddison & Maddison, 1992), 1000 randomly-generated equiprobable trees were generated for the full data set of 123 chimpanzee haplotypes and 10 outgroups. The length of each tree was determined from the molecular data to generate a null distribution of tree lengths. Lengths were then calculated for each of the distance trees described above. The UPGMA (length = 442), neighbour-joining (length = 408) and Fitch-Margoliash (length = 414) trees were all significantly shorter than the shortest random tree (length = 905), indicating that these distance methods do, in fact, arrive at significantly 'better' topologies than would be expected by chance.

Population trees

When analysed using MacClade (Maddison & Maddison, 1992), trees of haplotypes in which populations were constrained to be monophyletic were considerably less parsimonious than any of the reconstructed distance trees presented above (length of shortest population monophyly constraint tree = 522). This observation does not, however, imply that the populations lack a unique, recoverable history. Even if gene flow among eastern chimpanzee populations has been extensive, it may not have entirely obscured all genetic record of the order in which the populations were founded. The following analyses impose a population structure on the data *a priori* in an effort to reconstruct such a history.

Distance matrices were generated among the 19 geographically-defined populations in the study using frequencies of the 123 identified alleles (see Table 2). Distances were computed and converted into trees using the program PHYLIP (Felsenstein, 1993) with the neighbour-joining algorithm. Other tree-building methods (UPGMA, Fitch-Margoliash) yielded identical topologies. Results are presented in Figure 5. Three allelic distance measures available in PHYLIP were used: Cavalli-Sforza & Edwards (1967; top panel), Nei (1972; middle panel), and Reynolds, Weir & Cockerham (1983; bottom panel). Each measure has properties which render it uniquely informative for reconstructing phylogenetic relationships among populations (Cavalli-Sforza & Edwards, 1967; Nei, 1972; Reynolds *et al.*, 1983; Felsenstein, 1993). However, the likelihood that a correct topology was reconstructed using any measure is effectively zero due to the large number of taxa (19) and small number of genetic systems (1) investigated (Nei, Tajima & Tateno, 1983). The distance approach with gene frequency data should thus be viewed cautiously. It is potentially useful only for identifying particularly robust population groupings.

In the Cavalli-Sforza & Edwards (1967) and Reynolds *et al.* (1983) trees, locations within Budongo Forest (BOPI and BOSO) tend to cluster, as do locations within Semliki Forest (SIMU and SINI). In all three trees, Kibale Forest locations (KEDN, KEKA, KEKU, KENO) are paraphyletic, although KEDN, KEKA and KENO do tend to cluster on the Cavalli-Sforza & Edwards (1967) and Reynolds *et al.* (1983) trees. Few other strong similarities exist among the three trees. Branch lengths in the Nei (1972) tree are all approximately equal (zero), suggesting a 'star phylogeny' with little resolution. A similar pattern is also suggested by the Cavalli-Sforza & Edwards (1967) and Reynolds *et al.* (1983) trees; in both trees, central nodes are linked by short branch lengths while long branches appear only for terminal taxa.

The number of populations in this study (19) is considerably smaller than the number of haplotypes (123), which makes cladistic analysis using parsimony criteria

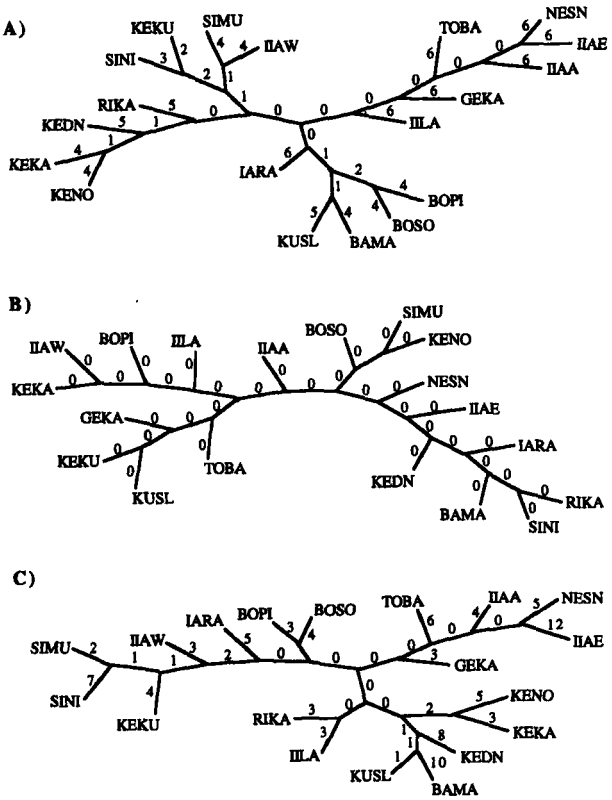

Figure 5. Relationships among populations based on allele frequencies. Three genetic distance matrices were calculated using the program PHYLIP (Felsenstein, 1993). Distance measures were those of (A) Cavalli-Sforza & Edwards (1967), (B) Nei (1972), and (C) Reynolds *et al.* (1983). Trees were constructed using PHYLIP with the neighbour-joining algorithm of Saitou and Nei (1987). Topologies were identical when other tree-building methods were used. Numbers above branches (not drawn to scale) are relative branch lengths.

possible at the population level. The analysis which follows attempts to reconstruct 'ancestral' alleles for each sampling location using cladistic inference; parsimony criteria are then used to reconstruct phylogenetic relationships among these alleles (Rice, Donoghue & Olmstead, in press). The analysis considers only 'endemic' alleles (alleles found exclusively in a single sampling location) under the assumption that they evolved locally from a unique ancestral sequence which was also geographically localized. This analysis further assumes that reconstructed ancestral alleles represent founder alleles from the initial population which colonized each sampling location. 'Eurytopic' alleles (alleles shared among two or more sampling locations) were excluded under the assumption that they represent recent migration and could confound phylogenetic reconstruction by introducing homoplasy.

For each population, a maximum-parsimony tree of endemic haplotypes was generated using the branch and bound algorithm of PAUP (Swofford, 1993) and a transition: transversion ratio of 10:1. This algorithm is an exact search method, practical only for numbers of taxa less than approximately 13. One central African chimpanzee, one west African chimpanzee and one bonobo sequence were included

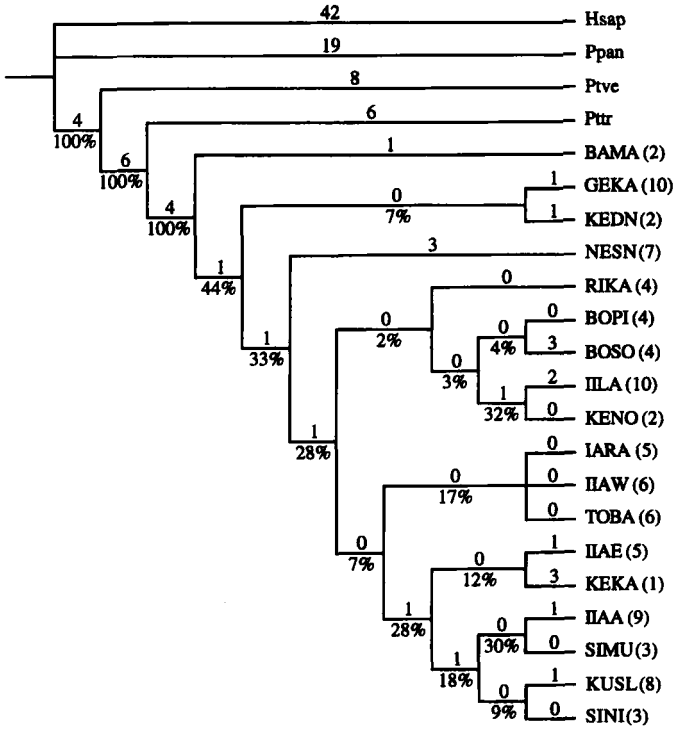

Figure 6. Maximum parsimony tree of ancestral haplotypes. The cladogram is a 50% majority-rule consensus tree computed from 1000 bootstrap replicates of the data using the heuristic search option in PAUP (Swofford, 1993). Numbers above branches are minimum numbers of inferred changes. Numbers below branches are bootstrap values. The tree is rooted using Human (Hsap), Bonobo (Ppan), *P. t. verus* (Ptve) and *P. t. troglodytes* (Pttr) as outgroups. Tree length = 139; rescaled consistency index = 0.588. Ancestral character states were reconstructed from cladistic relationships among endemic haplotypes within sampling locations (calculated individually using the branch and bound algorithm of PAUP). Numbers of endemic haplotypes are shown after taxon names. KEKU was excluded from the analysis because it lacked endemic haplotypes.

in each search as outgroups. In 17 of 19 cases, a single most parsimonious tree was identified. In the remaining 2 cases, more than one (2 and 4, respectively) equally-parsimonious tree was identified. PAUP was then used to reconstruct the ancestral state of each of the 368 characters (nucleotide positions) at the basal node linking the eastern chimpanzee haplotypes in each tree. Ancestral haplotypes were input into PAUP along with sequences from a human, bonobo, a western African chimpanzee, and a central African chimpanzee. PAUP was used with the heuristic search option (10 random-addition replicates) and a transition: transversion ratio of 10:1 to search for maximally parsimonious trees (transition: transversion ratios of 1:1 and 30:1 were also tested, but did not alter the results of the analysis). One thousand bootstrap replicates of the data (Felsenstein, 1985) were run to estimate statistical confidences around reconstructed groupings.

Results are presented in Figure 6. Because 1163 equally-parsimonious trees were found during the heuristic search, the cladogram presented is a 50% majority-rule bootstrap consensus tree in which all possible groupings were allowed. Although outgroup relationships were reconstructed with high confidence, ingroup relationships

were not. Bootstrap values within the eastern chimpanzee clade were universally low and internal branch lengths (minimum numbers of reconstructed changes) were all ≤ 1. The topology is also inconsistent with the predictions of Pleistocene refuge theory, as outlined in the Introduction. The hypothetical population tree presented in Figure 2 was a considerably less-parsimonious explanation of the data (length = 152) than the reconstructed tree (length = 139), when analysed using MacClade (Maddison & Maddison, 1992). 'Refuge' locations (Zaïrian locations and RIKA) are dispersed throughout the tree in Figure 6, rather than being outgroups, as predicted if other populations were founded by chimpanzees dispersing from them (see Fig. 2). BAMA (Bugoma), a peripheral forest which probably did not exist in its present location before approximately 300 years ago (Howard, 1991), is reconstructed as an outgroup to all other populations. Locations within Kibale Forest are paraphyletic, as are locations within Semliki Forest. The two Budongo Forest locations do cluster, but with only 4% bootstrap confidence.

Eurytopic haplotypes were excluded from the analysis described above because they were assumed to be represent recent migration. However, eurytopic haplotypes may be good indicators of relatedness between populations under the alternate assumption that they reflect recent common ancestry. This would be true under a model in which founder populations dispersing from a common ancestral population shared greater numbers of alleles than founding populations dispersing from different ancestral populations. This would also be true in the case of vicariance, under the assumption that recently-vicariated populations retain greater numbers of alleles in common than do distantly-vicariated populations. This reasoning is central to PAE, or 'parsimony analysis of endemicity,' a little-used and little-explored biogeographic technique founded by Rosen and Smith (Rosen, 1984, 1985; Rosen & Smith, 1988).

In its original formulation, PAE combined data on species distributions with the cladistic methodology of Hennig (1966) to build cladograms of geographic areas. This technique considers entire geographic areas as 'taxa' and species as characters. A species inhabiting areas A and B but not C would count as a synapomorphy supporting a clade linking A and B to the exclusion of C. Parsimony criteria can be applied in the usual sense to find the optimal area cladogram. Taxa may also be autapomorphic or symplesiomorphic, in which case they are uninformative for reconstructing area relationships (Rosen, 1988). Ambiguity in character state assignment (can a taxon ever be proven absent from an area?), character weighting, and the ecological interdependence of taxa are but some of the issues which have not been fully addressed for this method. Although its reasoning is sound, the technique should therefore be considered experimental.

PAE was applied to the chimpanzee genetic data at the allelic level. Sampling locations were defined as taxa and were scored for the presence or absence of each of the 24 eurytopic alleles identified in the study (see Table 2). PAUP (Swofford, 1993) was used with the heuristic search option find the most parsimonious tree(s) linking all 19 populations. To estimate statistical confidences associated with observed topological groupings, 1000 heuristic bootstrap replicates were run. Because 634 equally-parsimonious trees were found, a 50% majority-rule bootstrap consensus tree was calculated in which all possible groupings were allowed. The tree was rooted using Lundberg rooting (Lundberg, 1972) and an ancestral population containing no haplotypes. This rooting method is advocated by Rosen & Smith (1988) under the reasoning that it represents "a locality occurring so far back in geological time that none of the taxa in the sample set of localities had yet evolved".

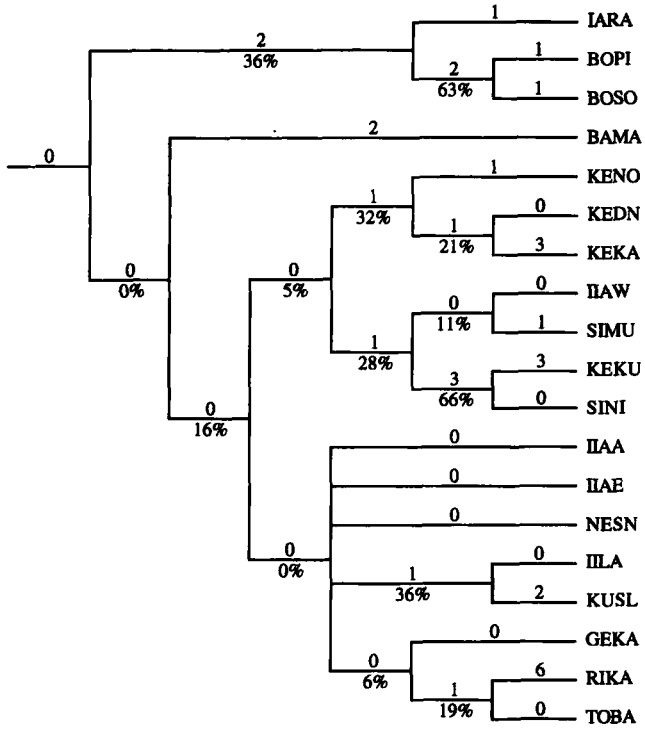

Figure 7. Parsimony analysis of endemism using presence/absence of 24 eurytopic haplotypes in 19 sampling locations. The cladogram is a 50% majority-rule consensus tree computed from 1000 bootstrap replicates of the data using the phylogenetic analysis program PAUP (Swofford, 1993) with the heuristic search option. Numbers above branches are minimum numbers of inferred changes. Numbers below branches are bootstrap values. The tree is rooted by the method of Lundberg (1972), using an ancestral population containing no haplotypes. Tree length = 45; rescaled consistency index = 0.207.

In the present context, it would represent a locality in which none of the alleles in the sample had yet evolved.

Results are presented in Figure 7. Bootstrap values on the tree are generally low, with two exceptions. A bootstrap value of 63% supports the monophyly of the two Budongo Forest locations. A bootstrap value of 66% supports a grouping of a Kibale Forest population (KEKU) with a Semliki Forest population (SINI). Monophyly of the Budongo Forest locations would be expected for obvious reasons. However, monophyly of locations within other multiply-sampled forests (Semliki and Kibale) is not supported.

The PAE cladogram does not support the hypothesis of population relationship schematized in Figure 2. Indeed, the topology presented in Figure 2 was a less parsimonious explanation of the data (length = 47) than the reconstructed PAE tree (length = 45) when analysed using MacClade (Maddison & Maddison, 1992). The PAE tree also lacks long internal branches. This property likely results from the small sample size of characters (24). Because this number is only slightly greater than the number of taxa (19), phylogenetic reconstructions will necessarily be problematic. It is, however, also possible that the pattern observed is real: namely, that the real phylogeny is a 'star phylogeny', and that resolution is low as a result

of the nearly simultaneous genesis of all present-day populations. If so, then area cladograms constructed from other, independent data should show this same pattern.

Fortunately, an independent data set is available for a subset of the populations from which chimpanzees were sampled. Howard (1991) compiled species lists for 12 Ugandan forests using published records, museum collections and extensive ground survey data. Howard recorded the presence or absence of 425 species of forest trees, 329 species of forest birds, 67 species of forest butterflies, and 12 species of forest primates in each of the 12 forests. The data are preliminary in that species absence were not, in most cases, absolutely confirmed. A more detailed description of the data is available in Howard (1991), and more extensive survey work is currently underway. Nine of Howard's 12 forests contain chimpanzees. Seven of these (Bugoma, Budongo, Itwara, Kibale, Kalinzu, Semliki and Rwenzori) were sampled during the present study.

Parsimony analysis of endemicity was performed on Howard's data using PAUP. The branch and bound search option was used, and 1000 branch and bound bootstrap replicates were run to obtain statistical confidences around reconstructed groupings. Characters were unordered and equally weighted. In the case of trees, butterflies and the combined data set, a single most parsimonious tree was found. In the case of birds and primates, more than one equally parsimonious tree was found (2 and 9, respectively). In these latter cases, 50% majority rule consensus trees were calculated. Cladograms were Lundberg rooted (Lundberg, 1972) using a hypothetical ancestral forest containing no taxa (Rosen & Smith, 1988).

Figures 8 and 9 show the PAE cladograms for trees, birds, butterflies and primates, and for the combined data set of 833 taxa, respectively. The cladograms are similar, with well-defined clades (long internal branch lengths) and high bootstrap values. Bootstrap values for the butterfly and primate cladograms were lower than for the tree and bird cladograms, probably because of relatively low numbers of butterflies and primates. In every case, Rwenzori Forest is most divergent. This would be expected considering Rwenzori Forest's unique ecology (Howard, 1991) and the fact that it was probably the location of a montane Pleistocene forest refuge (Hamilton, 1976; Struhsaker, 1981; Rodgers, Owen & Homewood, 1982; Colyn, 1991). The topology of the cladogram of combined biotas (Figure 9) is remarkably robust, having high bootstrap values and long internal branch lengths. It is also generally consistent with the geographic arrangement of forests along the Ugandan forest 'archipelago' (see Figure 2). Centrally-located forests (Itwara, Kibale) branch off after Rwenzori, followed by increasingly more 'peripheral' forests (Kalinzu, Bugoma, Budongo). Semliki is an exception in that, although central (it is contiguous with Rwenzori Forest), it clusters with the most peripheral forests (Budongo, Bugoma).

Howard (1991) presents dendrograms of relationships among these forests derived from the same data. Howard's distance-based analysis groups forests according to the overall similarity of their species compositions, which Howard interprets as indicative of ecological similarity among forests within groups. This analysis is thus fundamentally different from the cladistically-derived relationships presented here, which group forests on the basis of inferred shared ancestral characters (species). The topologies generated by the two analyses differ slightly, primarily in the relationships among 'ingroup' forests (Howard's dendrograms tend to cluster Kibale with Itwara and Semliki with Rwenzori); Rwenzori is maximally divergent by both analyses. This observation suggests that the PAE approach may be informative in

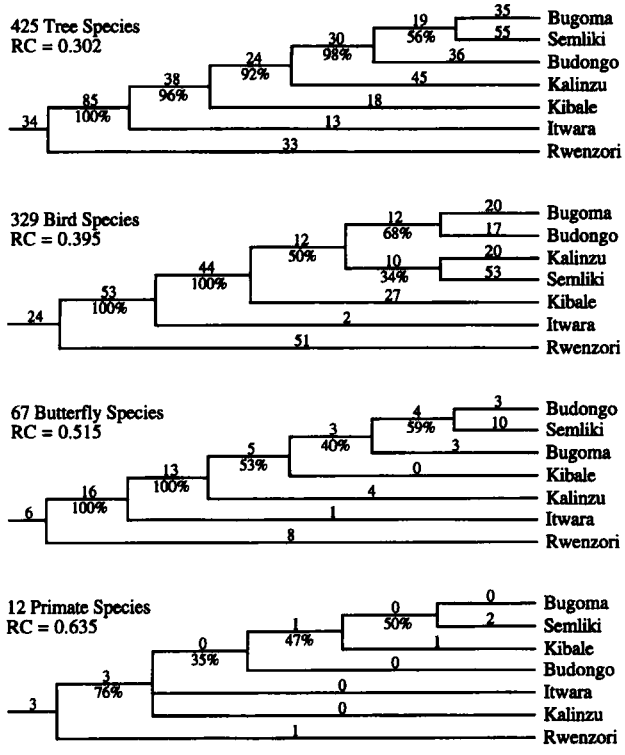

Figure 8. Parsimony analyses of endemism for seven Ugandan forests. Data were taken from Howard (1991) and consist of species lists from each forest, compiled from ground survey data. Cladograms are maximum parsimony trees found using the branch and bound algorithm of PAUP (Swofford, 1993) and rooted by the method of Lundberg (1972) with an ancestral location containing no taxa. Numbers above branches are minimum numbers of inferred changes. Numbers below branches are bootstrap values (1000 replicates). Species were equally weighted. Rescaled consistency indices (RC) are given for each cladogram.

qualitatively different ways than traditional biogeographic approaches, which are based on overall similarities among species assemblages (Humphries *et al.*, 1988).

In any case, the PAE cladogram of eurytopic chimpanzee haplotypes presented in Figure 7 bears little resemblance, either topologically or statistically, to the PAE cladograms of forest taxa presented in Figures 8 and 9. The genetic PAE tree in Figure 7 is even less similar to Howard's (1991) original distance-based ecological dendrograms, in that consistently-divergent forests by Howard's analysis appear as well-ensconced ingroups in Figure 7. To the extent that PAE and ecological grouping accurately reconstruct history, the forests analysed appear to be related by a unique, discernible history which is broadly concordant with the predictions of Pleistocene refuge theory. In contrast, chimpanzee populations inhabiting these forests do not conform to this pattern, and do not cluster into well-defined clades by any analysis. This observation may reflect the inadequacy of the genetic data to discerning such relationships. Alternatively, it may indicate that chimpanzee populations are related by a different history, or that relationships among them are truly ill-defined. The latter pattern would suggest either near simultaneity of population origin, or the obscuring of phylogenetic resolution due to extensive gene flow.

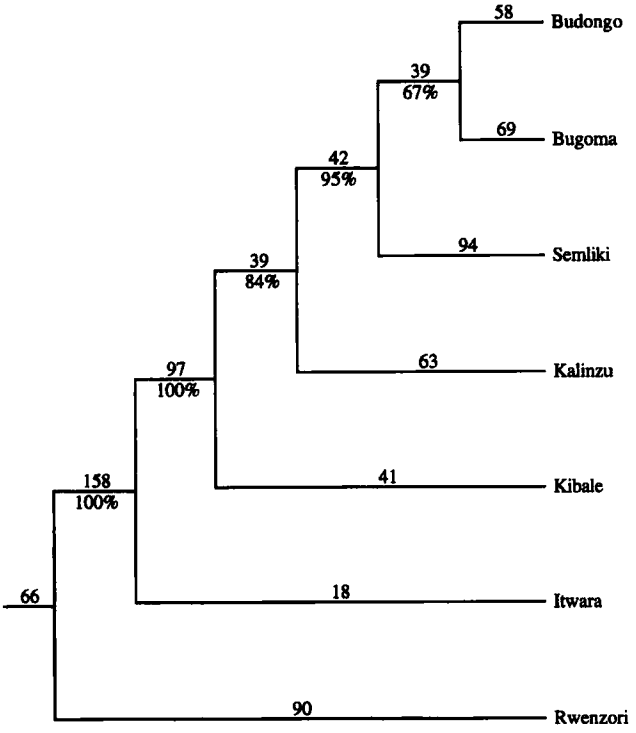

Figure 9. Parsimony analysis of endemism for seven Ugandan forests using combined species data for trees ($n=425$), birds ($n=329$), butterflies ($n=67$) and primates ($n=12$). The cladogram is a maximum-parsimony tree found using the branch and bound algorithm in PAUP (Swofford, 1993), rooted by the Lundberg method (Lundberg, 1972) with an ancestral location containing no taxa. Numbers above branches are minimum numbers of inferred changes. Numbers below branches are bootstrap values (1000 replicates). All characters (species) were equally weighted. Length = 1101; rescaled consistency index = 0.342.

DISCUSSION

Phylogeographic analyses of mitochondrial DNA (Avise *et al.*, 1987; Avise, 1989) have documented intraspecific phylogenetic 'gaps' in other species, often corresponding to obvious geographic barriers (Avise, 1994). Such major subdivisions typically characterize populations of small, relatively non-dispersive mammals (MacNeil & Strobeck, 1987; Plante, Boag & White, 1989; Riddle & Honeycutt, 1990; Prinsloo & Robinson, 1992), birds (Shields & Wilson, 1987; Avise & Nelson, 1989; Edwards & Wilson, 1990; Moore, Graham & Price, 1991; Degnan & Moritz, 1992; Zink & Dittmann, 1993; Seutin *et al.*, 1994), reptiles and amphibians (Densmore *et al.*, 1989; Wallis & Arntzen, 1989; Moritz, 1991; Lamb & Avise, 1992), fishes (Bermingham & Avise, 1986; Avise, 1992), and invertebrates (Sounders, Kessler & Avise, 1986; Hale & Singh, 1991; Murray, Stine & Johnson, 1991; Burton & Lee, 1994). However, even large and/or highly-vagile species can exhibit significant population subdivisioning (e.g. Carr *et al.*, 1986; Baker *et al.*, 1990; Cronin, Nelson & Pac, 1991; Wada, Kobayashi & Numachi, 1991). For example, mitochondrial haplotypes of humpback whales (*Megaptera novaeangliae*), a species of high vagility and

global distribution, sort into distinct geographic clades due to matrilineal fidelity to migratory routes (Baker *et al.*, 1990).

The analyses presented above indicate that *P. t. schweinfurthii* falls into the other category of taxa which do not show 'classical' phylogeographic subdivisioning (e.g. Ball *et al.*, 1988; Cronin *et al.*, 1991; Lehman & Wayne, 1991; Arnason Pálsson & Arason, 1992). Although phylogeographic breaks (representing subspecific divisions) do exist at the *species* level (Morin *et al.*, 1994), no such divergent clades exist within the eastern subspecies. Rather, the eastern chimpanzee mitochondrial pattern is one of low overall variability and little subclustering. This pattern most closely resembles that of humans on a global scale (Cann, Stoneking & Wilson, 1987; Merriwether *et al.*, 1991; Vigilant *et al.*, 1991). Like humans, chimpanzees are a highly-vagile species, as would be expected from their documented long-distance ranging ability (Goodall, 1986; Moore, 1992; Chapman & Wrangham, 1993).

Phylogeographic barriers such as major Zaïrian rivers, which demarcate biogeographic subdivisions within and among other primate genera (Colyn, 1991), have apparently not impeded gene flow within eastern chimpanzees to an extent that would create discontinuities on haplotype trees. However, it is worth noting that IIAA and IIAE (populations 7 and 8 in Figure 1) cluster together more frequently on the population trees presented in this study than do IIAA and IILA (populations 7 and 10) or IIAE and IIAW (populations 8 and 9). This observation is intriguing, since IIAA and IIAE are separated by a large geographic distance, but not by major rivers. The latter two aforementioned population pairs, however, are separated by very small geographic distances, but have large rivers running between them. It is possible, therefore, that rivers may act as barriers, but to a degree or over a time scale to which mitochondrial DNA sequence data is not highly sensitive (Goldberg, 1996).

In light of the lack of effective geographic barriers within *P. t. schweinfurthii*, it is not entirely surprising that population trees, in which geographic structure was imposed on the data, had low bootstrap values. All attempts to analyse the genetic data in this way resulted in topologies that were ambiguous, contradictory and of low statistical confidence. Any phylogenetic signal of the order in which these populations were originally founded (by dispersal or vicariance) has therefore been swamped by the noise of recent migration. This pattern stands in contrast to the inferred history of the forests which these populations occupy. PAE analysis of Howard's (1991) data suggest that relationships among forests are clear, consistent, and statistically well-supported. Furthermore, PAE trees of forest taxa are roughly consistent with the topological predictions of Pleistocene refuge theory.

The possibility does exist that the genetic data are inadequate to the task of detecting population-level evolutionary events of the type predicted by Pleistocene refuge theory. Other, more quickly-evolving loci (e.g. nuclear hypervariable repeat loci) might detect stronger population-level subdivisioning than would mitochondrial DNA, as has been demonstrated for humans (Bowcock *et al.*, 1994). However, mitochondrial DNA is, in many respects, optimal for detecting population subdivisioning (Avise *et al.*, 1987; Avise, 1994). Its pattern of matrilineal transmission would make it especially so for chimpanzees, a species characterized by female dispersal.

To the extent that the mitochondrial data do paint an accurate portrait of chimpanzee phylogeography, chimpanzee genetic evolution has apparently been decoupled from the historical biogeographic processes which have characterized

forest evolution within eastern Africa. This observation supports the notion of *P. t. schweinfurthii* as a highly vagile species, capable of moving across open spaces between forest patches (Kortlandt, 1983; Teleki, 1989; Moore, 1992). If *P. t. schweinfurthii* has existed historically in dry, marginal habitats, then a decoupling of forest history from population history would be expected. Perhaps for the same reasons, chimpanzee populations have remained mitochondrially undifferentiated despite biogeographic barriers such as large gaps between eastern forests, Zaïrian rivers, or the lacustrine system of lakes Edward and George and the Kasinga Channel.

Phylogeographic 'breaks' may currently exist within the subspecies, but, if so, the populations which would define them were not sampled in this study. Populations at the western extreme of the subspecies range (near the confluence of the Zaïre and Ubangi rivers) are likely candidates, since the major rivers which transect these western forests have been instrumental in the subspecific evolution of Colobine and Cercopithecine primates (Colyn, 1987, 1991). Other candidate populations are those in the extreme southwest of the subspecies range (near the headwaters of the Lualaba River), and certain (probably extinct) populations north of the Ubangi River in southern Sudan (Teleki, 1989).

ACKNOWLEDGEMENTS

Richard Wrangham, David Pilbeam and Mark Leighton provided invaluable intellectual advice. Colin and Lauren Chapman, John and Terese Hart, Tuhairwe Peter Atwooki, Katembo Mushgezni Vital, and the faculty and staff of the Makerere University Biological Field Station (Kanyawara, Uganda) and the Centre de Formation et de Recherche en Conservation Forestière (Epulu, Zaïre) were instrumental in facilitating sample collection. Philip Harrell, Elizabeth Williamson, Christine Manning, Ellen Jacobs, Onyach Onecimo and Phillip Morin also helped collect or provide samples.

Laboratory and field research was generously funded by the Wenner-Gren Foundation for Anthropological Research, the L.S.B. Leakey Foundation, the Charles A. and Anne Morrow Lindbergh Foundation, the Mellon Foundation, and the Department of Anthropology at Harvard University. Additional funding was received from Harvard University, the U.S. Department of Education, and the Cora Du Bois Charitable Trust. The Uganda National Council for Science and Technology and the government of Zaïre kindly permitted fieldwork.

REFERENCES

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG. 1981. Sequence and organization of the human mitochondrial genome. *Nature* **290:** 457–465.

Arnason E, Pálsson S, Arason A. 1992. Gene flow and lack of population differentiation in Atlantic cod, *Gadus morhua L.*, from Iceland, and comparison of cod from Norway and Newfoundland. *Journal of Fish Biology* **40:** 751–770.

Avise J, Nelson WS. 1989. Molecular genetic relationships of the extinct Dusky Seaside Sparrow. *Science* **243:** 646–648.

Avise JC. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution* **43:** 1192–1208.

Avise JC. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos* **63:** 62–76.

Avise JC. 1994. *Molecular Markers, Natural History and Evolution.* New York: Chapman and Hall.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC. 1987. Intraspecific Phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18:** 489–522.

Baker CS, Palumbi SR, Lambertsen RH, Weinrich MT, Calambokidis J, O'Brien SJ. 1990. Influence of seasonal migration on geographic distribution of mitochondrial DNA haplotypes in humpback whales. *Nature* **344:** 238–240.

Ball RM, Freeman S, James FC, Bermingham E, Avise JC. 1988. Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. *Proceedings of the National Academy of Science of the United States* **85:** 1558–1562.

Bermingham E, Avise JC. 1986. Molecular zoogeography of freshwater fishes in the south-eastern United States. *Genetics* **113:** 939–965.

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368:** 455–458.

Brundin LZ. 1988. Phylogenetic biogeography. In: Myers AA, Giller PS, eds. *Analytical Biogeography.* London: Chapman & Hall, 343–369.

Burton RS, Lee B-N. 1994. Nuclear and mitochondrial gene genealogies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigriopus californicus. Proceedings of the National Academy of Sciences of the United States* **91:** 5197–5201.

Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. *Nature* **325:** 31–36.

Carr SM, Ballinger SW, Derr JN, Blankenship LH, Bickham JW. 1986. Mitochondrial DNA analysis of hybridization between sympatric white-tailed deer and mule deer in west Texas. *Proceedings of the National Academy of Science of the United States* **83:** 9576–9580.

Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: Models and estimation procedures. *Evolution* **21:** 550–570.

Chapman CA. 1983. Speciation of tropical rainforest primates of Africa: insular biogeography. *African Journal of Ecology* **21:** 297–308.

Chapman CA, Wrangham RW. 1993. Range use of the forest chimpanzees of Kibale: implications for the understanding of chimpanzee social organization. *American Journal of Primatology* **31:** 263–273.

Colyn M. 1987. Les primates des forêts ombrophiles de la cuvette du Zaïre: interprétations zoogéographiques des modèles de distribution. *Revue de Zoologie Africaine* **101:** 183–196.

Colyn M. 1991. L'Importance Zoogéographique du bassin du fleuve Zaïre pour la spéciation: le cas des primates simiens. *Annales de la Musée Royale de l'Afrique Centrale, Tervuren, Belgique* **264:** I–249.

Croizat L. 1958. *Panbiogeography.* Caracas: Published by the author.

Croizat L. 1964. *Space, Time, Form: The Biological Synthesis.* Caracas: Published by the author.

Cronin MA, Amstrup SC, Garner GW, Vyse ER. 1991. Interspecific and intraspecific mitochondrial DNA variation in North American bears (*Ursus*). *Canadian Journal of Zoology* **69:** 2985–2992.

Cronin MA, Nelson ME, PacDF. 1991. Spatial heterogeneity of mitochondrial DNA and allozymes among populations of white-tailed deer and mule deer. *Journal of Heredity* **82:** 118–127.

Degnan SM, Moritz C. 1992. Phylogeography of mitochondrial DNA in two species of white-eyes in Australia. *The Auk* **109:** 800–811.

Densmore LD, Moritz CC, Wright JW, Brown WM. 1989. Mitochondrial DNA analyses and the origin and relative age of parthenogenic lizards (genus *Cnemidophorus*). IV. Nine *sexlineatus*-group unisexuals. *Evolution* **43:** 969–983.

Edwards SV, Wilson AC. 1990. Phylogenetically informative length polymorphism and sequence variability in the mitochondrial DNA of Australian songbirds (*Pomatostomus*). *Genetics* **126:** 695–711.

Emliani C. 1955. Pleistocene temperatures. *Journal of Geology* **63:** 538–578.

Endler JA. 1982. Pleistocene forest refuges: fact or fancy? In: Prance GT, ed. *Biological Diversification in the Tropics.* New York: Columbia University Press, 641–657.

Farris JS. 1983. The logical basis of phylogenetic analysis. In: Plantick N, Funk V, eds. *Advances in Cladistics.* New York: Columbia University Press, 7–36.

Farris JS. 1986. *HENNIG86.* State University of New York at Stony Brook, Department of Ecology and Evolution.

Felsenstein J. 1978. The number of evolutionary trees. *Systematic Zoology* **27:** 27–33.

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39:** 783–791.

Felsenstein J. 1993. *PHYLIP: Phylogenetic Inference Package, version 3.5c.* Department of Genetics, University of Washington, Seattle.

Fitch WM, Margoliash E. 1967. Construction of phylogenetic trees. *Science* **155:** 279–284.

Goldberg T. 1996. Genetics and Biogeography of East African Chimpanzees, *Pan troglodytes schweinfurthii.* Unpublished Ph. D. Thesis, Harvard University.

Goodall J. 1986. *The chimpanzees of Gombe: patterns of behavior.* Cambridge: Harvard University Press.

Grubb P. 1982. Refuges and dispersal in the speciation of African forest mammals. In: Prance GT, ed. *Biological Diversification in the Tropics.* New York: Columbia University Press, 537–553.

Haffer J. 1982. General aspects of the refuge theory. In: Prance GT, ed. *Biological Diversification in the Tropics.* New York: Columbia University Press, 6–24.

Hale LR, Singh RS. 1991. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster.* IV. Mitochondrial DNA variation and the role of history vs. selection in the genetic structure of geographic populations. *Genetics* **129:** 102–117.

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. *Palaeoecology of Africa* **9:** 63–97.

Hamilton AC. 1981. The quarternary history of African forests: its relevance to conservation. *African Journal of Ecology* **19:** 1–6.

Hamilton AC. 1988. Guenon evolution and forest history. In: Gautier-Hion A, Bourliere F, Gautier J-P, Kingdon J, eds. *A Primate Radiation: Evolutionary Biology of the African Guenons.* Cambridge: Cambridge University Press, 13–34.

Hennig W. 1966. *Phylogenetic Systematics.* Urbana: University of Illinois.

Horai S, Satta Y, Hayasaka K, Kondo R, Inoue T, Ishida T, Hayashi S, Takahata N. 1992. Man's place in the Hominoidea revealed by mitochondrial DNA genealogy. *Journal of Molecular Evolution* **35:** 32–43.

Howard PC. 1991. Nature Conservation in Uganda's Tropical Forest Reserves. Gland, Switzerland and Cambridge, U.K: IUCN.

Humphries CJ, Ladiges PY, Roos M, Zandee M. 1988. Cladistic biogeography. In: Myers AA, Giller PS, eds. *Analytical Biogeography.* London: Chapman & Hall, 371–404.

Kendall RL. 1969. An ecological history of the Lake Victoria basin. *Ecological Monographs* **39:** 121–176.

Kingdon J. 1981. Where have the colonists come from? A zoogeographical examination of some mammalian isolates in eastern Africa. *African Journal of Ecology* **19:** 115–124.

Kocher TD, Wilson AC. 1991. Sequence evolution of mitochondrial DNA in humans and chimpanzees. In: *Evolution of Life: Fossils, Molecules and Culture.* Tokyo: Springer, 391–413. Osawa S, Honjo T, eds.

Kortlandt A. 1972. *New Perspectives in Ape and Human Evolution.* Amsterdam: Stichting voor Psychobiologie.

Kortlandt A. 1983. Marginal habitats of chimpanzees. *Journal of Human Evolution* **12:** 231–278.

Lamb T, Avise JC. 1992. Molecular and population genetics aspects of mitochondrial DNA variability in the diamondback terrapin, *Malaclemys terrapin. Journal of Heredity* **83:** 262–269.

Lehman N, Wayne RK. 1991. Analysis of coyote mitochondrial DNA genotype frequencies: estimation of the effective number of alleles. *Genetics* **128:** 405–416.

Lillefors HW. 1967. On the Kolmolgorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association* **64:** 399–402

Lundberg JG. 1972. Wagner networks and ancestors. *Systematic Zoology* **21:** 398–413.

MacNeil D, Strobeck C. 1987. Evolutionary relationships among colonies of Columbian ground squirrels as shown by mitochondrial DNA. *Evolution* **41:** 873–881.

Maddison DR, Ruvolo M, Swofford DL. 1992. Geographic origins of human mitochondrial DNA: phylogenetic evidence from control region sequences. *Systematic Biology* **44:** 111–124.

Maddison WP, Maddison DR. 1992. *MacClade: Analysis of phylogeny and character evolution.* Sunderland, Massachusetts, Sinauer Associates.

Mayr E, O'Hara RJ. 1986. The biogeographic evidence supporting the Pleistocene forest refuge hypothesis. *Evolution* **41:** 55–67.

McGrew WC, Baldwin PJ, Tutin CEG. 1981. Chimpanzees in a hot, dry and open habitat: Mt. Assirik, Senegal, West Africa. *Journal of Human Evolution.* **10:** 227–244.

Merriwether DA, Clark AG, Ballinger SW, Schurr TG, Soodyall H, Jenkins T, Sherry ST, Wallace DC. 1991. The structure of human mitochondrial DNA variation. *Molecular Evolution* **33:** 543–555.

Moeyersons J, Roche E. 1982. Past and present environment. In: Van Noten F, ed. *The Archaeology of Central Africa.* Graz/Austria: Akademische Druck u. Verlagsanstalt, 15–26.

Moore J. 1992. 'Savanna' chimpanzees. In: Nishida T, McGrew WC, Marler P, Pickford M, deWaal FBM, eds. *Topics in Primatology: Human Origins.* Tokyo: University of Tokyo Press, 99–118.

Moore WS, Graham JH, Price JT. 1991. Mitochondrial DNA variation in the Northern Flicker (*Colaptes auratus*, Aves). *Molecular Biology and Evolution* **8:** 327–344.

Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, Woodruff DS. 1994. Kin selection, social structure, gene flow and the evolution of chimpanzees. *Science* **265:** 1193–1201.

Moritz C. 1991. The origin and evolution of pathogenesis in *Heteronotia binoei* (Gekkonidae): evidence for recent and localized origins of widespread clones. *Genetics* **129:** 211–219.

Morrone JJ, Carpenter JM. 1994. In search of a method for cladistic biogeography: An empirical comparison of component analysis, brooks parsimony analysis and three-area statements. *Cladistics* **10:** 99–153.

Murray J, Stine OC, Johnson MS. 1991. The evolution of mitochondrial DNA in *Partula. Heredity* **66:** 93–104.

Nei M. 1972. Genetic distance between populations. *American Naturalist* **106:** 238–292.

Nei M, Tajima F, Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19:** 153–170.

Nelson G. 1981. *Systematics and biogeography/cladistics and vicariance.* New York: Columbia University Press.

Plante Y, Boag PT, White BN. 1989. Microgeographic variation in the mitochondrial DNA of meadow voles (*Microtus pennsylvanicus*) in relation to population density. *Evolution* **43:** 1522–1537.

Platnick NI, Nelson G. 1984. Composite areas in vicariance biogeography. *Systematic Zoology* **33:** 328–335.

Prinsloo P, Robinson TJ. 1992. Geographic mitochondrial DNA variation in the rock hyrax, *Procavia capensis. Molecular Biology and Evolution* **9:** 447–456.

Reynolds J, Weir BS, Cockerham CC. 1983. Estimation of the coancestry coefficient: Basis for a short-term genetic distance. *Genetics* **105:** 767–779.

Rice KA, Donoghue MJ, Olmstead RG. In press. Analyzing large datasets: *rbc*L 500 revisited. *Systematic Biology.*

Riddle BR, Honeycutt RL. 1990. Historical biogeography in North American arid regions: an approach using mitochondrial-DNA phylogeny in grasshopper mice (genus *Onychomys*). *Evolution* **44:** 1–15.

Rodgers WA, Owen CF, Homewood KM. 1982. Biogeography of East African forest mammals. *Journal of Biogeography* **9:** 41–54.

Rogers AR, Jorde LB. 1995. Genetic evidence on modern human origins. *Human Biology* **67:** 1–36.

Rosen BR. 1984. Reef coral biogeography and climate through the late Cainozoic: just islands in the sun or a critical pattern of islands? In: Brenchley PJ, ed. *Fossils and Climate.* London: *Geol. J. Special Issue,* 201–262.

Rosen BR. 1985. Geological hierarchies and biogeographic congruence in the Caribbean. *Annals of the Missouri Botanical Garden* **72:** 636–659.

Rosen BR. 1988. From fossils to Earth history: applied historical biogeography. In: Myers AA, Giller PS, eds. *Analytical Biogeography.* London: Chapman & Hall, 437–481.

Rosen BR, Smith AB. 1988. Tectonics from fossils? Analysis of reef coral and sea urchin distributions from late Cretaceous to Recent, using a new method. In: Audley-Charles MG, Hallam A, eds. *Gondwana and Tethys.* London: Geological Society of London Special Publication, 275–306.

Ruvolo M, Zehr S, vonDornum M, Pan D, Chang B, Lin J. 1993. Mitochondrial COII sequences and modern human origins. *Molecular Biology and Evolution* **10:** 1115–1135.

Saitou N, Nei M. 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406–425.

Seutin G, Klein NK, Ricklefs RE, Bermingham E. 1994. Historical biogeography of the bananaquit (*Coereba flaveola*) in the Caribbean region: a mitochondrial DNA assessment. *Evolution* **48:** 1041–1061.

Shields GF, Wilson AC. 1987. Subspecies of the Canada Goose (*Branta canadensis*) have distinct mitochondrial DNAs. *Evolution* **41:** 662–666.

Sober E. 1988. *Reconstructing the Past: Parsimony, Evolution and Inference.* Cambridge, MA: MIT Press.

Sokal RR, Michener CD. 1958 A statistical method for evaluating systematic relationships. *University of Kansas Scientific Bulletin* **28:** 1409–1438.

Sounders NC, Kessler LG, Avise JC. 1986. Genetic variation and geographic differentiation in mitochondrial DNA of the horseshoe crab, *Limulus polyphemus*. *Genetics* **112:** 613–627.

Struhsaker TT. 1981. Forest and primate conservation in East Africa. *African Journal of Ecology* **19:** 99–114.

Swofford DL. 1993. *PAUP: Phylogenetic Analysis Using Parsimony*. Illinois Natural History Survey, Champaign.

Taylor KC, Lamorey GW, Doyle GA, Alley RB, Grootes PM Mayewski PA, White JWC, Barlow LK. 1993. The 'flickering switch' of late Pleistocene climate change. *Nature* **361:** 432–436.

Teleki G. 1989. Population status of wild chimpanzees (*Pan troglodytes*) and threats to survival. In: Heltne PG, Marquardt LA, eds. *Understanding Chimpanzees*. Cambridge: Harvard University Press, 312–353.

van Zinderen Bakker EM, Coetzee JA. 1972. A re-appraisal of late-Quaternary climatic evidence from tropical Africa. *Palaeoecology of Africa* **7:** 151–181.

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. 1991. African populations and the evolution of human mitochondrial DNA. *Science* **253:** 1503 –1507.

Wada S, Kobayashi T, Numachi K-I. 1991. Genetic variability and differentiation of mitochondrial DNA in Minke whales. In: Hoelzel AR, ed. *Genetic Ecology of Whales and Dolphins*. Cambridge, England: Black Bear Press, 203–215.

Wallis GP, Arntzen JW. 1989. Mitochondrial-DNA variation in the crested newt superspecies: limited cytoplasmic gene flow among species. *Evolution* **43:** 88–104.

Wilkinson L, Hill M, Vang E. 1992. *SYSTAT: Statistics*. Evanston, IL, Systat, Inc.

Zink RM, Dittmann DL. 1993. Gene flow, refugia and evolution of geographic variation in the song sparrow (*Melospiza melodia*). *Evolution* **47:** 717–729.