

Retention of an Ancestral Polymorphism in the Mbuna Species Flock (Teleostei: Cichlidae) of Lake Malawi¹

Paul Moran and Irv Kornfield

Department of Zoology and Center for Marine Studies, University of Maine

Molecular genetic variation was examined within and among species of the rock-dwelling cichlid fishes (mbuna) of Lake Malawi. Phylogenetic relationships among mbuna mitochondrial DNA (mtDNA) haplotypes were estimated by using restriction-fragment-length polymorphisms. The distribution of these lineages among mbuna species is of particular significance for phylogenetic systematic study of this fauna. Some species were found to be polymorphic for divergent haplotypes that substantially predate their isolation from sister taxa. Repeated speciation events among numerous closely related taxa appear to have been so recent that mtDNA lineage sorting among species is incomplete. Thus, the mtDNA gene tree is not congruent with the putative species tree. These results indicate that analysis of mtDNA alone will not be sufficient for resolution of phylogenetic relationships in the mbuna. Clarification of these relationships will require examination of multiple nuclear loci, because many of these new markers are also likely to retain ancestral polymorphisms.

Introduction

The haplochromine cichlid fishes of the East African Great Lakes provide an extreme example of adaptive radiation among vertebrates. In Lake Malawi, >500 cichlid species have radiated to fill virtually every ecological niche (Fryer and Iles 1972; Eccles and Trewavas 1989). Recent molecular systematic studies indicate that the large and diverse endemic Malawi fauna arose within the lake from a single or very few progenitors (Meyer et al. 1990; Moran et al., accepted). It is remarkable that the radiation must have occurred in <2 Myr, the age of Lake Malawi (Banister and Clarke 1980).

The rock-dwelling cichlids of Lake Malawi, the mbuna, are of particular interest to the study of evolutionary biology and behavioral ecology (Fryer and Iles 1972, pp. 496–499; Ribbink et al. 1983; Reinthal 1990a). This group includes ~300 species distributed among 13 genera (with the addition of *Aulonocara*, *Alticorpus*, and *Lethrinops*; Moran et al., accepted). Many of these species are narrow endemics, restricted to single isolated locations that were nonexistent before a post-Pleistocene rise in lake level (Scholz and Rosendahl 1988); many species have thus radiated within the past 25,000 years, perhaps even as recently as the past 200 years (Owen et al. 1990, p. 544). The mbuna represent a premier example of a species flock and provide the opportunity for examination of critical aspects of speciation and radiation (Echelle and Kornfield 1984). To address many specific evolutionary questions regarding modes of speciation and patterns of radiation in the mbuna, a detailed and reliable phylogeny

1. Key words: mitochondrial DNA, lineage sorting, cichlids, species flocks.

Address for correspondence and reprints: Paul Moran, NMFS/CZES, Genetics, 2725 Montlake Boulevard East, Seattle, Washington 98112.

Mol. Biol. Evol. 10(5):1015–1029. 1993.

© 1993 by The University of Chicago. All rights reserved.
0737-4038/93/1005-0006\$02.00

of the group must be available (Kornfield 1991). In particular, phylogenetic information is needed to fully evaluate the significance of morphological and ecological variation.

Various techniques have provided information on phylogenetic systematics and evolutionary ecology of the mbuna, including morphometrics and meristics (Regan 1922; Trewavas 1935; Liem 1981; Reinthal 1990*b*), color and pigment diversity (Ribbink et al. 1983; Eccles and Trewavas 1989, p. 27), and ecology (Fryer 1959; Fryer and Iles 1972, p. 173; Reinthal 1990*a*). However, two factors have complicated the reconstruction of relationships among the haplochromines in general and among the mbuna in particular: (1) the paucity of synapomorphies and (2) the existence of extensive parallelism.

Although the mbuna exhibit substantial diversity in both coloration and trophic morphology, there are few unambiguous synapomorphies that can be used to identify sister-group relationships (Reinthal 1990*b*). Moreover, widespread convergence among the other African Great Lakes suggests a high potential for parallelism within lakes, particularly in trophic morphology (Fryer and Iles 1972, pp. 512–517). Examples of convergence in morphology include the evolution of a crushing pharyngeal mill, hypertrophied lips, enlarged sensory canals, and inclined gape associated with pedophagy (Eccles and Trewavas 1989, pp. 23–24). In addition to morphology, there are multiple examples of ecological convergence, e.g., *Tropheus moorei* (Tanganyika)–*Pseudotropheus zebra* (Malawi) or *Petrochromis* (Tanganyika)–*Petrotilapia* (Malawi) (see Meyer et al. 1990). Construction of an accurate phylogeny for the mbuna requires characters that are less subject to the problems of parallelism (Reinthal 1987, p. 26).

East African haplochromines have been characterized by using a variety of molecular techniques. Allozyme analysis has revealed high genetic similarity among mbuna genera, indicative of their recent origin and rapid radiation [intergeneric I_N values (Nei 1972) = 0.91–0.99; Kornfield 1978]. Inferences about sister-group relationships have been limited by this high similarity, although allozyme analysis has been useful to confirm reproductive isolation between sympatric sibling species (Kornfield 1978; McKaye et al. 1984).

Mitochondrial DNA (mtDNA) has been used to study evolutionary relationships of African cichlid fishes. Restriction-fragment-length polymorphisms (RFLPs) in mtDNA have been used to resolve the relationships among the subspecies of *Oreochromis niloticus* (Seyoum and Kornfield 1992) and to construct a partial phylogeny for members of several taxa in Lake Victoria (Dorit 1986). More recently, mtDNA sequence data were used by Meyer et al. (1990) to produce a phylogeny for cichlids from each of the African Great Lakes, including 24 species in 12 genera from Lake Malawi. Using RFLP analysis of mtDNA, we have previously examined relationships among mbuna and nonmbuna genera from Lake Malawi (Moran et al., accepted); both sequencing and RFLP studies of mtDNA have provided a clearer picture of higher-order systematic relationships among the Malawi cichlids. However, while some organization was evident within the mbuna, little detail was provided by this previous research.

The current study expands the results of previous work, with a more detailed view of the mbuna, by examining RFLPs in mtDNA of more species and by sampling more individuals within species. Although substantial differentiation was detected among mbuna haplotypes, much of the potential phylogenetic information from mtDNA may be compromised because of incomplete mtDNA lineage sorting and the retention of ancestral polymorphism.

Material and Methods

RFLPs were used to characterize genetic variation among mtDNA lineages of the rock-dwelling mbuna of Lake Malawi. Thirty species representing 11 genera were collected from 13 locations (table 1). This sample represents all currently recognized genera of mbuna except *Gephyrochromis*. Each species was represented by one to eight individuals (82 total). Two additional taxa (*Astatotilapia calliptera* and *Copadichromis*

Table 1
Material Examined and Collection Locations

Acronym (<i>N</i>)	Taxon	Collection Location ^a
Mbuna taxa:		
ADE (1)	<i>Aulonocara</i> 'deep water' ^b	Maledeco Fishery
AJA (1)	<i>Aulonocara jacobfreibergeri</i> ^b	(Aquarium trade)
AMI (1)	<i>Aulonocara</i> 'minutus' ^{b,c}	Monkey Bay
ANY (1)	<i>Aulonocara nyassae</i> ^b	Thumbi Island West
COB (2)	<i>Cyathochromis obliquidens</i>	Thumbi Island West
CAM (1)	<i>Cynotilapia afra</i>	Thumbi Island West
GME ₁ (1)	<i>Genyochromis mento</i>	Thumbi Island West
GME ₂ (1)	<i>G. mento</i>	Otter Point
ISP (2)	<i>Iodotropheus sprengerae</i>	(Chinyamwezi)
LFU (2)	<i>Labeotropheus fuelleborni</i>	(Mbenji Island)
LFR (1)	<i>Labidochromis freibergeri</i>	Thumbi Island West
LVE (1)	<i>Labidochromis vellicans</i>	Thumbi Island West
LGO (2)	<i>Lethrinops gossei</i> ^a	Monkey Bay
MAU (1)	<i>Melanochromis auratus</i>	(Mbenji Island)
PTR (2)	<i>Petrotilapia tridentiger</i>	Nkhata Bay
PBA (3)	<i>Pseudotropheus barlowi</i>	Maleri Island
PLE (1)	<i>Pseudotropheus elegans</i>	Chimumbo Bay
PEL (2)	<i>Pseudotropheus elongatus</i>	Nkhata Bay
PHE (3)	<i>Pseudotropheus heteropictus</i>	Thumbi Island West
PTB (2)	<i>Pseudotropheus tropheops</i> 'black'	Nkhata Bay
PTO (2)	<i>Pseudotropheus tropheops</i> 'orange chest'	Thumbi Island West
PWI (2)	<i>Pseudotropheus williamsi</i>	Nakantenga Island
PXA (6)	<i>Pseudotropheus xanostomachus</i>	Maleri Island
PZA (2)	<i>Pseudotropheus zebra</i> 'aggressive'	Maleri Island
PZC (1)	<i>Pseudotropheus zebra</i> 'BB'	Chilumba Point
PZM (4)	<i>Pseudotropheus zebra</i> 'BB'	Mumbo Island
PZN (5)	<i>Pseudotropheus zebra</i> 'BB'	Nkhata Bay
PZT (4)	<i>Pseudotropheus zebra</i> 'BB'	Thumbi Island West
PZB (5)	<i>Pseudotropheus zebra</i> 'black dorsal'	Maleri Island
PZL (6)	<i>Pseudotropheus zebra</i> 'blue'	Maleri Island
PZO (2)	<i>Pseudotropheus zebra</i> 'cobalt'	Nkhata Bay
PZG (2)	<i>Pseudotropheus zebra</i> 'gold'	Nkhata Bay
PZE (2)	<i>Pseudotropheus zebra</i> 'mbenji'	Mbenji Island
PZR (8)	<i>Pseudotropheus zebra</i> 'red dorsal'	Nakantenga Island
Nonmbuna outgroup taxa:		
ACA (1)	<i>Astatotilapia calliptera</i>	Thumbi Island West
CML (1)	<i>Copadichromis mloto</i>	Chimumbo Bay

^a For specimens obtained through the aquarium trade, the original collection locations, where available, are in parentheses.

^b Considered a member of the mbuna complex, on the basis of the results of previous mtDNA restriction analysis; bootstrap resampling indicated that this species was embedded within a monophyletic group that contained 11 mbuna species representing seven genera (Moran et al., accepted).

^c Identified in the field as *Trematocranus* 'minutus'; however, we refer to it here as *Aulonocara*, following the subsequent revision by Eccles and Trewavas (1989).

mloto) were included for outgroup comparison. mtDNA, isolated from fresh and frozen tissue, was purified by density gradient ultracentrifugation (Lansman et al. 1981; Dowling et al. 1990). Twelve restriction enzymes were selected that showed synapomorphic variation within the mbuna, on the basis of results of earlier work (Moran et al., accepted). Restriction fragments were separated by agarose gel electrophoresis (0.5%–2.0%) and were visualized by 5' end-labeling with α - ^{32}P or ^{35}S dNTPs (Drouin 1980; Dowling et al. 1990) and autoradiography (XO-mat AR film; Kodak). Fragment sizes were estimated by comparison to a size standard included on each gel. Haplotype designations (haploid genotypes) were based on the unique composite of all restriction profiles (table 2). Thus, the presence or absence of each fragment observed for all enzymes constitutes the character states that define an individual's haplotype. Many of the individual restriction profiles were observed previously in a broader survey of the Malawi cichlids (Moran et al., accepted); the designations used in that study are retained here for consistency. To assure fragment homology, alternative restriction profiles were visualized on the same gels, usually side by side. Furthermore, nearly all restriction profiles could be explained by single site changes from one or more other profiles, placing additional confidence in the assignment of fragment homologies.

Relationships among haplotypes were examined both phenetically and cladistically. Phenetic analyses were conducted by using estimates of nucleotide diversity among haplotypes (the number of base substitutions per nucleotide, d_{ij} ; Nei and Li 1979). These relationships were visualized by using both clustering and ordination. Clustering results were presented as the strict consensus of 20 single- and complete-linkage dendrograms. This conservative view of phenetic relationships forms groups in which the members are closer to each other than to any taxon outside the cluster. These groups (ball clusters) are likely to emerge from most clustering procedures (Rohlf 1990). To examine the genetic distances among haplotypes, we conducted principal coordinates analysis and projected the haplotypes on the first three principal axes. A minimum spanning tree (MST) was also calculated and superimposed on the principal coordinates projection. The MST is useful because it can reveal local deformations in principal coordinates space. Haplotypes that appear rather distinct on the basis of the first three principle axes yet that are connected by the MST may be closer in principle coordinates space when additional dimensions are considered (Rohlf 1970). Clustering and ordination were achieved by using NTSYS 1.6 (Rohlf 1990). Input files for all analyses were constructed by using REAP (McElroy et al. 1992).

The principal coordinates projection is important to this study, for two reasons. First, a consensus method was required for clustering, and thus no meaningful scale of genetic distance can be assigned to the levels of the nodes in the single- and complete-linkage dendrogram (fig. 1). Principal coordinates analysis provides a three-dimensional interpretation of relative genetic distances. Second, agglomerative methods tend to produce clusters regardless of the structure of the data, whereas ordination maximizes the orthogonal separation among haplotypes. Focusing attention on the relationships that emerge from both clustering and ordination can help prevent overinterpretation of these results.

Cladistic analysis was achieved by using restriction fragments as characters and Wagner parsimony. Appropriate outgroup taxa were selected on the basis of a previous study (Moran et al., accepted). The cladistic results are presented as the strict consensus of 1,000 equal length trees produced with the tree-bisection-reconnection algorithm of PAUP, version 3.0 (Swofford 1989). Bootstrap resampling was used to heuristically evaluate the strength of specific clades of the resulting tree (PHYLIP 3.1; Felsenstein

Table 2

Composite mtDNA Genotypes Observed in Mbuna Taxa and Outgroup Taxa

Designation	Composite Genotype ^a												Species Acronym(s) ^b	
Mbuna taxa:														
α1	C	C	C	C	C	B	C	C	C	C	C	C	LFR, LVE, PZR, PZN, PZM	
α2	C	C	8	C	C	B	C	C	C	C	C	C	PZL	
α3	C	C	G	C	C	B	C	C	C	C	C	C	PXA	
α4	C	C	C	C	E	B	C	C	C	C	C	C	PEL	
α5	C	C	C	C	C	B	C	C	H	C	C	C	PEL	
α6	C	C	C	C	C	B	D	C	B	C	C	C	PTB, PZC	
α7	C	C	C	C	C	B	C	C	C	E	C	C	PZA	
α8	C	C	C	C	C	B	C	C	I	I	C	C	PZG	
α9	C	C	C	C	C	B	D	C	J	C	C	C	PZO	
α10	C	C	8	C	D	B	C	C	L	C	C	C	PZL	
α11	V	C	C	C	C	B	D	C	B	C	C	C	PHE	
α12	V	C	C	C	C	B	D	F	B	C	C	C	PHE, PBA	
α13	C	C	G	C	C	B	D	F	B	C	C	C	PZN	
α14	G	C	H	C	C	B	D	C	B	F	C	C	PZO	
α15	C	D	C	C	C	B	C	C	C	E	B	C	PZM	
β ₁	D	D	D	E	C	B	C	C	K	E	B	D	PZL	
β ₂	D	D	D	E	C	B	C	C	T	E	B	D	PZL	
β ₃	D	D	D	E	C	C	C	C	F	E	B	D	PXA	
β ₄	J	D	D	E	C	C	C	C	F	E	B	D	LFU	
β ₅	J	D	6	E	C	C	C	C	F	E	B	D	PZM	
β ₆	J	D	D	E	C	C	C	C	W	E	B	D	COB	
β ₇	J	D	D	E	C	C	C	C	A	E	B	D	ISP	
β ₂₁	D	D	C	C	C	C	C	C	A	E	B	D	ADE, LFU, LGO, PWI, PXA, PZL	
β ₂	D	D	C	C	C	C	C	C	A	G	B	D	AJA	
β ₂₃	D	D	G	C	C	C	C	C	A	E	B	D	LGO	
β ₂₄	D	D	N	C	C	C	C	C	A	E	B	D	CAF	
β ₂₅	F	D	C	C	C	C	C	C	A	E	B	D	PWI	
β ₂₆	D	D	C	E	C	C	C	G	A	E	B	D	ISP	
β ₂₇	J	D	C	C	C	C	C	C	A	E	B	D	COB	
β ₂₈	D	D	C	C	C	C	C	C	U	E	B	B	AMI	
β ₂₉	K	D	C	C	C	C	C	C	D	E	B	D	PTB	
β ₂₁₀	D	B	I	C	C	C	C	C	V	E	B	D	ANY	
β ₃₁	F	D	M	C	D	C	C	C	E	E	B	D	GME, PZB	
β ₃₂	F	D	M	C	D	C	C	D	E	E	B	D	PHE, PZR, PBA	
β ₃₃	F	D	M	C	D	C	C	D	G	E	B	D	PZR	
β ₃₄	F	D	M	C	D	C	E	C	G	E	B	D	PTR	
β ₃₅	F	D	F	C	D	C	E	C	E	E	B	D	PTR	
β ₃₆	E	D	M	C	D	C	C	C	E	E	B	D	PZT	
β ₃₇	E	D	M	C	D	C	C	F	E	E	B	D	PTO	
β ₃₈	1	D	M	C	D	C	C	F	E	E	B	D	PTO	
β ₃₉	E	D	M	C	D	C	C	C	G	E	B	D	PZT	
β ₃₁₀	E	D	S	C	D	C	C	C	C	E	E	B	D	PZT
β ₃₁₁	D	D	M	C	D	C	C	C	E	E	B	D	PZE	
β ₃₁₂	D	D	M	C	C	C	C	C	E	E	B	D	PZE	
β ₃₁₃	W	D	M	C	D	C	C	C	E	E	B	D	PLE	
β ₃₁₄	F	D	E	E	D	C	C	C	E	I	B	D	MAU	
Nonmbuna outgroup taxa:														
ACA	I	G	W	F	J	C	C	I	X	N	B	H	<i>Astatotilapia calliptera</i>	
CML	B	H	A	E	D	C	C	D	Y	L	B	H	<i>Copadichromis mloti</i>	

^a Alphanumeric designations (left to right) refer to the restriction profiles for *ApaI*, *AvaI*, *Avall*, *BamHI*, *BclII*, *DraI*, *EcoRV*, *HindIII*, *MbolI*, *NciI*, *SmaI*, and *StyI*, respectively.

^b Refer to table 1.

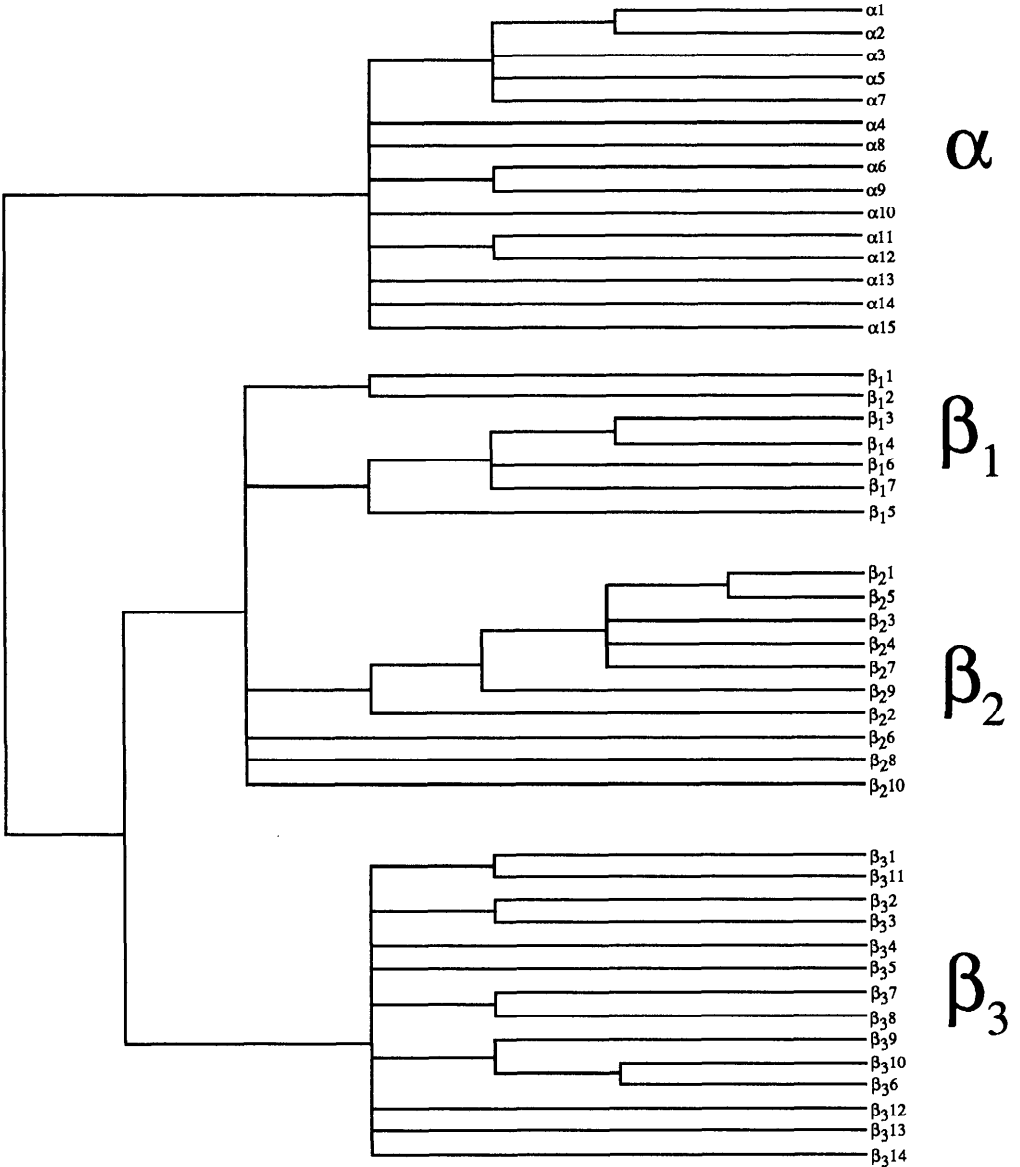


FIG. 1.—Phenetic relationships based on mtDNA sequence divergence (Nei and Li 1979) among mbuna haplotypes. Clustering is presented as the strict consensus of 20 single- and complete-linkage phenograms. This is a conservative clustering method, which yields groups in which all members are closer to each other than to any taxon outside the cluster (Rohlf 1990). Because this is a consensus, the scale of the levels is arbitrary, and only subset information is depicted. For relative genetic distances among haplotypes, see fig. 2.

1988). Five hundred bootstrap replicates were conducted, with variation in both the random-number seed and the input order of taxa. Although the statistical legitimacy of bootstrap analysis has been questioned (Swofford and Olsen 1990), we feel that this procedure is useful in evaluating the *relative* strengths of particular clades.

Results

The 12 restriction enzymes used in this study generated an average of 70 fragments for each of the 82 individuals examined. In aggregate, this sample represents ~ 355 bp, or 2.2% of the 16,300-bp mbuna mtDNA genome. No heteroplasmy or variation in genome size was observed. Forty-six unique haplotypes were revealed among the 30 species of mbuna (table 2). Fragment data are available on request.

Cladistic and phenetic analyses both consistently distinguished two major groupings of haplotypes within the mbuna. We refer to these groups as " α " and " β ." Together, the α and β groups constitute the B (mbuna) clade identified in two earlier studies (Meyer et al. 1990; Moran et al., accepted). The average sequence divergence between haplotypes of the α and β groups was 1.8%. The nucleotide diversity estimates within the α and β groups were both $\sim 1.1\%$. These estimates are upwardly biased somewhat because, on the basis of earlier work, enzymes that produced monomorphic and autapomorphic restriction fragments were excluded from consideration (Moran et al., accepted). The α/β distinction was evident in both clustering (fig. 1) and ordination (fig. 2). Further, these phenetic analyses suggested the presence of three subgroups within β (referred to as " β_1 "–" β_3 "). The designation of these groups is meant to emphasize the distinctiveness of α from all members of β and to indicate substantially lower levels of divergence between the β subgroups, particularly between β_1 and β_2 . Within both the α and β groups there are examples of local deformations in principle coordinate space revealed by the connection of nonadjacent OTUs with the MTS. This result is due to the fact that a small but significant amount of information is contained in the next five principle axes. The uncertainty regarding relationships within

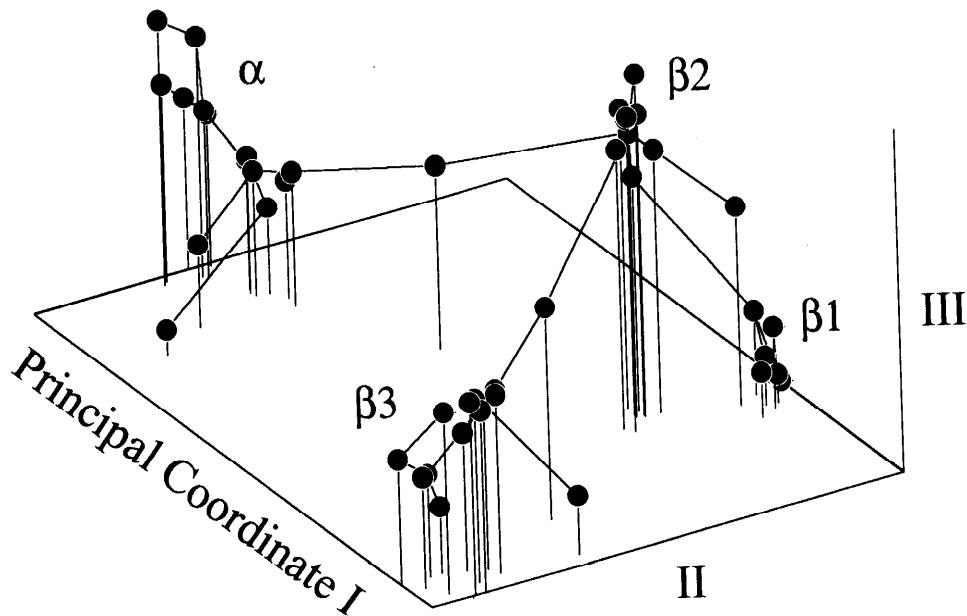


FIG. 2.—Principal coordinates projection of sequence divergence among mtDNA haplotypes. The first three principle axes describe 72% of the variation observed. The next five eigenvectors are significant and account for an additional 13% of the total variation. A minimum spanning tree is superimposed on the projection to assist in visualization of phenetic relationships (Rohlf 1970). The separation of α and β haplotypes is evident, as is the reduced differentiation among the β subgroups.

the major α and β isolates is further reflected by the lack of resolution in the complete- and single-linkage consensus cluster analysis (fig. 1).

Cladistic results were broadly concordant with phenetic relationships. The haplotypes of the α group were all very closely related and clearly divergent from members of the β group (bootstrap level 100%; fig. 3). As in the phenetic analysis, the β group formed a loose assemblage of three subgroups. The lower level of divergence among the β subgroups was clearly evident from the relatively low bootstrap values obtained for those clades.

Although the discrimination of the α and β haplotypes was unequivocal, one specimen (*Pseudotropheus zebra* 'BB' from Mumbo Island) possessed a haplotype (α_{15}) that was almost exactly intermediate between α and β_2 (fig. 2). Of five enzymes whose fragments were diagnostic for the α/β dichotomy, three revealed α restriction profiles in this individual, whereas two showed β profiles. [It is possible to explain the position of this haplotype with only two or three site reversals; yet, the fragment profiles for two of the three enzymes in which the reversals might have occurred (*AvaI* and *SmaI*) were highly conserved, and, with the exception of a single fish, each enzyme revealed only two fragment profiles, one corresponding to α and the other corresponding to β .]

The relatively divergent α and β lineages exhibited an unexpected distribution among the mbuna; many species contained both lineages, often within the same population (table 3). That is, the α and β mtDNA lineages represent an intraspecific polymorphism that extends across species boundaries.

Discussion

The most significant result of this study is the observation that many mbuna are polymorphic for divergent mtDNA lineages (table 3). A number of species that exhibit this polymorphism are restricted to habitats in southern Lake Malawi that certainly did not exist before the late-Pleistocene low-lake-level stand (25 ka >200 m lower; Scholz and Rosendahl 1988). Some habitats now containing strict local endemics may even have been unavailable during the low stand of the late 18th century (Owen et al. 1990). Thus, divergence of the α and β mtDNA lineages predates the radiation of many local endemic species. This finding has significant implications for the study of evolutionary processes in the mbuna and for future application of mtDNA analysis in this fauna.

Previous mtDNA studies in Lake Malawi have indicated the presence of two major mtDNA lineages separated by 2.5%–3.5% sequence divergence (A and B; Meyer et al. 1990). The A lineage consists of a diverse group of sand-dwelling and pelagic species, while the rock-dwelling mbuna constitute the B lineage. In addition to these two groups, a few additional "oligotypic" lineages have been identified, each of which is characterized by only one or a few species (Moran et al., accepted). The mtDNA restriction results presented here describe further structuring within the mbuna lineage. The mbuna α lineage represents a distinct and unified clade, strongly supported by bootstrap analysis (fig. 3) and evident in both clustering and ordination.

Shared Polymorphism among Species

Sharing of diverse mtDNA polymorphisms across species boundaries is often explained as the result of secondary contact and introgressive hybridization (Ferris et al. 1983; Dowling et al. 1989; Dowling and Hoeh 1991). In many cases this conclusion is supported by behavioral observations, allozyme data, or the presence of morpho-

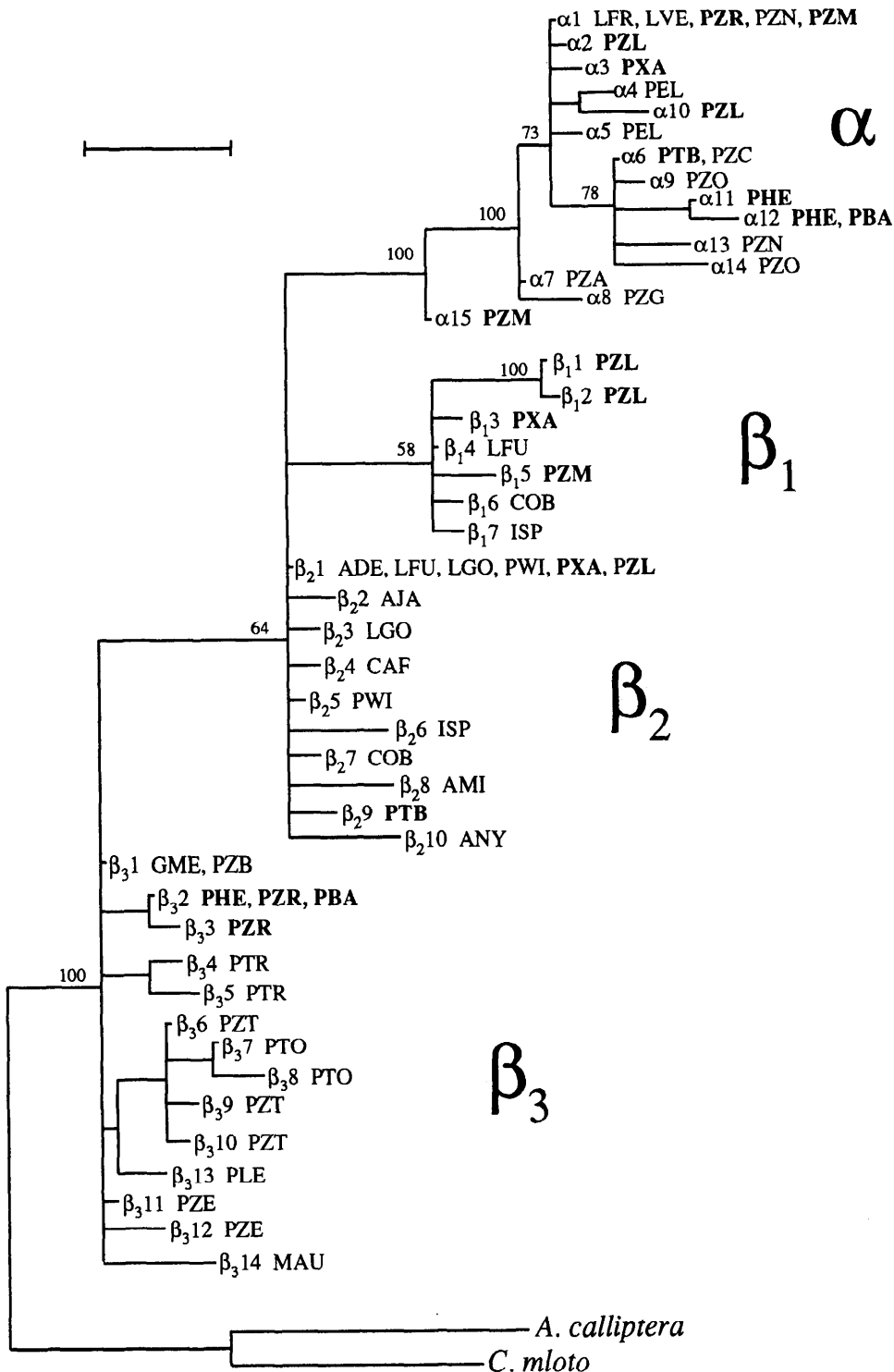


FIG. 3.—Wagner parsimony dendrogram illustrating cladistic relationships among mbuna mtDNA haplotypes. A strict consensus is presented of 1,000 equal length trees (231 steps) produced by the tree-bisection-reconnection algorithm (Swofford 1989). Branch lengths are drawn to approximate scale (except that, for readability, zero lengths are shown as half steps), and the values at selected nodes indicate the percentage of 500 bootstrap replicates that united the haplotypes of that clade (Felsenstein 1988). The scale bar indicates 10 character-state changes. Species acronyms (from table 1) shown in boldface designate those taxa that were polymorphic for α and β lineages (table 3). The tree is rooted by using *Astatotilapia calliptera* and *Copadichromis mloto*.

Table 3
Mbuna Taxa Polymorphic for α and β mtDNA Lineages

Taxon	Location	α	β_1	β_2	β_3
<i>Pseudotropheus barlowi</i> . . .	Maleri Island	1	0	0	2
<i>P. heteropictus</i>	Thumbi Island West	2	0	0	1
<i>P. tropheops</i> 'black'	Nkhata Bay	1	0	1	0
<i>P. xanstromachus</i>	Maleri Island	2	2	2	0
<i>P. zebra</i> 'BB'	Mumbo Island	3 ^a	1	0	0
<i>P. zebra</i> 'blue'	Maleri Island	2	3	1	0
<i>P. zebra</i> 'red dorsal'	Nakantenga Island	1	0	0	7

NOTE.—A subsequent population study has revealed that *P. zebra* 'BB' at Nkhata Bay is also polymorphic for α and β lineages (authors' unpublished data).

^a Includes one individual that was approximately intermediate between members of the α and β_2 mtDNA lineages (see text).

logical intermediates in contact zones. However, theoretical and empirical work, particularly of Avise, Nei, and their respective co-workers, has directed increasing attention toward mtDNA lineage sorting and the retention of ancestral polymorphism as an explanation for this phenomenon (Avise et al. 1983; Tajima 1983; Neigel and Avise 1986; Nei 1987, p. 277; Pamilo and Nei 1988). Other explanations for shared polymorphism, such as species misidentification or convergence (Dowling et al. 1989), can be virtually ruled out here by the fact that the α/β polymorphism is so widespread both taxonomically and geographically.

Lineage sorting is the elimination of ancestral polymorphism such that sister species become monophyletic with respect to mtDNA. A polymorphic ancestor with large effective female population size ($N_{f(e)}$) and multiple divergent haplotypes may give rise to sister species, two or more of which inherit the same polymorphism. Until these taxa reach monophyly, and until all mtDNA diversity postdates speciation, lineage sorting is incomplete, and the relationships implied by the mtDNA gene tree may be inconsistent with the species tree (Pamilo and Nei 1988), particularly when small samples sizes are used.

We propose that the α/β polymorphism is shared among mbuna species because of incomplete lineage sorting rather than because of introgressive hybridization. Introgression is unlikely, for at least three reasons. First, interspecific courtship has never been reported in the mbuna, despite extensive in situ observation (Holzberg 1978; Marsh et al. 1981; P. N. Reinthal, personal communication; I. Kornfield, unpublished data). Second, mbuna hybrids have not been recovered, despite extensive collecting (Fryer and Iles 1972, p. 167). Some artificially produced hybrids exhibit distinctive, mosaic morphology (McElroy and Kornfield, accepted), yet these phenotypes have not been reported. Indeed, numerous pairs of sympatric sibling species show significant differences in allele (Kornfield 1978; McKaye et al. 1984) or mtDNA haplotype frequencies (authors' unpublished data), consistent with reproductive isolation. Finally, the α/β polymorphism was observed in seven of nine taxa that were examined for more than two individuals, and it was distributed over a broad geographic area. Thus, limited hybridization events in a few taxa could not explain these results. Hybridization episodes under past ecological conditions cannot be ruled out (Crapon de Caprona 1986). However, hybridization, if it actually occurred, is probably an old event, predating the origin of many species. Regardless, current behavior, morphology, and

genetic differentiation are not consistent with hybridization as an explanation for shared lineages among species. Instead, we suggest that these species retain a polymorphism that was present in the common ancestor of the mbuna species flock. The retention of this ancestral polymorphism implies that insufficient time has elapsed, since speciation, for lineage sorting to have reached completion.

Computer modeling has demonstrated the significant role played by demographics in determining the persistence of an ancestral polymorphism (Avise et al. 1984). Two critical factors for retention of multiple ancestral lineages are population growth rate (r) and population size. Population size is critical both initially and at carrying capacity; simply put, large or expanding populations are likely to retain multiple lineages. In stable populations the survival probability of multiple lineages declines continually with time until, at $4N_{f(e)}$ generations, the probability is essentially 0. In expanding populations, however, the probability declines initially but levels off and reaches an asymptote at a value that is a function of r and $N_{f(e)}$. In this case, a large carrying capacity is far more significant than the initial number of founders, in maintaining a high probability of retained polymorphism. An additional factor important to the probability of survival of multiple mtDNA lineages is variation in female fecundity. Low variation in fecundity increases the probability that multiple ancestral lineages will be retained (Avise et al. 1984).

Population characteristics of mbuna can be examined in order to explore the likelihood that incomplete lineage sorting explains the shared α/β polymorphism among species. The more consistent that those characteristics are with the conditions that promote the retention of ancestral polymorphisms, the more compelling incomplete lineage sorting becomes. However, because of the stochastic nature of the lineage-sorting process, these observations can only be considered heuristic, suggesting avenues for future research. The work of Avise et al. (1984), described above, leads us to examine the following questions: How large were founding populations? Have populations increased substantially in size since founding? and, How much variation do mbuna species exhibit in female fecundity?

It is unclear how large the founding populations were likely to have been for most species of mbuna. In many cases, however, mbuna populations occupy relatively new habitats that became available some time after the post-Pleistocene rise of the lake level (Scholz and Rosendahl 1988). These populations may have exhibited a net expansion for some time after introduction. Current population sizes of many mbuna species are quite high, and, although habitat area is often confined by inappropriate substrate, densities within those areas may exceed 8 fish/m² (Ribbink et al. 1983). Indeed, populations of some species (e.g., *Pseudotropheus zebra* 'BB') frequently number in the hundreds of thousands. Thus, it appears that, for at least some population parameters, a number of mbuna species exhibit characteristics that are consistent with prolonged survival of ancestral lineages.

In contrast to large population size and recent expansion, which increase the probability of lineage survival, some mbuna species appear to exhibit high variability in female fecundity—which should have the opposite effect, tending to accelerate lineage extinction. For example, the number of eggs contained in the ovaries of female *P. zebra* 'BB' at Monkey Bay were 13–41 (mean 26.7, standard deviation = 10.18; Kornfield 1974). However, fecundity in haplochromines is correlated with body size (Goldschmidt 1989, p. 102). Further, because adult females are not territorial, it is probable that most individuals who survive to maturity will reproduce. Thus, intra-specific variation in female fecundity may be less extreme than the coefficient of vari-

ation (i.e., 38) might suggest. The cumulative evidence against hybridization, combined with the population parameters described above, suggests that incomplete lineage sorting is more parsimonious than secondary contact, as an explanation for the shared α/β polymorphism among species.

Evolution of the Mbuna Flock

The 13 currently described genera of mbuna exhibit substantial trophic and ecological diversity (Trewavas 1935; Fryer and Iles 1972, pp. 66–72), yet the estimated sequence divergence between mbuna haplotypes for some pairs of morphologically divergent species is minute (Moran et al., accepted). Highly derived species such as *Genyochromis mento*, a specialized scale and fin eater, were found to contain mtDNA haplotypes that were identical to those of primitive algivores such as *P. zebra* 'black dorsal.' *Lethrinops gosseii*, a deep-water species not previously recognized as a member of the mbuna flock, had the same mtDNA haplotype as did both *P. xanostomachus* and *P. zebra* 'blue' (table 2). The low levels of mtDNA nucleotide divergence among these species underscore both the relative recency of the mbuna radiation and the rapidity of morphological diversification.

Further Systematic Application of mtDNA

There are four major implications of the retention of ancestral polymorphism. First, mtDNA will be of limited utility for reconstruction of phylogenetic relationships among the mbuna. mtDNA analysis has been useful for higher systematic studies of the more divergent genera of Lake Malawi and of relationships of cichlid faunas among lakes in East Africa (Meyer et al. 1990; Sturmbauer and Meyer 1992; Moran et al., accepted). However, additional molecular markers will have to be derived for the closely related mbuna in order to achieve adequate sensitivity for complete phylogenetic reconstruction.

Second, apparent regional fixation for specific haplotypes must be interpreted with caution. Because haplotype recovery is sample-size dependent, many taxa may prove to be polymorphic on further sampling (table 3). For example, our initial findings suggested that the distribution of α and β haplotypes was geographically structured. Multiple species appeared to be fixed for α in the northern part of the lake and for β in the south. However, expanded sampling revealed that the polymorphism was both geographically and taxonomically widespread. Haplotype frequency distribution may reflect patterns of evolutionary divergence; drift appears to have produced some regional differences in haplotype frequencies (authors' unpublished data). However, the possibility for convergence of frequencies is considerable, and it remains unclear how much systematic information these differences will provide (Crother 1990). The results presented here suggest that the α/β polymorphism may be widespread among the mbuna. Further, in light of the similar levels of sequence divergence among taxa within the nonmbuna A lineage of Lake Malawi (Moran et al., accepted), some of those species may also exhibit incomplete lineage sorting.

Third, greater resolution of mtDNA variation through the use of additional restriction enzymes, site mapping, or direct sequencing may increase the number of alleles sampled but will be of only limited utility; mtDNA is nonrecombining and therefore behaves as a single locus. Examination of additional alleles at a single locus provides a more accurate gene tree, but only through the examination of multiple, independent loci can a more accurate species tree be obtained (Pamilo and Nei 1988).

Higher resolution of the mtDNA gene tree will not necessarily yield resolution of the mbuna species tree.

Finally, nuclear loci are subject to the same limitations resulting from incomplete lineage sorting as those described above for mtDNA. Indeed, nuclear loci are considerably slower to reach allelic fixation, because of their fourfold-larger effective population size (Nei 1987, pp. 186–187). Thus, studies of multiple nuclear loci are likely to reveal varying degrees of lineage sorting. At most nuclear loci, mbuna species will probably not exhibit monophyly. Allozyme results demonstrate relatively few fixed alleles among common species of mbuna (Kornfield 1978). Multiple independent molecular markers will be required for a complete phylogenetic reconstruction of the mbuna species flock.

Acknowledgments

We acknowledge the support and cooperation of the Malawi Department of Fisheries, Malawi Department of Parks, and the National Research Council of Malawi. S. Grant provided continuous assistance with fieldwork. P. Reinthal assisted in field collection, logistic support, identification of specimens, and encouragement in this research. We thank T. Dowling for introducing one of us (I.K.) to mtDNA analysis and for assisting with preliminary research. We thank W. Glanz, D. McElroy, and J. Ringo for commenting on the manuscript. Preliminary communication of the results of this research was made at the ASIH annual meetings in New York, June 1991. This research was supported by the National Science Foundation grant BSR 84-16131, by the Smithsonian Oceanographic Sorting Center, and by the University of Maine. Manuscript preparation was supported by National Marine Fisheries Service, Coastal Zone and Estuarine Studies Division, Northwest Fisheries Science Center.

LITERATURE CITED

- AVISE, J. C., J. E. NEIGEL, and J. ARNOLD. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.* **20**:99–105.
- AVISE, J. C., J. F. SHAPIRA, S. W. DANIEL, C. F. AQUADRO, and R. A. LANSMAN. 1983. Mitochondrial DNA differentiation during the speciation process in *Peromyscus*. *Mol. Biol. Evol.* **1**:38–56.
- BANISTER, K. E., and M. A. CLARKE. 1980. A revision of the large *Barbus* (Pisces, Cyprinidae) of Lake Malawi with a reconstruction of the history of the southern African Rift Valley lakes. *J. Nat. Hist.* **14**:483–543.
- CRAPON DE CAPRONA, M.-D. 1986. Are 'preferences' and 'tolerances' in cichlid mate choice important for speciation? *J. Fish Biol.* **29**:151–158.
- CROTHER, B. I. 1990. Is "some better than none" or do allele frequencies contain phylogenetically useful information? *Cladistics* **6**:277–281.
- DORIT, R. L. 1986. Molecular and morphological variation in Lake Victoria haplochromine cichlids (Perciformes: Cichlidae). Ph.D. diss., Harvard University, Cambridge, Mass.
- DOWLING, T. E., and W. R. HOEH. 1991. The extent of introgression outside the contact zone between *Notropis cornutus* and *Notropis chrysocephalus* (Teleostei: Cyprinidae). *Evolution* **45**:944–956.
- DOWLING, T. E., C. MORITZ, and J. D. PALMER. 1990. Nucleic acids. II: Restriction site analysis. Pp. 250–317 in D. M. HILLIS and C. MORITZ, eds. *Molecular systematics*. Sinauer, Sunderland, Mass.
- DOWLING, T. E., G. R. SMITH, and W. M. BROWN. 1989. Reproductive isolation and introgression between *Notropis cornutus* and *Notropis chrysocephalus* (Family Cyprinidae): comparison of morphology, allozymes, and mitochondrial DNA. *Evolution* **43**:620–634.

- DROUIN, J. 1980. Cloning of human mitochondrial DNA in *Escherichia coli*. *J. Mol. Biol.* **140**:15–34.
- ECCLES, D. E., and E. TREWAVAS. 1989. Malawian cichlid fishes: the classification of some haplochromine genera. Lake Fish Movies, Herten, Germany.
- ECHELLE, A. A., and I. KORNFIELD. 1984. Evolution of fish species flocks. University of Maine Press, Orono.
- FELSENSTEIN, J. 1988. PHYLIP: phylogenetic inference package, version 3.1. University of Washington, Seattle.
- FERRIS, S. D., R. D. SAGE, C.-M. HUANG, J. T. NIELSEN, U. RITTE, and A. C. WILSON. 1983. Flow of mitochondrial DNA across a species boundary. *Proc. Natl. Acad. Sci. USA* **80**:2290–2294.
- FRYER, G. 1959. Some aspects of evolution in Lake Nyasa. *Evolution* **13**:440–451.
- FRYER, G., and T. D. ILES. 1972. The cichlid fishes of the Great Lakes of Africa. Oliver Boyd, Edinburgh.
- GOLDSCHMIDT, T. 1989. An ecological and morphological field study on the haplochromine cichlid fishes (Pisces, Cichlidae) of Lake Victoria. Ph.D. diss., Rijksuniversiteit, Leiden.
- HOLZBERG, S. 1978. A field and laboratory study of the behaviour and ecology of *Pseudotropheus zebra* (Boulenger), an endemic cichlid of Lake Malawi (Pisces, Cichlidae). *Z. Zool. Syst. Evol. Forsch.* **16**:171–187.
- KORNFIELD, I. L. 1974. Evolutionary genetics of endemic African cichlids. Ph.D. diss., State University of New York, Stony Brook.
- . 1978. Evidence for rapid speciation in African cichlid fishes. *Experientia* **34**:335–336.
- . 1991. Genetics. Pp. 103–128 in M. H. A. KEENLEYSIDE, ed. *Cichlid fishes: behaviour, ecology and evolution*, Chapman & Hall, London.
- LANSMAN, R. A., R. O. SHADE, J. F. SHAPIRA, and J. C. AVISE. 1981. The use of restriction endonucleases to measure DNA sequence relatedness in natural populations. *J. Mol. Evol.* **17**:214–226.
- LIEM, K. F. 1981. A phyletic study of the Lake Tanganyika cichlid genera *Asprotilapia*, *Ectodus*, *Lestradea*, *Cunningtonia*, *Ophthalmochromis* and *Ophthalmotilapia*. *Bull. Mus. Comp. Zool.* **149**:191–214.
- MCELROY, D. M., and I. KORNFIELD. Novel jaw morphology in hybrids between *Pseudotropheus zebra* and *Labeotropheus fuelleborni* (Pisces: Cichlidae) from Lake Malawi, Africa. *Teleostei Copeia* (accepted).
- MCELROY, D. M., P. MORAN, E. BERMINGHAM, and I. KORNFIELD. 1992. REAP: an integrated environment for the manipulation and phylogenetic analysis of restriction data. *J. Hered.* **83**:157–158.
- MCKAYE, K. R., T. KOCHER, P. REINTHAL, R. HARRISON, and I. KORNFIELD. 1984. Genetic evidence for allopatric and sympatric differentiation among morphs of a Lake Malawi cichlid fish. *Evolution* **38**:215–219.
- MARSH, A. C., A. J. RIBBINK, and B. A. MARSH. 1981. Sibling species complexes in sympatric populations of *Petrotilapia* Trewavas (Cichlidae, Lake Malawi). *Zool. J. Linnean Soc.* **71**:253–264.
- MEYER, A., T. D. KOCHER, P. BASASIBWAKI, and A. C. WILSON. 1990. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* **347**:550–553.
- MORAN, P., I. KORNFIELD, and P. N. REINTHAL. Molecular systematics and radiation of the haplochromine cichlids (Teleostei: Perciformes) of Lake Malawi. *Copeia* (accepted).
- NEI, M. 1972. Genetic distance between populations. *Am. Nat.* **106**:283–292.
- . 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.
- NEI, M., and W.-H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* **76**:5269–5273.
- NEIGEL, J. E., and J. C. AVISE. 1986. Phylogenetic relationships of mitochondrial DNA under

- various demographic models of speciation. Pp. 515–534 in S. KARLIN and E. NEVO, eds. *Evolutionary processes and theory*. Academic Press, Orlando, Fla.
- OWEN, R. B., R. CROSSLEY, T. C. JOHNSON, D. TWEDDLE, S. DAVIDSON, I. KORNFIELD, D. H. ECCLES, and D. E. ENGSTROM. 1990. Major low levels of Lake Malawi and implications for speciation rates in cichlid fishes. *Proc. R. Soc. Lond. [Biol.]* **240**:519–533.
- PAMILO, P., and M. NEI. 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.* **5**:568–583.
- REGAN, C. T. 1922. The cichlid fishes of Lake Nyasa. *Proc. Zool. Soc. Lond.* **1921**:675–727.
- REINTHAL, P. N. 1987. Morphology, ecology, and behavior of a group of rock dwelling cichlid fishes (Cichlidae: Perciformes) from Lake Malawi. Ph.D. diss., Duke University, Durham, N.C.
- . 1990a. The feeding habits of a group of herbivorous rock-dwelling cichlid fishes (Cichlidae: Perciformes) from Lake Malawi, Africa. *Environ. Biol. Fishes* **27**:215–233.
- . 1990b. Morphological analysis of the neurocranium of a group of rock-dwelling cichlid fishes (Cichlidae: Perciformes) from Lake Malawi, Africa. *Zool. J. Linnean Soc.* **98**:123–139.
- RIBBINK, A. J., A. C. MARSH, C. C. RIBBINK, and B. J. SHARP. 1983. A preliminary survey of the cichlid fishes of rocky habitats in Lake Malawi. *S. Afr. J. Zool.* **18**:149–310.
- ROHLF, F. J. (1970). Adaptive hierarchical clustering schemes. *Syst. Zool.* **19**:58–82.
- . 1990. NTSYS: numerical taxonomy and multivariate analysis system, version 1.6. Exeter, Setauket, N.Y.
- SCHOLZ, C. A., and B. R. ROSENDAHL. 1988. Low lake stands in Lakes Malawi and Tanganyika, east Africa, delineated with multifold seismic data. *Science* **240**:1645–1648.
- SEYOUM, S., and I. KORNFIELD. 1992. Taxonomic notes on the *Oreochromis niloticus* subspecies complex (Pisces: Cichlidae), with a description of a new subspecies. *Can. J. Zool.* **70**:2161–2165.
- STURMBAUER, C., and A. MEYER. 1992. Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature* **358**:578–581.
- SWOFFORD, D. L. 1989. PAUP: phylogenetic analysis using parsimony, version 3.0. Illinois Natural History Survey, Champaign.
- SWOFFORD, D. L., and G. J. OLSEN. 1990. Phylogeny reconstruction. Pp. 411–500 in D. M. HILLIS and C. MORITZ, eds. *Molecular systematics*. Sinauer, Sunderland, Mass.
- TAJIMA, F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**:437–460.
- TREWAVAS, E. 1935. A synopsis of the cichlid fishes of Lake Nyasa. *Annu. Mag. Nat. Hist.* **10**:65–118.

WESLEY M. BROWN, reviewing editor

Received August 6, 1992; revision received April 19, 1993

Accepted April 19, 1993