Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 5107–5110, April 1999
Evolution

Phylogeny of a rapidly evolving clade: The cichlid fishes of Lake Malawi, East Africa

(adaptive radiation/sexual selection/speciation/amplified fragment length polymorphism/lineage sorting)

R. C. ALBERTSON, J. A. MARKERT, P. D. DANLEY, AND T. D. KOCHER†

Department of Zoology and Program in Genetics, University of New Hampshire, Durham, NH 03824

Communicated by John C. Avise, University of Georgia, Athens, GA, March 12, 1999 (received for review December 17, 1998)

ABSTRACT **Lake Malawi contains a flock of >500 species of cichlid fish that have evolved from a common ancestor within the last million years. The rapid diversification of this group has been attributed to morphological adaptation and to sexual selection, but the relative timing and importance of these mechanisms is not known. A phylogeny of the group would help identify the role each mechanism has played in the evolution of the flock. Previous attempts to reconstruct the relationships among these taxa using molecular methods have been frustrated by the persistence of ancestral polymorphisms within species. Here we describe results from a DNA fingerprinting technique that overcomes this problem by examining thousands of polymorphisms distributed across the genome. The resulting dendrogram averages the evolutionary history of thousands of genes and should accurately reflect the evolutionary history of these species. Our tree resolves relationships among closely related Lake Malawi cichlids and provides insights into the pattern of speciation in this group. We demonstrate that adaptive divergence in trophic morphology played an important role during the early history of the lake. Subsequent species diversity has arisen with little change in trophic morphology, which suggests that other forces are responsible for the continued speciation of these fishes.**

Lake Malawi and its fish fauna have attracted the attention of ecological, behavioral, and evolutionary biologists for over a century (1). Lake Malawi cichlids exhibit spectacular diversity in trophic morphology, including specialist algal scrapers, planktivores, insectivores, piscivores, paedophages, snail crushers, and fin biters. As rich as they are in trophic diversity, these fishes are even more striking in their array of color patterns (2). Their complex mating behavior, polyandrous mating systems (3), and tendency to breed in leks also are intriguing to behaviorists.

Both adaptive and nonadaptive processes likely have played a role in the explosive radiation of East African cichlids, but the importance of each mechanism has been the focus of considerable debate (4). On one hand, it has been suggested that morphological adaptation, particularly of the feeding apparatus, is a "key innovation" for rapid evolutionary change (5). An opposing hypothesis suggests that adaptive morphological change is not the primary cause of speciation but occurs after the establishment of genetic isolation by other means (6, 7). These authors argue that sexual selection is the force that creates reproductive isolation, allowing subsequent morphological differentiation. The competing hypotheses could be evaluated if a robust phylogeny of these taxa was available. If morphological adaptation is critical to speciation, then we expect that even closely related species will show significant divergence in trophic morphology. If sexual selection is responsible for speciation, then we expect that sister taxa will frequently differ in color pattern but not morphology.

Most attempts to determine the relationships among cichlid species have used morphological characters, which may be prone to convergence (8). Molecular sequences normally provide the independent estimate of phylogeny needed to infer evolutionary mechanisms. The Lake Malawi cichlids, however, are speciating faster than alleles can become fixed within a species (9, 10). The coalescence of mtDNA haplotypes found within populations predates the origin of many species (11). In such situations, phylogenies based on a single gene reflect only the history of that marker and may not accurately reflect the true relationships among species (12). Studies using several unlinked microsatellite DNA loci also have failed to resolve relationships among even small sets for taxa (13). Only a method that assays variation at a large number of loci is likely to overcome the effects of incomplete sorting of alleles between speciation events (9).

Here, we use a DNA fingerprinting method, amplified fragment length polymorphisms (AFLP) (14), to study variation at thousands of sites distributed across the genome. This technique is simple and robust and allows the typing of thousands of independently segregating loci. Unlike traditional restriction fragment length polymorphisms, characters generated by AFLP are largely independent. The loss of a restriction site removes the fragment from the profile rather than changing its size. Our results demonstrate that the AFLP technique may be used to resolve phylogenetic structure in rapidly evolving systems in which persistence of ancestral polymorphisms makes phylogenetic reconstruction difficult by other means.

MATERIALS AND METHODS

Specimens. We focused our study on the rock-dwelling cichlids, known locally as "mbuna." This clade consists of 13 genera and ≈300 species (2). We sampled two or more individuals from each of 14 populations, representing a total of nine species in four genera. Each of these genera is characterized by a unique jaw morphology. *Metriaclima* has a moderately sloped head and a terminal, isognathus mouth that it uses to comb food from attached algae while perpendicular to the rock surface. *Labeotropheus* is distinguished from other mbuna by a protuberant fleshy snout and an inferior-subterminal mouth that it uses to scrape algae while oriented parallel to the substrate. Members of the *Pseudotropheus tropheops* species complex have a steeply descending facial profile and a slightly subterminal mouth. *Melanochromis* originally was defined by jaw characters, although it is now recognized by the presence of strong horizontal stripes in the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviation: AFLP, amplified fragment length polymorphism.

†To whom reprint requests should be addressed at: Department of Zoology and Program in Genetics, University of New Hampshire, 216 Rudman Hall, 46 College Road, Durham, NH 03824. e-mail: Tom.Kocher@unh.edu.

color pattern (15). Within genera, species differ primarily in color pattern. For example, male *Metriaclima zebra* have a series of 6–8 vertical black bars on a brilliant blue background whereas its sympatric congener *Metriaclima benetos* lacks the black bars. A third species, *Metriaclima sandaracinos*, differs by having a red dorsal fin (16).

We also included in our study a nested set of outgroups to provide a context for interpreting the mbuna results. *Copadichromis eucinostomus,* another haplochromine cichlid from Lake Malawi, is a member of the sand-dwelling sister clade to the mbuna. *Tropheus moorii*, from Lake Tanganyika, belongs to the tribe (*Tropheini*) thought to be a sister group to the Lake Malawi flock (17). *Neolamprologus brichardi*, a lamprologine cichlid from Lake Tanganyika, is more distantly related to the Lake Malawi flock. Finally, the tilapiine cichlid *Oreochromis niloticus* diverged from these other taxa ≈10 million years ago (18). All animals used were caught from the wild, with the exception of *O. niloticus*, representatives of which were obtained from a commercial tilapia farm. DNA was extracted from fin clips stored in ethanol.

Molecular Methods. We followed the AFLP procedure as originally described (14). Genomic DNAs were cut with *Eco*RI and *Mse*I. Double-stranded adapters specific for each site were ligated to each end, altering the recognition sequence and preventing a second restriction. Preselective amplification was performed with one selective base on each primer (*Eco*RI-A and *Mse*I-C), reducing the number of displayed fragments ≈16-fold. Eleven different selective amplifications were performed by using primers with an additional two-base extension (E-AGG, M-CTT, M-CTG; E-ACT, M-CTA, M-CAG, M-CAT; E-ACA, M-CAA, M-CAG; E-AGC, M-CAG, M-CAT; E-ACC, M-CAA, M-CAC). Fragments were visualized by labeling the selective *Eco*RI primer with a fluorescent dye. Each selective primer combination amplified ≈150 fragments in the range of 100–500 bp from each fish. These fragments were resolved on a 6% acrylamide gel in an Applied Biosystems 373A automated DNA sequencer. Fragment data were collected by using Applied Biosystems GENESCAN 2.0.2 software.

Distance Analysis. The list of fragment sizes for each individual was transferred into a spreadsheet and was manually aligned among individuals. Because sizing of fragments on the automated sequencer has a standard deviation of 0.15 bp (19), fragments were inferred to be homologous if they differed by no more than 0.5 base pairs. The data then were coded to indicate the presence/absence of each fragment in each individual. Dendrograms were constructed from the mean character distance by using the neighbor-joining algorithm of PAUP* 4.0 (20).

RESULTS AND DISCUSSION

Our data set consists of 2,247 characters derived from 11 selective primer pair combinations. This corresponds to a marker about every 0.5 centimorgans over the genome (21). Of these, 1,205 were polymorphic among the taxa studied here. The dendrogram shown in Fig. 1 reflects phylogenetic structure for both recent and ancient divergences.

Consistent with recent evidence that mbuna populations are highly structured (22–24), replicate individuals from nearly every population cluster together. The bootstrap values that unite populations are among the most highly supported (unit-

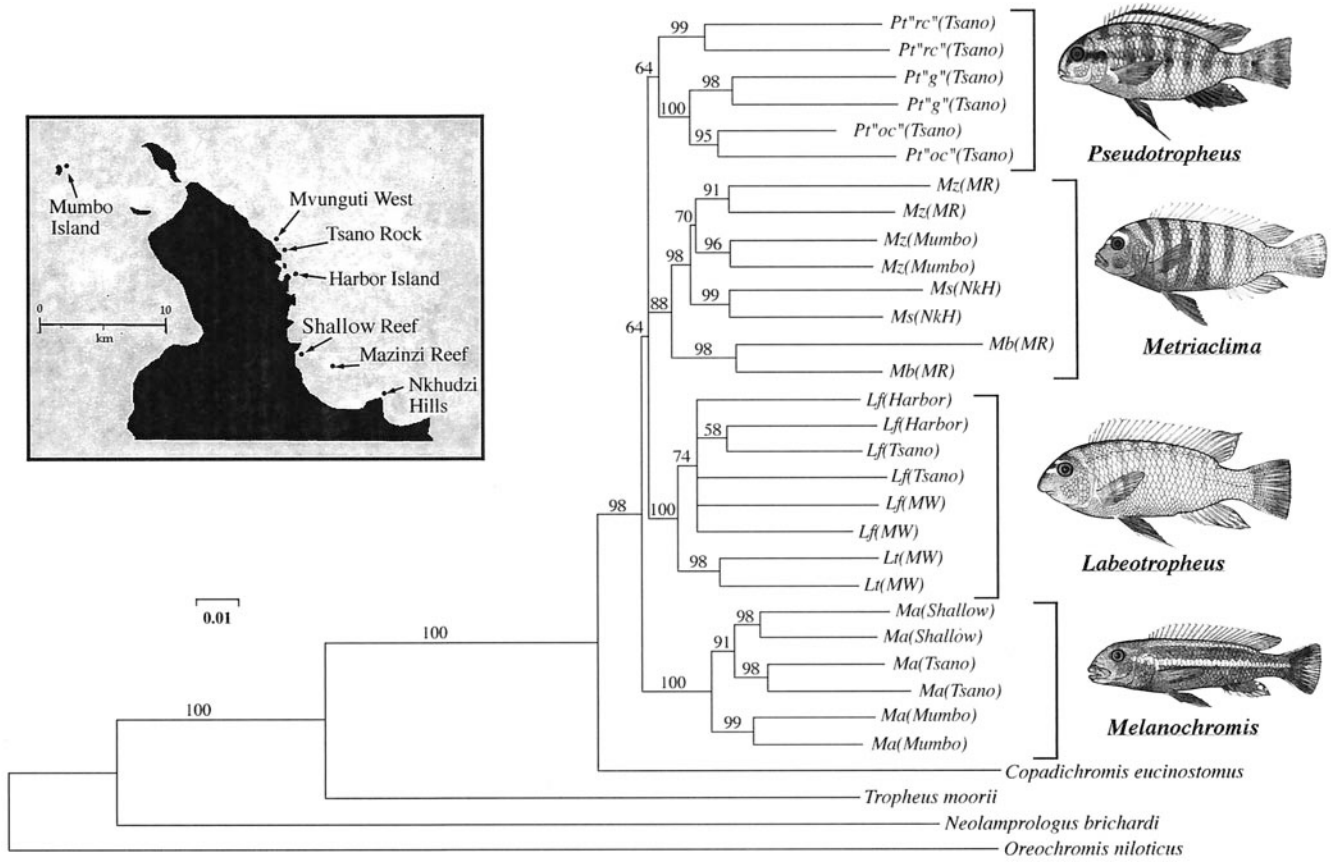

FIG. 1. Dendrogram describing the relationships among Lake Malawi cichlid fishes. Shown is a neighbor-joining tree based on the similarity of AFLP banding profiles. Species abbreviations: Pt, *P. tropheops*; "rc," red cheek; "g," *gracilior*; "oc," orange chest; Mz, *M. zebra*; Ms, *M. sandaracinos*; Mb, *M. benetos*; Lf, *L. fuelleborni*; Lt, *L. trewavasae*; Ma, *M. auratus*. Localities (in parentheses) refer to the inset map of sites near the Nankumba Peninsula in southern Lake Malawi. Numbers indicate the proportion of 1,000 bootstrap samples in which a particular clade was found. Scale bar indicates 1% character difference.

ed in an average of 97% of the bootstrap resamplings). The only exception is the clustering of individual *Labeotropheus fuelleborni* from Tsano Rock, Harbor Island, and Muvunguti West. These localities are within 5 km of one another, and analysis of microsatellite DNA markers suggests that gene flow is high among these particular populations (23). Thus, these three localities may be considered effectively one population.

Intrageneric phylogenies provide a framework for studies of speciation mechanisms. Our data resolve such relationships in Lake Malawi cichlids for the first time. Our tree rejects the hypothesis that the endemic unbarred species *M. benetos* arose from the black-barred species *M. zebra* at Mazinzi Reef. Rather, *M. benetos* is a relatively distinct lineage of *Metriaclima* that must have arisen long before this particular locality was submerged 200 years ago (25). We also studied three sympatric species of *P. tropheops* that show very little anatomical differentiation and differ primarily in color pattern. Our data show strong support (100% bootstrap) for the clustering of *P. tropheops* "gracilior" and *P. tropheops* "orange chest" to the exclusion of *P. tropheops* "red cheek."

If trophic adaptation occurred after speciation events, we would expect to find that closely related species differed in jaw morphology. In contrast, our tree shows the integrity of the four morphologically defined genera. The four *Metriaclima* species are united in 88% of the bootstrap resamplings. The two *Labeotropheus* species are found in 100% of the bootstrap resamples, and the three *P. tropheops* species are clustered in 64%. There is no evidence for convergence of feeding morphology among distantly related species; closely related species differ primarily in color pattern, not jaw morphology.

Moreover, our tree suggests that hybridization has not been rampant during the evolution of the flock (26). Each morphologically defined genus forms a discrete clade. If hybridization were a frequent occurrence, such phylogenetic structure would not have been observed. Although mbuna will hybridize under artificial conditions (27), only a single instance of natural hybridization has ever been observed (28). We cannot rule out a role for hybridization during the early radiation of the flock, but the integrity of genera observed here suggests that hybridization is now a rare event.

Our tree also recovers the predicted relationships among the outgroup taxa. *Copadichromis eucinostomus* occupies its expected position as sister to the mbuna clade. *Tropheus moorii* appears sister to the Lake Malawi flock. The overall topology is consistent with previous phylogenetic work on this group (17, 18, 29). The high bootstrap values at each node indicate that this tree is robust.

The branch lengths in the tree reflect the distribution of genetic variation within and among species. Consistent with previous evidence for incomplete lineage sorting (9), the vast majority of genetic variation resides within populations. Only a small proportion of the variation is distributed among species and genera, suggesting that speciation has occurred without significant population bottlenecks. Consistent with this idea, surveys of mitochondrial DNA haplotypes (30) and microsatellite loci (13) have found high allelic diversity within most populations.

Rapid radiations are expected to result in "starburst" phylogenies lacking internal structure, but such bushy trees also may be produced when the data lack power to resolve relationships among distantly related taxa (the "saturation" effect). The incomplete resolution of relationships among these four *mbuna* genera is not caused by a weakness of the technique because our data resolve both earlier and later divergences. We find strong support for recent nodes in our tree, describing the relationships among populations with divergence times of a few hundred years (25); our analysis also resolves the relationships of outgroups that diverged >10 million years ago.

To determine how the resolution of our tree might improve with the addition of more characters, we plotted the number of resolved nodes against the number of primer pairs used (Fig. 2*a*). The 26 nodes in Fig. 1 are resolved with the first seven primer combinations (≈700 informative characters), but mean bootstrap values continue to increase as more characters are scored (Fig. 2*b*). The full potential of the AFLP technique has not yet been realized, and we expect that bootstrap values will continue to improve as we increase the number of characters scored.

The AFLP method may find broad applicability in evolutionary studies of rapidly speciating groups. It has significant advantages over other nuclear markers, in terms of cost and speed. The binary nature of the allelic data allow the estimation of genetic distances with far fewer samples than are needed for highly polymorphic microsatellite markers. The AFLP method should find application wherever single-marker studies conflict because of incomplete lineage sorting. Examples that come to mind include the relationships among hominoid primates and the radiation of Darwin's finches.

The search for "key innovations" has dominated studies of the East African cichlids for many years. Our data argue for a more pluralistic explanation for the origins of this taxonomic diversity. Morphological adaptation clearly played a role in the radiation of the *mbuna*, eventually resulting in 13 protogenera differing in jaw morphology. But only minor morphological differences are observed within genera (16). Subsequent diversification into >300 species is associated with variation in other characters, most notably the color pattern of adult males. Recent evidence suggests that sexual selection plays an important role in maintaining the diversity of East African cichlids (31). Our results imply that sexual selection is also the source of that diversity. We also contend that neither hybridization nor population bottlenecks played a significant role in

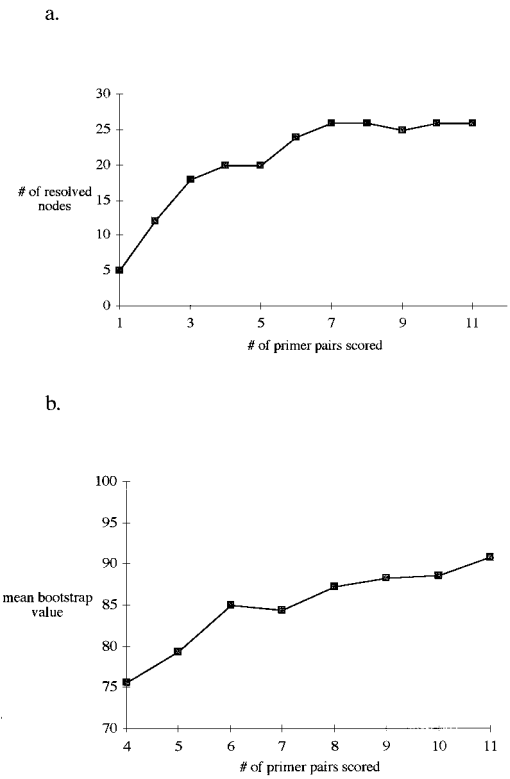

FIG. 2. Improved resolution of nodes with increasing number of characters. (*a*) Number of resolved nodes versus the number of selective AFLP primer pairs used. (*b*) Mean bootstrap values versus number of primer pairs scored for nodes resolved in 50% or more of the bootstrap samples.

the radiation of the *mbuna*. The ability to reconstruct the phylogenetic relationships among closely related populations and species will provide a historical context within which we may understand the patterns and mechanisms of speciation in these remarkable fishes.

We thank M. Arnegard, J. Stauffer, A. Ambali, the University of Malawi, and the Malawi Government for assistance with collection of specimens and J. Bolker, K. Carleton, S. Cohen, J. Curole, N. Garnhart, and M. Scott for discussion and critical reading of the manuscript. This work was supported by grants from the National Geographic Society and the National Science Foundation.

1. Fryer, G. & Iles, T. D. (1972) *The Cichlid Fishes of the Great Lakes of Africa: Their Biology and Evolution* (Oliver and Boyd, Edinburgh).
2. Konings, A. (1995) *Malawi Cichlids in Their Natural Habitat* (Cichlid Press, Lauenau, Germany).
3. Kellogg, K. A., Markert, J. A., Stauffer, J. R., Jr. & Kocher, T. D. (1995) *Proc. R. Soc. London Ser. B* **260,** 79–84.
4. Greenwood, P. H. (1990) in *Cichlid Fishes: Behavior, Ecology and Evolution*, ed. Keenleyside, M. H. A. (Chapman & Hall, New York), pp. 86–102.
5. Liem, K. F. (1974) *Syst. Zool.* **22,** 425–441.
6. Dominey, W. J. (1984) in *Evolution of Fish Species Flocks*, eds. Echelle, A. A. & Kornfield, I. (Univ. of Maine Press, Orono, ME), pp. 231–249.
7. Sage, R. D., Loiselle, P. V., Basasibwaki, P. & Wilson, A. C. (1984) in *Evolution of Fish Species Flocks*, eds. Echelle, A. A. & Kornfield, I. (Univ. of Maine Press, Orono, ME), pp. 185–197.
8. Kocher, T. D., Conroy, J. A., McKaye, K. R. & Stauffer, J. R., Jr. (1993) *Mol. Phylogenet. Evol.* **2,** 158–165.
9. Moran, P. & Kornfield, I. (1993) *Mol. Biol. Evol.* **10,** 1015–1029.
10. Klein, J., Sato, A., Nagl, S. & O'hUigin, C. (1998) *Annu. Rev. Ecol. Syst.* **29,** 1–21.
11. Parker, A. & Kornfield, I. (1997) *J. Mol. Evol.* **45,** 70–83.
12. Avise, J. C. (1994) *Molecular Markers, Natural History and Evolution* (Chapman & Hall, New York).
13. Kornfield, I. & Parker, A. (1997) in *Molecular Systematics of Fishes*, eds. Kocher, T. D. & Stepien, C. A. (Academic, San Diego), pp. 25–37.
14. Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., *et al.* (1995) *Nucleic Acids Res.* **23,** 4407–4414.
15. Trewavas, E. (1983) *Rev. Fr. Aquariol.* **10,** 97–106.
16. Stauffer, J. R., Jr., Bowers, N. J., Kellogg, K. A. & McKaye, K. R. (1997) *Proc. Acad. Nat. Sci. Philadelphia* **148,** 189–230.
17. Meyer, A. (1993) *Trends Ecol. Evol.* **8,** 279–284.
18. Kocher, T. D., Conroy, J. A., McKaye, K. R., Stauffer, J. R., Jr. & Lockwood, S. F. (1995) *Mol. Phylogenet. Evol.* **4,** 420–432.
19. Lazaruk, K., Walsh, P. S., Oaks, F., Gilbert, D., Rosenblum, B. B., Menchen, S., Scheibler, D., Wenz, H. M., Holt, C. & Wallin, J. (1998) *Electrophoresis* **19,** 86–93.
20. Swofford, D. L. (1998) PAUP* 4.0 *Beta Version* (Sinauer, Sunderland, MA).
21. Kocher, T. D., Lee, W.-J., Sobolewska, H., Penman, D. & McAndrew, B. (1998) *Genetics* **148,** 1225–1232.
22. van Oppen, M. J. H., Deutsch, J. C., Turner, G. F., Rico, C., Ibrahim, K. M., Robinson, R. L. & Hewitt, G. M. (1997) *Proc. R. Soc. Lond. Ser. B* **264,** 1803–1812.
23. Arnegard, M. E., Markert, J. A., Danley, P. D., Stauffer, J. R., Jr. & Kocher, T. D. (1999) *Proc. R. Soc. Lond. Ser. B* **266,** 1–12.
24. Markert, J. A., Arnegard, M. E., Danley, P. D. & Kocher, T. D. (1999) *Mol. Ecol.*, in press.
25. Owen, R. B., Crossley, R., Johnson, T. C., Tweddle, D., Davidson, S., Kornfield, I., Eccles, D. H. & Engstrom, D. E. (1990) *Proc. R. Soc. Lond. Ser. B* **240,** 519–533.
26. Crapon de Caprona, M.-D. (1986) *J. Fish Biol.* **29,** 151–158.
27. McElroy, D. M. & Kornfield, I. (1993) *Copeia* **1993,** 933–945.
28. Stauffer, J. R., Jr., Bowers, N. J., Kocher, T. D. & McKaye, K. R. (1996) *Copeia* **1996,** 203–208.
29. Meyer, A., Kocher, T. D., Basasibwaki, P. & Wilson, A. C. (1990) *Nature (London)* **347,** 550–553.
30. Moran, P. & Kornfield, I. (1995) *Mol. Biol. Evol.* **12,** 1085–1093.
31. Seehausen, O., van Alphen, J. J. M. & Witte, F. (1997) *Science* **277,** 1808–1811.