Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes

Christian Sturmbauer & Axel Meyer

Department of Ecology and Evolution, State University of New York, Stony Brook, New York 11794-5245, USA

SINCE their discovery at the turn of the century[1], the species assemblages of cichlid fishes in the East African Lakes Victoria, Malawi and Tanganyika have fascinated evolutionary biologists. Many models have attempted to account for the 'explosive' evolution of several hundred species within these lakes[2-7]. Here we report a case of surprisingly large genetic divergence among populations of the endemic *Tropheus* lineage of Lake Tanganyika. This lineage of six species contains twice as much genetic variation as the entire morphologically highly diverse cichlid assemblage of Lake Malawi and six times more variation than the Lake Victoria species flock. Although it is highly variable in coloration, this group of species and its closest relatives have not undergone appreciable morphological change. The observed geographic pat-
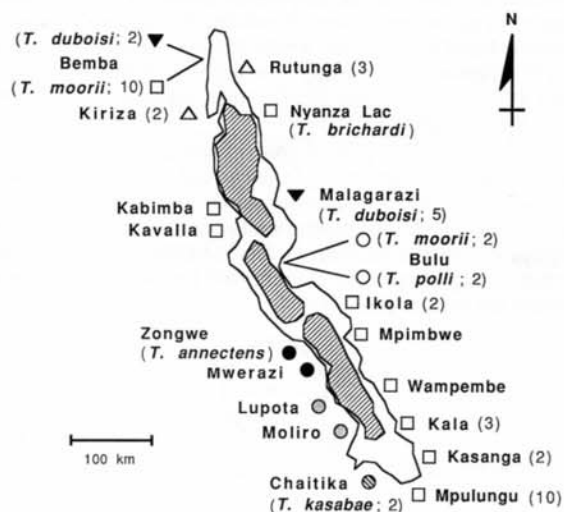

FIG. 1 Localities of 21 *Tropheus* populations from Lake Tanganyika included in this study. Fifty-four individuals, assigned to six currently recognized species by meristic characters (for example, by the number of anal spines) and coloration, were sequenced[10]: *Tropheus moori* Boulenger 1898, *T. annectens* Boulenger 1898, *T. duboisi* Marlier 1959, *T. brichardi* Nelissen and Thys 1976, *T. kasabae* Nelissen 1977 and *T. polli* Axelrod 1978. The populations are identified by the name of the nearest village[13]. The species names are given according to the current taxonomic assignments. Numbers of individuals analysed from each locality are given in parentheses if larger than one. All populations in which the species name was omitted are presently classified as *T. moorii*; this name is only included for those localities in which *T. moorii* is sympatric with other *Tropheus* species. The symbols indicate genetically distinct lineages based on the phylogenetic analyses of mitochondrial DNA sequence data (Fig. 4). The molecular phylogeny is in partial conflict with the present taxonomy. We suggest a revision of the taxonomy of *Tropheus*, which should include molecular data rather than rely solely on taxonomic characters like coloration or the number of spines in the anal fin that are currently used. Stippled areas indicate the locations of the three separate basins during a period of low lake level[16,17]. In all populations from the geographically widespread lineage for which sequences were determined (squares in Figs 1 and 4) only two substitutions were observed (Fig. 4a) in cytochrome *b* and up to 4% (corrected sequence divergence, see Fig. 3) in the control region. In the central part of the lake two populations from opposite shorelines (Kabimba and Ikola) have a smaller sequence divergence in the control region (0.9%) than to other populations from the same coastline (for example, Kabimba−Bemba=2.1%; Ikola−Nyanza=2.4%).

tern of genetic variation suggests that major lake level fluctuations affected the distribution and speciation of this lineage of cichlid fishes.

Speciation and morphological evolution are two independent processes that usually occur synchronously but may proceed independently[8]. The morphology of 'living fossils', like horseshoe crabs, has remained essentially unchanged for millions of years, although these organisms exhibit normal levels of molecular evolution[9]. In contrast, hundreds of species of cichlid fishes (species flocks) from the African Great Lakes, particularly Lakes Victoria and Malawi, originated extremely rapidly through adaptive radiations and are genetically exceptionally closely related[7], yet these fishes have undergone major morphological diversifications. The oldest of these lakes, Lake Tanganyika, harbours 171 species assigned to 49 endemic genera which belong to at least seven old, evolutionarily distinct lineages[10,11]. Because of their long independent evolutionary history, these cichlids are morphologically, ecologically and behaviourally the most diverse species flock of the African lakes[4,12].

In most Tanganyikan cichlid lineages, many geographically distinct populations have been described, distinguishable by only minor if any morphological variation but by pronounced differences in coloration[10]. The best known examples of this phenomenon are the six described species of the *Tropheus* lineage. Of these, more than 50 distinctly coloured 'races' are currently known, some of which have overlapping distributions[13]. *Tropheus* are strictly confined to rocky habitats for foraging and mating and have limited capacity for dispersal across open water[13,14].

Mitochondrial DNA of 54 individuals from 21 populations of all six described species of *Tropheus* was sequenced to investigate the distribution of genetic variation within and among populations and species in relation to taxonomy, coloration and geographic distribution (Fig. 1). Up to 842 base pairs (bp) from the control region and the cytochrome *b* gene were sequenced for each individual (Fig. 2).
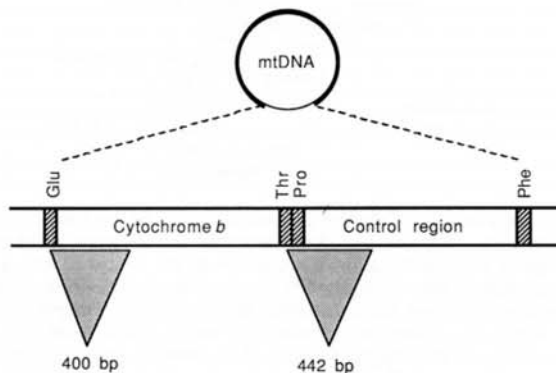
FIG. 2 Portion of the mitochondrial genome showing the sequenced gene segments. A 442-base-pair (bp) segment of the major non-coding region (control region), including a part of the threonine transfer RNA gene (12 bp) and the proline tRNA gene (71 bp), was amplified by the polymerase chain reaction (PCR) and directly sequenced[7,26] in all 54 individuals. In addition, a 400-bp segment of the cytochrome *b* gene was sequenced for all six species (15 specimens from 13 populations). DNA was extracted from white muscle tissue of frozen or ethanol-preserved specimens. PCR amplifications were done in 25-μl volumes of TRIS buffer (67 mM, pH 8.8) containing 2 mM MgCl$_2$, 1 mM of each dNTP, 1 μM of each primer and *Taq* polymerase (1 unit; Cetus). The primers used for the amplification of part of the cytochrome *b* gene were L14724 and H15148, and those for the tRNA and the control region were L15926 and H16498 (details of the protocol and primer sequences have been reported in refs 7, 26). Sequences have been deposited in Genbank under accession numbers Z12047–Z12100 (*Tropheus*, control region); Z12030–Z12044 (*Tropheus*, cytochrome *b*), Z12045 (*Simochromis babaulti*, cytochrome *b*), Z12046 (*Oreochromis tanganicae*, cytochrome *b*). Swedish Museum of Natural History, NRM 12815. The aligned sequences have also been deposited as supplementary information at the *Nature* office.

The amount of sequence divergence in the mitochondrial control region within the *Tropheus* lineage was compared to that within the cichlid species flocks of Lakes Malawi and Victoria. This comparison revealed a surprisingly high degree of genetic variation in *Tropheus* (Fig. 3). With a maximum corrected sequence divergence[15] (Fig. 3) of 14.5%, the *Tropheus* lineage contains almost twice as much variation as the whole Malawi species flock of about 500 species (8.2%) (ref. 7, and P. Reinthal and A.M. unpublished results) and more than six times as much variation as the Victoria flock of about 300 species (2.3%)[7]. This estimate is also supported by the variation in cytochrome *b*, for which there is up to 5% variation within the *Tropheus* lineage compared with less than 2% in the whole Malawi flock[7]. The species flock of Lake Malawi is estimated to be around 700,000 years old and that of Lake Victoria 200,000
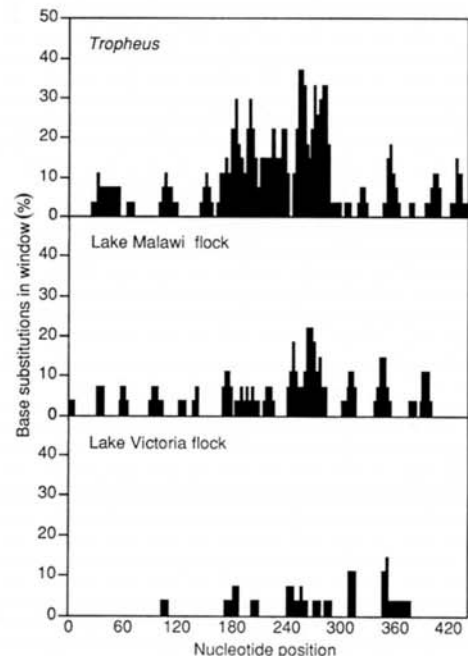
FIG. 3 Sliding window analysis of 442 base pairs (bp) of the mitochondrial control region (window: 9 bp, 3 bp overlap) of *Tropheus* and representative species of the endemic species flocks of lakes Malawi (500–1,000 species) and Victoria (about 300 species)[12]. Twenty-one Tropheus (*n* =54) populations from Lake Tanganyika were compared with 24 Malawian species from 12 genera (*n* =45) and 14 species of 9 genera (*n* =32) from Lake Victoria[7]. Note the extraordinarily large sequence divergence in *Tropheus*. The genetic variation is given in per cent of the 27 possible (window width, 9 bp) base substitutions from a consensus sequence. Transversion mutations are less frequently observed than transitions among closely related species and therefore trace phylogenetic events more reliably[15]. Transversions will cover up previous transitions and estimates of sequence divergence must be corrected for those multiple substitutions. The sequence divergences of the control region were corrected for multiple substitutions according to a formula[15] based on the observed ninefold higher frequency of transitions than transversions. The frequency of transversions was determined by a linear regression, plotting transitions versus transversions of 53 pairwise comparisons of species or populations of up to 8 per cent uncorrected sequence divergence (regression: $y =0.11*x +0.19$, $r^2 =0.66$, $1/0.11 =9.1$). The amount of genetic variation within populations was surveyed in two populations. At Bemba, eight out of ten fish were identical; one individual differed by one transition, another by two transitions. At Mpulungu, eight out of ten fish had sequence differences of between 1 and 17 substitutions (3.8%). This population contains several distinct genotypes that are closely related to those found in neighbouring populations at Kasanga, Kala and Wampembe. These findings agree with data from sediment cores from that area, showing the drainage of Mpulungu Bay by two repeated drops in lake level by 400 m 30,000–20,000 years ago[27], probably causing admixture of those populations.
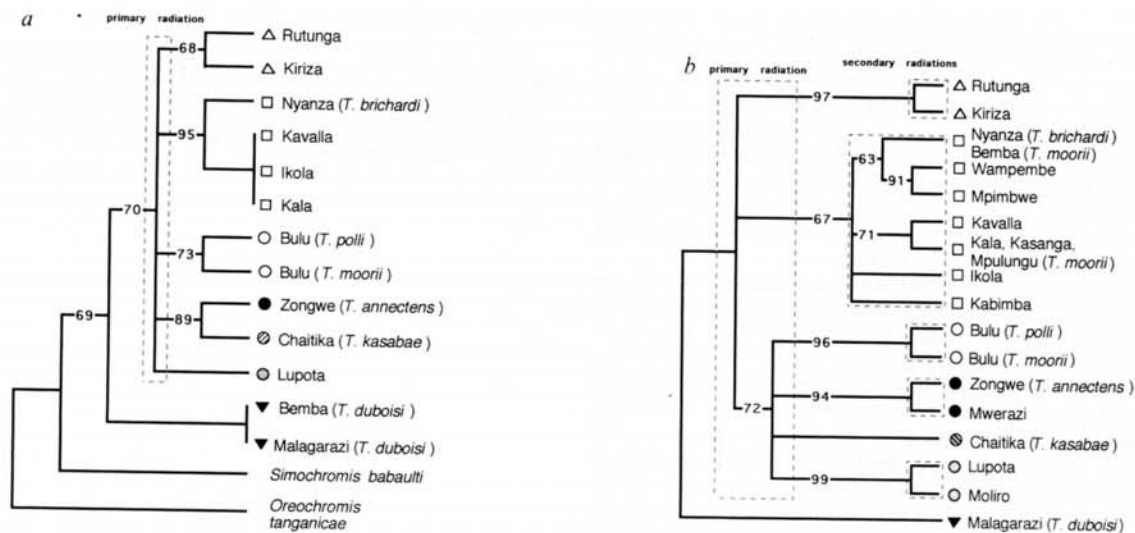
FIG. 4 Molecular phylogeny of the genus *Tropheus*. *a*, Bootstrap consensus tree (100 replications, 50% majority rule) based on cytochrome *b* sequences. *Oreochromis tanganicae*, a member of the tilapiine cichlids, a distantly related species to *Tropheus* (and most other endemic Tanganyikan genera (refs 11 and 28; also our unpublished observation), was used as outgroup. Transitions and transversions were equally weighted because of the low frequency of transversions. Thirteen populations of six species from all coastal regions were included in this analysis, one shortest tree was found (heuristic search, ACCTRAN option; tree length, 93; consistency index, 0.75; consistency index excluding uninformative characters, 0.56). *Tropheus duboisi* is the sister group of all other *Tropheus* species; monophyly of the *Tropheus* lineage seems supported by the inclusion of *Simochromis babaulti*, a member of the tribe Tropheini[10]. *b*, Bootstrap consensus tree (50% majority rule) based on the control region. Transversions were weighted nine times over transitions, according to their frequency among closely related species (Fig. 3). Three insertions were observed; each was weighted as a single transversion. This analysis included 16 representative populations (heuristic search, 100 replications, ACCTRAN option; weighted tree length, 386; unweighted tree length, 171; consistency index, 0.61; consistency index excluding uninformative characters, 0.49). To keep the number of taxa manageable for the bootstrap analysis[29], each population was represented by only one individual. Parsimony analysis[29] of all individvals (N=36; individuals with identical sequences were excluded) resulted in 24 equally short trees (heuristic search, ACCTRAN option; weighted tree length, 444; unweighted tree length, 197; consistency index, 0.58; consistency index excluding uninformative sites, 0.5) which agree with the topology shown and varied only slightly in the arrangements of terminal branches. Parsimony analysis and bootstrap procedures were conducted with PAUP (Version 3.0rβ[29]). The closely related populations from Nyanza and Bemba and the populations from Kala, Kasanaga and Mpulungu were each represented by only one individual. The primary radiation and the succeeding more recent secondary radiations of each lineage are indicated by boxes. A neighbour-joining analysis[30] (54 individuals) with weighted distance matrix (weighing transversions and gaps nine times over transitions) resulted in a tree in agreement with the branching order shown here.

years[7]. The geological ages of the lakes themselves are gauged to be somewhat greater[4,16,17]. On the basis of these estimates and the assumption of comparable rates of molecular divergence among the members of each of the African species flocks, the *Tropheus* lineage might therefore be about 1.25 million years old. Although the absolute ages of these species assemblages remain approximations, the relative divergence demonstrates the ancienty of the *Tropheus* lineage compared to other morphologically highly diverse species assemblages.

Tropheus is morphologically distinct from the other genera (for example, *Simochromis babaulti*; Fig. 4*a*) assigned to the same tribe, Tropheini[10]. Despite the large amount of sequence divergence and inferred old age, all species of *Tropheus* are morphologically largely identical even where they live sympatrically. Genetically, *Tropheus duboisi* is the most divergent species of the genus, but morphologically it differs only slightly from all other *Tropheus* by a narrower mouth and slight dentitional differences[18].

The extreme levels of morphological diversity among the main lineages of the Tanganyikan species flock, which lack intermediate morphotypes, led Greenwood[12] to hypothesize that these lineages were formed rapidly after the colonization of the lake by several ancestors filling available niches, but then remained morphologically unchanged. Lacking crucial fossil data, this scenario had remained untested but seems now to be supported by the molecular data for *Tropheus*. Among various possible explanations for stasis in morphology[8,19–21], the action of stabilizing selection, as a consequence of species as part of 'tightly packed' communities[21], might best explain the situation

in the *Tropheus* lineage. Future molecular investigation of other Tanganyikan lineages might determine whether the morphological stasis and the mode of evolution inferred from the *Tropheus* lineage is a general attribute of the species flock of Lake Tanganyika.

The evolutionary relationships among *Tropheus* were reconstructed based on DNA sequences. The phylogenetic information obtained from the more slowly evolving cytochrome *b* gene indicated the relationship of *Tropheus* to morphologically dissimilar genera (such as *Simochromis*; Fig. 4*a*). The rapidly evolving control region resolved more recent phylogenetic events within the genus (Fig. 4*b*). *Tropheus duboisi* appears to be the closest living relative of all other *Tropheus* species. After this first split of the *Tropheus* line, two rapid creations of lineages (radiations) occurred. In a primary radiation (Fig. 4*a*, *b*) probably all rocky shores of the lake were colonized. Six lineages with similar genetic divergence (Fig. 4*b*) originated during the primary radiation; they are separated by an average corrected sequence divergence in the control region of 8.2% (standard deviation 1.4%; 314 pairwise comparisons). Each of these lineages has undergone a 'secondary radiation' (Fig. 4*b*).

These secondary radiations might have been triggered by a large fluctuation in lake level. The similar sequence divergences among members of each of these radiations (average in the control region: 2.7%, standard deviation 1.0%, 195 pairwise comparisons between populations from each lineage) of all the secondary radiations suggests that they might have arisen at a similar time in the geological history of the lake. Seismic data show that about 200,000 (ref. 16) to 75,000 (ref. 17) (C. A.

Scholz, personal communication) years ago the level of Lake Tanganyika dropped 600 m below its present level, probably splitting the lake into three sublakes for several tens of thousands of years (Fig. 1). This event must have had trenchant effects on the biogeographic distribution of many species and probably led to widespread extinctions and fusions of formerly isolated populations. Because *Tropheus* are confined to rock habitats[14,18] many populations might have disappeared in shallow areas of the lake. Populations at rocky steeply sloping shores probably survived the drop in lake level intact.

The geographic distribution of five of the observed six primary lineages is limited, but members of one lineage (square symbols in Figs 1 and 4) are widespread throughout the lake. They have colonized newly available habitats, some of which are geographically separate and were probably previously occupied by other *Tropheus* lineages (the southeastern part of the lake for example). The distribution of genetically closely related populations on both sides of the lake (Fig. 1) suggests that colonizations not only took place along the same side of the lake, but these fish seemed to have crossed from one shore to the other. The northern and southern shorelines of the sublakes during the low lake level might have permitted the crossing of the lake. The molecular divergence data suggest that abiotic factors, such as major lake level fluctuations, might be partly responsible for speciation in *Tropheus*.

The DNA data show that genetically closely related populations can differ considerably in coloration. For example, the two sympatric, genetically closely related species (3.1% sequence divergence) at Bulu are remarkably differently coloured. Behavioural observations demonstrated the importance of coloration during conspecific interactions[14,22]. Sexually and socially selected traits might undergo more rapid divergence than might traits that are under natural selection[3,6,23,24]. Owen *et al.*[25] reported on extremely young (estimated to be less than 1,000 years old), colorationally distinct species from the Lake Malawi cichlid species flock. *Tropheus* may be a model of how rapidly differences in sexually selected traits such as coloration might arise as reproductive barriers (without concordant morphological diversification) during periods of isolation, possibly aiding the 'explosive' pace of speciation in East African cichlid species flocks. □

1. Boulenger, C. L. *Trans. Zool. Soc. Lond.* **15**, 1–30 (1898).
2. Avise, J. C. *Nature* **347**, 512–513 (1990).
3. Kosswig, C. *Nature* **159**, 604–605 (1947).
4. Fryer, G. & Iles, T. D. *The Cichlid Fishes of the Great Lakes of Africa* (Oliver & Boyd, Edinburgh, 1972).
5. Mayr, E. in *Evolution of Fish Species Flocks* (eds Echelle, A. A. & Kornfield, I.) 3–11 (University of Maine Press, Orono, 1984).
6. Dominey, W. J. in *Evolution of Fish Species Flocks* (eds Echelle, A. A. & Kornfield, I.) 231–249 (University of Maine Press, Orono, 1984).
7. Meyer, A., Kocher, T. D., Basasibwaki, P. & Wilson, A. C. *Nature* **347**, 550–553 (1990).
8. Larson, A. in *Speciation and its Consequences* (eds Otte, D. & Endler, J. A.) 579–598 (Sinauer, Sunderland, Massachusetts, 1989).
9. Selander, R. K., Yang, S. Y., Lewontin, R. C. & Johnson, W. E. *Evolution* **24**, 402–414 (1970).
10. Poll, M. *Ac. Roy. de Belg. Memories de la classe des sciences. Collection in 8°–2e série, T.XLV, Fasc.* **2**, 1–16 (1986).
11. Nishida, M. *Experientia* **47**, 974–979 (1991).
12. Greenwood, P. H. in *Evolution of Fish Species Flocks* (eds Echelle, A. A. & Kornfield, I.) 141–154 (University of Maine Press, Orono, 1984).
13. Brichard, P. *Cichlids of Lake Tanganyika,* (TFH Publications Inc., Neptune City, New Jersey, 1989).
14. Sturmbauer, C. thesis, Univ. Innsbruck (1990).
15. Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. *J. molec. Evol.* **18**, 225–239 (1982).
16. Tiercelin, J.-J., Mondeguer, A. in *Lake Tanganyika and its Life* (ed. Coulter, G. W.) 7–48 (Oxford Univ. Press, London, 1991).
17. Scholz, C. A. & Rosendahl, B. R. *Am. Assoc. Petrol. Geol: Memoir* **50**, 151–168 (1990).
18. Marlier, G. *Rev. Zool. Bot. Afr.* **59**, 164–183 (1959).
19. Charlesworth, B., Lande, R. & Slatkin, M. *Evolution* **36**, 474–498 (1982).
20. Charlesworth, B. & Lande, R. *Nature* **296**, 610–611 (1982).
21. Turner, J. R. in *Patterns and Processes in the History of Life* (eds Raup, D. M. & Jablonski, D.) 183–207 (Springer, Heidelberg, 1986).
22. Nelissen, M. *Rev. Zool. Afr.* **90**, 17–29 (1976).
23. West-Eberhard, M. J. *Q. Rev. Biol.* **58**, 155–183 (1983).
24. Coyne, J. A. *Nature* **355**, 511–515 (1992).
25. Owen, R. B. *et al. Proc. R. Soc.* B **240**, 519–553 (1990).
26. Kocher, T. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 6196–6200 (1989).
27. Gasse, F., Ledee, V., Massault, M. & Fontes, J. C. *Nature* **342**, 57–59 (1989).
28. Kornfield, I. in *Cichlid Fishes: Behavior, Ecology and Evolution* (ed. Keeleyside, M. H. A.) 103–128 (Chapman & Hall, London, 1991).
29. Swofford, D. L. *Phylogenetic Analysis Using Parsimony, Version 3.0r* (Illinois Natural History Survey, Champaign, 1991).
30. Saitou, N. & Nei, M. *Molec. Biol. Evol.* **4**, 406–425 (1987).