22 Bakker, T.C.M. *et al.* (1999) Sexual selection – condition-related mate choice in sticklebacks. *Nature* 401, 234
23 ten Cate, C. and Vos, D.R. (1999) Sexual imprinting and evolutionary processes in birds: a reassessment. *Adv. Stud. Behav.* 28, 1–31
24 Kendrick, K.M *et al.* (1998) Mothers determine sexual preferences. *Nature* 395, 229–230
25 Laland, K.N. (1994) On the evolutionary consequences of sexual imprinting. *Evolution* 48, 477–489
26 Weary, D.M. *et al.* (1993) A product of discriminative learning may lead to female preferences for elaborate males. *Evolution* 47, 333–336
27 Witte, K. *et al.* (2000) Sexual imprinting on a novel adornment influences mate preferences in the Javanese Mannakin *Lonchura leucogastroides. Ethology* 106, 349–363
28 Tomlinson, I.P.M. and O'Donald, P. (1996) The influence of female viability differences on the evolution of mate choice. *Heredity* 77, 303–312
29 Moore, A.J. *et al.* (1998) The influence of direct and indirect genetic effects on the evolution of behavior: social and sexual selection meet maternal effect. In *Maternal Effects as Adaptations* (Mousseau, T.A. and Fox, C.W., eds), pp. 22–41, Oxford University Press
30 Wolf, J.B. *et al.* (1999) The role of maternal and paternal effects in the evolution of parental quality by sexual selection. *J. Evol. Biol.* 12, 1157–1167
31 Price, T. *et al.* (1993) Sexual selection when the female directly benefits. *Biol. J. Linn. Soc.* 48, 187–211
32 Birgersson, B. and Ekvall, K. (1997) Early growth in male and female fallow deer fawns. *Behav. Ecol.* 8, 493–499
33 Qvarnström, A. (1998) Genotype-by-environment interactions in the determination of the size of a secondary sexual character in the collared flycatcher (*Ficedula albicollis*). *Evolution* 53, 1564–1572

New markers for new species: microsatellite loci and the East African cichlids

Jeffrey A. Markert, Patrick D. Danley and Matthew E. Arnegard

The radiation of the East African cichlid fishes has engaged biologists for over a century. Because so much taxonomic diversity has evolved recently, they are an ideal natural system in which to study the process of speciation. Hypervariable microsatellite loci have been used to verify multiple paternity and maternity in cichlid broods, to quantify the fitness of cooperative breeders and reproductive parasites, to estimate effective population sizes in captive populations, and to illuminate the spatial and temporal scale of gene flow among natural populations. The patterns that have emerged from these studies often reflect important biological differences among taxa. The cichlid species of East Africa represent a large amount of taxonomic and adaptive diversity all neatly packaged into a single lineage and confined to a modest geographical area. Data from microsatellite loci are now providing us with the means to understand one of the world's most intriguing and instructive comparative evolutionary systems.

Jeffrey A. Markert*
Dept of Genetics, Nelson Labs, Rutgers University, Piscataway, NJ 08854, USA.
***e-mail: markert@mbcl.rutgers.edu**

Patrick D. Danley
Dept of Zoology, Rudman Hall, University of New Hampshire, Durham, NH 03824, USA.

Matthew E. Arnegard
Dept of Neurobiology and Behavior, Seeley G. Mudd Hall, Cornell University, Ithaca, NY 14853, USA.

In a 1998 *TREE* article, Galis and Metz[1] asked 'Why are there so many cichlid species?' Their answer involved an evolutionarily and developmentally plastic feeding apparatus, which permits the exploitation of many distinct trophic niches[2], acting in concert with rapidly evolving mate-recognition systems[3]. This hypothesis is consistent with recent theoretical studies suggesting that reproductive isolation might evolve in sympatry when a genetic correlation develops between a component of the mate recognition system and a competitively adaptive trait[4,5].

Although theoretical advances can help us understand the origins of species richness among the East African cichlids, the diversity of this group affords extraordinary opportunities for the direct observation of animal speciation as a process. While cichlids share most aspects of their basic bauplan, the lineage contains a diversity of ecological adaptations and reproductive strategies. Furthermore, each of several behavioral, ecological and biogeographical descriptors encompasses numerous species, many of which transcend phylogenetic or biogeographical boundaries. Thus, the diversity of the East African cichlids provides an ideal and replicated comparative system in which to test the predictions of speciation theories, and might serve as a paradigm for evolutionary biologists and ecologists studying other systems.

Before the mid-1990s, the available genetic markers provided insufficient resolution to adequately describe cichlid mating strategies, population genetics and phylogeny. Although the 'rich, speculative'[6] cichlid literature contains a wealth of research on the phenotypic raw material present, it has until recently lacked evidence for the phylogenetic and population genetic underpinnings needed to understand the divergence of lineages. Here, we focus on recently obtained MICROSATELLITE data, which have overcome many of the persistent difficulties inherent in the application of other markers.

The trouble with cichlids

The evolution of taxonomic diversity has been prolific and recent in the East African cichlids, with thousands of endemic species having evolved in the three Great Lakes since the late Pleistocene[7]. This

http://tree.trends.com 0169–5347/01/$ – see front matter PII: S0169-5347(00)02064-4

rampant splitting of lineages provides numerous opportunities for studying the divergence process *in situ*. Ironically, certain aspects of cichlid evolution are difficult to resolve, precisely because speciation is so recent. In species from the two youngest Great Lakes (Malawi and Victoria), allozymes and mtDNA contain low levels of allelic or sequence polymorphism. The brief periods between speciation events cause what little variation exists at these markers to be shared among taxa[8]. These types of marker are not definitive in studies of assortative mating[9], only mildly informative for population studies[10,11], and are phylogenetically informative in only the oldest lineages. By contrast, simple sequence repeat loci[12] (SSRs, also known as microsatellites), with their high mutation rates and high levels of polymorphism, have clarified aspects of the reproductive behaviour and population genetics in this most speciose vertebrate family.

Mating systems and microsatellites

Paternity and maternity

A high level of parental care is characteristic of the cichlids. In Lakes Malawi and Victoria, and in many Tanganyikan lineages, care is provided by the female who incubates eggs and fry in her mouth until the juveniles are well developed. Generally males form leks on the open sand or in the water column or defend small, spatially clustered territories among the rocks, from which they court passing females. Behavioral observations suggested that gravid females are courted by multiple males in sand-dwelling species[13], but neither the reproductive success of these males, nor whether rock-dwelling females mate with more than one male, was known. Kellogg *et al.*[14] and Parker and Kornfield[15] used SSR loci to estimate the number of males contributing to a brood in several sand- and rock-dwelling species from Lake Malawi. Most broods surveyed from both habitat guilds are sired by two or more males. More surprisingly, brood mixing was documented in the Malawi cichlid *Protomelas* cf. *spilopterus* in four of six broods surveyed, an example of multiple maternity[16]. This was found to be true even for broods that were apparently too young to have been released on brief foraging trips, thus suggesting brood mixing might be under maternal control. The adaptive significance of multiple paternity and maternity in Lake Malawi has yet to be explored experimentally.

Mate recognition

Mate recognition is widely thought to be of fundamental importance to animal speciation. In the East African Cichlidae, putative sibling morphospecies are often distinguished by striking differences in male breeding coloration, whereas shape differences are often subtle[17,18]. In laboratory settings, hybrids between taxa from the same lake are often fully fertile. Behavioral experiments suggest that females mate assortatively on the basis of male color patterns[3]. Using a combination of behavioral observations and SSR genotypes, van Oppen *et al.*[19] tested mating preferences between sympatric color morphs (putative species) in the field. High and statistically significant interspecies fixation indices allowed the authors to reject the null hypothesis of genetic panmixia between sympatric color forms and to conclude that premating isolation is well developed among these morphospecies. Many alleles are shared between taxa, but it is not obvious whether this is a result of shared ancestry or a result of low levels of ongoing gene flow (*sensu* Wang *et al.*[20]).

Brood-care helpers and reproductive parasites

Microsatellite and other genetic fingerprinting methods have been used to document both reproductive parasitism and cooperative breeding in two substrate-spawning species from Lake Tanganyika. *Lamprologus callipterus* females incubate their eggs inside snail shells, which have been gathered into nests by the larger of two male morphs[21]. Smaller 'dwarf' males act as nest parasites by hiding inside shells and fertilizing eggs[22]. Analysis of SSR alleles indicates that these 'satellite' males are successful, siring up to 90% of the progeny when they are present in a shell (P. Meidl, PhD thesis, University of Vienna, 1999).

Fitness benefits for 'helpers at the nest' have been documented in the social Tanganyikan cichlid *Neolamprologus pulcher*. Breeding units in this species consist of a pair of large reproductive adults and several smaller individuals who assist with brood care and defense[23]. DNA fingerprinting techniques in laboratory experiments demonstrate that when a mature male helper is present, he sires about 10% of the offspring[24]. Microsatellite analysis of wild *N. pulcher* populations further demonstrated that these family groups are largely matrilineal and that levels of inbreeding within a group are minimized as a result of high levels of male exchange between groups (M. Taborsky, pers. commun.).

Population structure

Levels of migration

It has long been suspected that gene flow between populations in many cichlid species is low and that this has an important influence on the speciation process, either as a result of distinct selective environments or through sexual selection[25]. Among the endemic and colorful rock-dwelling 'MBUNA' species of Lake Malawi, lack of larval dispersal, heavy predation pressure[26], narrow depth tolerance[27] and the patchy distribution of preferred habitat are all believed to contribute to high levels of population subdivision. Early biochemical and molecular studies were consistent with limited gene flow, but the markers employed could not adequately define the scale of genetic structure or the nature of any barriers to migration[11]. Recently, SSR loci have allowed a more detailed understanding of the scope

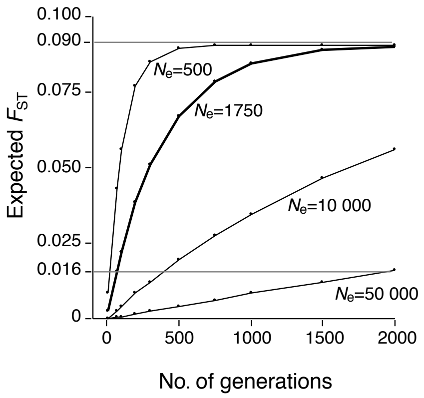
Box 1. Mbuna and the problem of excess polymorphism

Based on their highly significant but small F_{ST} estimates, van Oppen *et al.*[a] speculated that mbuna species can be divided into 'thousands' of semi-isolated genetic units and that this isolation might be an important aspect of the speciation process in cichlids, a statement echoed in subsequent papers[b–d]. Citing difficulties with the interpretation of migration indices when data from highly polymorphic loci are used, Hedrick[e] has recently questioned this prevailing view. F_{ST} values are constrained by within-population homozygosity, which is often rather low at microsatellite markers. These markers also provide researchers with large amounts of statistical power, permitting the detection of allele frequency differences that are statistically real, but which might have little biological importance.

Are levels of genetic difference detected in these recent mbuna studies biologically important? The migration estimates in these studies are high enough to prevent fixation of alternative alleles between adjacent populations through drift alone. However, an alternative estimator of migration, Barton and Slatkin's *N*m(Ref. f), suggests that some, more widely separated mbuna populations are indeed genetically decoupled[d]. Moreover, Danley *et al.*[c] have discussed the relevance of divergence with gene flow models to the evolution of the East African cichlids.

Reduced migration between habitat patches is important biologically if selective forces are strong enough to overcome the homogenizing effects of interdemic migration (*sensu* Endler[g]). Selective forces might include different modes of sexual selection or adaptation to variable community structures or abiotic factors that can vary widely between habitat patches[h].

I



TRENDS in Ecology & Evolution

The sampling formulae of Jin and Chakraborty[i] can be used to place migration estimates between recently separated populations in context for a given level of heterozygosity (following van Oppen *et al.*[a]). *Pseudotropheus tropheops* 'mauve' populations have an estimated F_{ST} of 0.016. Figure I shows the null expectation of F_{ST} in the complete absence of gene flow over time for four different N_e values. The F_{ST} value estimated by van Oppen *et al.* for this species is consistent with an N_e of 1750 or fewer individuals in populations that have been completely isolated for 67 generations. A recent re-evaluation of geological data[j] suggests that the populations could have been isolated for 2000 generations instead of the 67 generations assumed by van Oppen *et al.*, consistent with an N_e of 50 000 individuals, or far fewer individuals than this assuming extensive gene flow. In such a large population, deterministic forces might become more important than stochastic forces.

It is difficult to make inferences about the evolution of cichlids from survey data alone. The problem is compounded because divergence times, census and effective population sizes, SSR mutation rates and generation times – each crucial components of the balance between selection and migration – are so far a matter of informed conjecture. Although our present knowledge does not permit us to reliably estimate absolute levels of gene flow, the absence of panmixia between proximate populations indicates that philopatry is an aspect of cichlid evolution worthy of more formal exploration.

References

a van Oppen, M.J.H. *et al.* (1997) Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proc. R. Soc. London B Biol. Sci.* 264, 1803–1812
b Arnegard, M.E. *et al.* (1999) Population structure and colour variation of the cichlid fish *Labeotropheus fuelleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. *Proc. R. Soc. London B Biol. Sci.* 266, 119–130
c Danley, P. *et al.* (2000) Divergence with gene flow in the rock-dwelling cichlids of Lake Malawi. *Evolution* 54, 1725–1737
d Markert, J.A. *et al.* (1999) Biogeography and population genetics of the Lake Malawi cichlid *Melanochromis auratus*: habitat transience, philopatry and speciation. *Mol. Ecol.* 8, 1013–1026
e Hedrick, P. (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53, 313–318
f Barton, N.H. and Slatkin, M. (1986) A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56, 409–415
g Endler, J.A. (1973) Gene flow and population differentiation. *Science* 179, 243–250
h Ribbink, A.J. *et al.* (1983) A preliminary survey of the cichlid fishes of the rocky habitats in Lake Malawi. *S. Afr. J. Zool.* 18, 149–310
i Jin, L. and Chakraborty, R. (1995) Population structure, stepwise mutations, heterozygote deficiency and their implications in DNA forensics. *Heredity* 74, 274–285
j Nicholson, S.E. (1998) Fluctuations of rift valley lakes Malawi and Chilwa during historical times: a synthesis of ecological, archaeological and historical information. In *Environmental Change and Response in East African Lakes* (Lehman, J.T., ed.), pp. 207–231, Kluwer Academic Press

and nature of gene flow in both Lake Malawi and the Lake Victoria Region.

At Lake Malawi, van Oppen *et al.*[28] sampled four mbuna species from habitat patches along a 3 km stretch of the western shore. They calculated small but statistically significant F_{ST} values. The estimated number of migrants (*N*M) exchanged between populations ranges from 6.3 to 8.3 individuals per generation, on a scale of less than 1 km. On the basis of these estimates, the authors speculated that mbuna populations can be divided into 'thousands' of semi-independent genetic subunits (Box 1). If these preliminary estimates are roughly correct, modest selective pressures could lead to phenotypic divergence in long-lived populations, assuming that effective population sizes (N_e) are in the order of tens of thousands.

The biogeographical aspects of population structure in mbuna species (Fig. 1) were explored in surveys conducted in southern Lake Malawi[29–31].

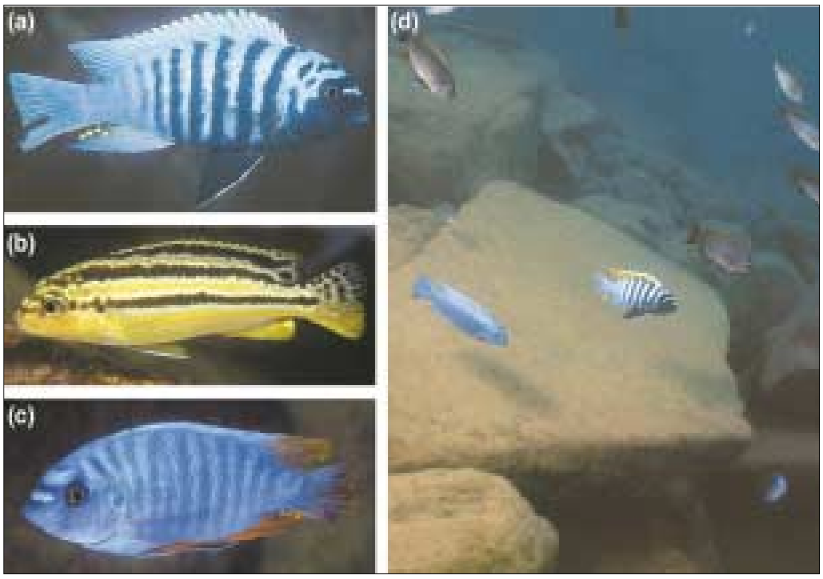


Fig. 1. Three rock-dwelling mbuna species from Thumbi Island West (Cape Maclear, southern Lake Malawi): (a) ***Metriaclima zebra*****; (b)** ***Melanochromis auratus*****; and (c)** ***Labeotropheus fuelleborni*****. (d) Typical mbuna habitat at Nakantenga Island.** ***Reproduced, with permission, from Ad Konings.***

Samples were collected from each major rocky habitat patch along a 42 km stretch of shoreline adjacent to the Nankumba Peninsula. All four species showed evidence of restricted gene flow between populations. *Melanochromis auratus* has the highest level of population differentiation, with an F_{ST} estimate of 0.15 overall. In fact, the *N*m between the two most distant populations sampled is estimated to be below the 0.5 individuals per generation necessary to allow the fixation of alternate alleles through drift alone. The other three species were estimated to have higher levels of gene flow, with statistically significant F_{ST} estimates ranging from 0.063 in *Labeotropheus fuelleborni* populations to 0.016 in the three *Metriaclima sandaracinos* populations (the latter species has a limited distribution in the study area). A comparison of estimates of gene flow between adjacent habitat patches suggests ecologically important differences between these taxa (Fig. 2). For example, in *M. auratus* and *L. fuelleborni* the lowest migration estimates were obtained between populations separated by deep water. This finding is consistent with physiological studies that have demonstrated that these benthic fishes are not capable of rapidly adjusting to changes in water depth[27]. These same deep water channels did not restrict migration in *Metriaclima zebra* to the same degree, leading Danley *et al.*[30] to speculate that *M. zebra* might cross them while feeding in the water column.

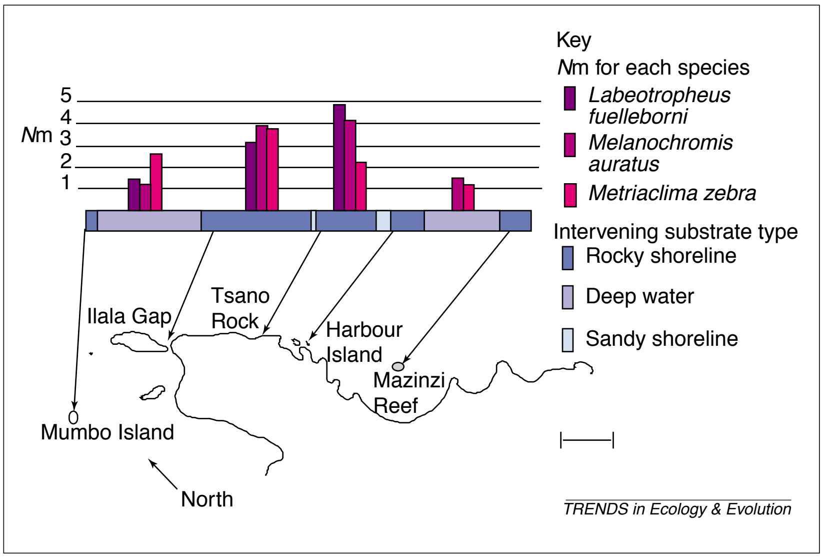


Fig. 2. Relative migration rates between three mbuna species along the shores of Lake Malawi's Nankumba Peninsula. The relative proportions of different substrate types are shown on the horizontal axis of the graph[18]. ***N*****M estimates shown here were determined using Barton and Slatkin's[46] private alleles method. Although all three species occupy the same rocky habitat, the constraining effects of different substrate types on levels of migration appear to vary between taxa.** ***Labeotropheus fuelleborni*** **is not present at Mazinzi Reef. Scale bar = 5km. Abbreviation:** ***N*****M, estimated number of migrants.**

Although a picture of limited migration in mbuna is emerging, species-specific factors might influence the relative levels of migration. Barrier and habitat types were not well replicated within the surveyed transect; thus these observations merit further study. Nevertheless, a picture of high levels of mbuna philopatry regulated by species-specific dispersal is intriguing, given that each of the Great Lakes has a large endemic assemblage of rock-fish and a patchy distribution of habitats (but see Box 2).

A Lake Victoria Region (LVR) species, surveyed with different SSR loci, illustrates a different pattern of population structure. Wu sampled several populations of the LVR cichlid *Astatoreochromis alluaudi*, a habitat generalist found in several LVR lakes and streams (L. Wu, PhD thesis, Ohio State University, 1999). No clear evidence of population structure was found within lakes. However, moderate levels of genetic structure were observed in all of the between-lake comparisons. Overall, the scale of genetic subdivision in *A. alluaudi* is much larger than in the Malawi mbuna. Wu notes that *A. alluaudi* has an extremely broad habitat distribution and that it belongs to one of the least speciose haplochromine genera. Indeed, this genus does not seem to have speciated at all within the LVR, suggesting that philopatry might be a precondition for high levels of divergence among the East African cichlids.

Evidence for philopatry also exists within habitat patches. Knight *et al.*[32] sampled *Metriaclima* from a 20 × 20 m grid. Using SSR-based estimates, they found a weak correlation between proximity and relatedness for females but not for males, providing evidence of male-biased dispersal and female philopatry. The authors reason that in a sexually selected species this could lead to prolific speciation. Migrating males would suffer greatly reduced mating success in foreign demes if females preferred male morphotypes from their own deme, and interdemic gene exchange would be relatively low. By contrast, male philopatry and female dispersal would result in females 'settling' for heterotypic males, who would readily mate with them, leading to the collapse of population-specific mating patterns. They hypothesized that inbreeding avoidance is a consequence of male dispersal behavior in this species

Box 2. Sympatry, speciation and cichlids

The importance of geographic isolation to the speciation process has been a long running and controversial theme in cichlid studies. Mayr[a], for example, eventually supported an older 'microallopatric' model in which speciation occurs at isolated habitat patches within a lake, abandoning a stricter view in which divergence originated in separate bodies of water.

The apparent monophyly of many lineages within the Great Lakes[b] (but see Nagl *et al.*[c]), the patchy distribution of rocky, sandy and muddy habitats within the lakes[d], and recent studies of population structure in Lake Malawi, which suggest that inter-population migration is restricted[e–h], are all compatible with the microallopatric view of speciation. However, none of these documented patterns exclude alternative modes of speciation.

In 1994, Schliewen *et al.*[i] presented data supporting strictly sympatric speciation in cichlid lineages within two small crater lakes in West Africa. Unlike the East African Great Lakes, the substrate in the crater lakes is homogeneous, presumably providing few opportunities for microgeographical isolation. mtDNA sequence data suggest that the several species within each crater lake arose recently from a single common ancestor, leading the authors to infer that these taxa evolved in the absence of geographic barriers.

Dominey[j] and other authors have stressed the importance of sexual selection as a driving force in the cichlid speciation process. Seehausen and van Alphen[k] have suggested that sexual selection might permit speciation, even in the absence of microgeographical isolation. In Lake Victoria, fish species with nearly identical morphologies but with distinct male breeding coloration often coexist in the same habitat patches. Based on experiments that document the importance of these color differences in mate recognition, and on the overlap of geographic distributions of congeners, Seehausen and van Alphen[k] argue that much speciation occurs in strict sympatry as a result of sexual selection. Although the distribution overlap observed by Seehausen and van Alphen[k] is consistent with sympatric speciation, Bouton[l] has argued that a similar pattern would result from speciation in allopatry followed by parallel colonization of linearly arrayed habitat.

At present, we are left with several unresolved questions and few excluded hypotheses. Does speciation occur primarily in allopatry on fine geographic scales as a result of divergent selection, either in the form of adaptation or through the evolution of site specific mating systems? Is there a role for reinforcement in these taxa that appear to have few post-mating reproductive barriers? Does rampant speciation in sympatry through sexual selection explain the rapid evolution of thousands of cichlid species? Or are a variety of mechanisms operating in parallel to generate this extraordinary taxonomic diversity?

References

a Mayr, E. (1963) *Animal Species and Evolution* Belknap Press

b Meyer, A. *et al.* (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* 347, 550–553

c Nagl, S. *et al.* (2000) The origin and age of haplochromine fishes in Lake Victoria, East Africa. *Proc. R. Soc. London B Biol. Sci.* 267, 1049–1061

d McKaye, K.R. and Gray, W.N. (1984) Extrinsic barriers to gene flow in rock-dwelling cichlids of Lake Malawi: macrohabitat heterogeneity and reef colonization. In *Evolution of Fish Species Flocks* (Echelle, A.A. and Kornfield, I., eds), pp. 169–184, University of Maine at Orono Press

e van Oppen, M.J.H. *et al.* (1997) Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proc. R. Soc. London B Biol. Sci.* 264, 1803–1812

f Arnegard, M.E. *et al.* (1999) Population structure and colour variation of the cichlid fish *Labeotropheus fuelleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. *Proc. R. Soc. London B Biol. Sci.* 266, 119–130

g Danley, P. *et al.* (2000) Divergence with gene flow in the rock-dwelling cichlids of Lake Malawi. *Evolution* 54, 1725–1737

h Markert, J.A. *et al.* (1999) Biogeography and population genetics of the Lake Malawi cichlid *Melanochromis auratus*: habitat transience, philopatry and speciation. *Mol. Ecol.* 8, 1013–1026

i Schliewen, U.K. *et al.* (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* 368, 629–632

j Dominey, W.J. (1984) Effects of sexual selection and life history on speciation: species flocks in African cichlids and Hawaiian *Drosophila*. In *Evolution of Species Flocks* (Echelle, A.A. and Kornfield, I., eds), pp. 231–249, University of Maine at Orono Press

k Seehausen, O. and van Alphen, J. (2000) Can sympatic speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.* 2, 262–271

l Bouton, N. (2000) Progressive invasion and allopatric speciation can also explain distribution patterns of rock-dwelling cichlids from southern Lake Victoria: a comment on Seehausen and van Alphen. *Ecol. Lett.* 3, 166–169

and that male dispersal might be a result of the limited number of available male breeding territories.

Genetic variation, habitat age and secondary contact

As a result of fluctuations in lake water level, shoreline habitats in the East African Great Lakes are chronically unstable. The rocky habitats in southern Lake Malawi would have been dry land during the most recent major drop in lake water level, which might have occurred in the past few hundred years[7] (but see Ref. 33). Habitats that today have rocks extending to a great depth would have become available for colonization earlier, as the lake basin refilled. A correlation between heterozygosity and the depth of the rock–sand interface (a proxy for habitat age and stability) was observed in both *M. zebra* and *M. auratus*[30,31]. The reduction in heterozygosity at shallower sites is not large. However, it does suggest that each new habitat patch was founded by a moderate number of individuals in a pattern of serial colonization from deeper to shallower sites. Evidence of the occurrence of genetic bottlenecks during this process is lacking, an interpretation consistent with previous studies[34]. Deep habitats are located in the northern half of the transect surveyed and shallower habitats are found to the south. Because deeper habitats might be capable of supporting larger populations, alternative explanations for this pattern are possible.

Secondary contact might be responsible for a cline in male breeding coloration in the early colonizer *L. fuelleborni*, which was observed along the same transect. Based on a qualitative assessment of allele

Box 3. Microsatellite mutation models and estimates of divergence

There has been an explosion of theoretical and empirical treatments of microsatellite mutation processes over the past decade. Currently, estimates of levels of genetic divergence or patterns of relatedness between populations and species rely on statistics that have been developed from two classic models.

Stepwise Mutation Model (SMM)

SMM has emerged as the 'null model' of microsatellite evolution[a]. The strict SMM assumes that mutations involve the loss or gain of a single repeat unit (thereby retaining some 'memory' of previous allelic states) and that allele sizes are unconstrained, which might hold true over short time spans[b,c]. SMM-based statistics require that a stepwise mutation process has more influence on allele frequency distributions than genetic sampling does. Based on evidence of non-stepwise mutational events at some microsatellite loci, several investigators have called for the identification and selection of mutationally 'well behaved' microsatellites[d] for which to apply estimators of genetic divergence based on the SMM.

Infinite Alleles Model (IAM)

IAM was developed by Kimura and Crow[e], and assumes that every mutational event creates a unique allele. As with the SMM, the IAM also assumes that there are no constraints on allele size. Statistics based on this mutational model violate *a priori* the well supported notion that there is some memory in the microsatellite mutation process. However, they might still perform well when the time since divergence is extremely short, when stochastic sampling of alleles far outweighs mutation in importance, and when the frequency of mutational events involving relatively large changes is high[b].

More complex microsatellite mutation models

Evidence is accumulating that there are several complications involved in the mutation of microsatellites, including: larger mutational events, a dependence of mutation rates on the structure of the tandem repeat motif, a predisposition for microsatellite expansion over contraction, allele size constraints (e.g. reflecting or absorbing boundaries), and the possibility of heterozygote instability (reviewed by Amos[a] and Estoup and Cornuet[d]). Di Rienzo *et al.*[f] introduced the two-phase mutation model (TPM) in which a mutation has a probability *p* of being a one step change and a probability 1–*p* of being a larger change, drawn from a symmetric geometric distribution with specified variance. Numerous additional models have been formulated more recently. Difficulties associated with the estimation of required parameters impede the development of estimators of genetic divergence and population demographic statistics from these more complex mutation models.

References

a Amos, W. (1999) A comparative approach to the study of microsatellite evolution. In *Microsatellites: Evolution and Applications* (Goldstein, D.B. and Schlötterer, C., eds), pp. 66–79, Oxford University Press

b Slatkin, M. (1995) Measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139, 457–462

c Goldstein, D.B. *et al.* (1999) The use of microsatellite variation to infer population structure and demographic history in a natural model system. *Genetics* 151, 797–801

d Estoup, A. and Cornuet, J-M. (1999) Microsatellite evolution: inferences from population data. In *Microsatellites: Evolution and Applications* (Goldstein, D.B. and Schlötterer, C., eds), pp. 49–65, Oxford University Press

e Kimura, M. and Crow, J.F. (1964) The number of alleles that can be maintained in a finite population. *Genetics* 49, 725–738

f Di Rienzo, A. *et al.* (1998) Heterogeneity of microsatellite mutations within and between loci, and implications for human demographic histories. *Genetics* 148, 1269–1284

frequency and biogeographical data, Arnegard *et al.*[29] hypothesized that this cline might represent introgression between divergent lineages that came into contact via migration along a historical shoreline route, rather than originating *in situ*.

Phylogeny quest

In Lakes Malawi and Victoria, short divergence times and the problem of retained ancestral polymorphisms have made the determination of phylogenetic relationships difficult[8]. The situation in these lakes contrasts with Lake Tanganyika, which appears to be a reservoir of ancient lineages[35]. Amplified fragment length polymorphism (AFLP)-based techniques have been used to compensate for these problems in Lake Malawi by generating an abundance of characters, but they appear to have little power to resolve intergeneric relationships in this system, perhaps because many genera evolved so rapidly[36]. Kornfield and Parker[37] and Sültmann and Mayer[38] have suggested that microsatellite loci might have the potential to overcome this difficulty owing to their high mutation rates and the potential information content of SSR alleles evolving under a stepwise mutation model (Box 3). In a limited pilot study using two species in each of three morphologically distinct genera, Kornfield and Parker[37] found that they could cluster species into generic dyads for two out of the three genera. Further, they proposed that microsatellite 'saltines' (widely disjunct clusters of alleles) might be interpretable in a cladistic manner as synapomorphies. Sültmann and Mayer[38] generated a phylogeny that roughly reflects two trophic guilds within Lake Victoria. Additionally, Wu used 14-locus SSR genotypes to estimate phylogenetic relationships among 23 taxa from the LVR (L. Wu, PhD thesis, Ohio State University, 1999). Although encouraging, these results are sensitive to mutation model assumptions. Unfortunately, independent estimates of phylogenetic relationships against which to compare these results are not available.

Microsatellite loci might ultimately be useful in estimating phylogenies, although the analytical and sampling challenges are considerable. The complexities of microsatellite mutations violate the strict assumptions of the genetic distance measures

Glossary

Amplified fragment 'length' polymorphism (AFLP)[a]: a technique for finding genetic variation at many loci simultaneously. Widely used in genome mapping, it has recently been used in phylogeny estimation and might be useful in population studies. In AFLP, genomic DNA is cut with restriction enzymes. Restriction site specific oligonucleotide adaptors are then ligated onto the entire population of restriction fragments. Next, PCR is performed using labeled primers that recognize the adaptors and several additional base pairs. This permits rapid visualization of a subset of restriction sites using high resolution polyacrylamide gel electrophoresis.

F_{ST}: an index of population differentiation, commonly used to infer levels of migration between populations. F_{ST} ranges between zero and one. Higher F_{ST} values are interpreted to indicate a lower level of gene exchange between populations. Originally defined by Sewall Wright as $F_{ST} = (H_T - \bar{H}_S)/H_T$ where $\bar{H}_S$ is the average expected heterozygosity within each population and H_T is the expected total heterozygosity calculated from the average of population allele frequencies. Wright's original example was for a single locus with two alleles. Numerous multi-locus, multi-allelic analogues have been developed[b].H

Homoplasy: evolutionary convergence. In molecular and biochemical data sets homoplasy refers to identity by character state that is not a result of evolutionary descent. Homoplasy may plague microsatellite studies where individual alleles are typically defined solely by the size of a PCR product.

Mbuna: a habitat guild composed of a diverse group of small, typically colorful, rock-dwelling cichlid species from Lake Malawi. At present, 11 genera are placed in this group: *Cyathochromis*, *Cynotilapia*, *Genyochromis*, *Gephyrochromis*, *Iodotropheus*, *Labeotropheus*, *Labidochromis*, *Melanochromis*, *Metriaclima* (alternatively known as *Maylandia*), *Petrotilapia* and *Pseudotropheus*. This group is widely believed to be monophyletic, but phylogenetic analyses have yet to conclusively demonstrate this because of the difficulties resulting from the rapid speciation of East African cichlids in general.

Microsatellite locus: a class of molecular marker consisting of a short DNA motif (< 6 bp) repeated many times. Mutations at these loci are most often the result of an increase or decrease in the number of times the motif is repeated. Ideally, alleles at these loci can be distinguished by the number of repeated units (but see Ref. c). Typically, a large number of alleles are present at these loci, making them an ideal tool for population genetics, genomic mapping and paternity assignment. Because these repeated motifs are usually flanked by single-copy DNA embedded in non-repeating regions, the areas adjacent to the repeated motif can be used to design a set of PCR primers specific to a particular locus, providing an easy way to determine genotypes at these loci. These loci are also referred to as simple sequences, simple sequence repeat loci (SSRs), simple tandem repeat polymorphisms (STRPs) or variable number of tandem repeats (VNTRs).

Nm: Barton and Slatkin's[d] index of migration, which derives an estimate of the number of migrants exchanged between populations in each generation from the number of alleles which are unique to a particular population. This is referred to as the 'private alleles' or 'rare alleles' method. This method might be especially useful for analyzing data from highly polymorphic loci such as SSRs.

References

a Vos, P. *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23, 4407–4414

b Hedrick, P. (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53, 313–318

c van Oppen, M.J.H. *et al.* (2000) Extensive homoplasy, nonstepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi cichlids. *Mol. Biol. Evol.* 17, 489–498

d Barton, N.H. and Slatkin, M. (1986) A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56, 409–415

that are currently available. Size HOMOPLASY, which is expected to increase with the mutation rate and the time since divergence[39], is another major problem. In addition, retained ancestral polymorphisms would result in phylogenetic noise at microsatellite loci[40], as with other classes of molecular markers, when closely related taxa are being compared. The high mutation rate and putative mutational 'memory' encoded in allele sizes[41] might provide a partial solution to these problems. If these methods are to succeed in revealing relationships between recently diverged taxa, homology between frequently employed complex repeat arrays must be better established by sequence analysis[40,42], and the appropriateness of distance estimators needs to be evaluated on a taxon- and locus-specific basis (Box 3).

Conclusions

Why are there so many cichlids? The question remains largely unanswered, but after a century of study the theoretical and genetic tools necessary to address it might at last be available. The data derived from microsatellite markers are contributing to the framework needed for a more comprehensive understanding of cichlid evolution, revealing tantalizing differences between these closely related taxa. New insights into reproductive behavior and patterns of assortative mating have been obtained using these markers, and reproductive success and relatedness have been estimated with far higher levels of confidence than could have been achieved with behavioral observations alone.

The implications of the mating system characteristics quantified using SSR loci are not yet clear. Multiple paternity[14,15] might imply that polyandry decreases the variance in male reproductive output (and thus increases the effective population size, N_e) by providing more mating opportunities for males. However, the element of female choice implied by these findings and supported by behavioral studies[43] might mean that mating success is skewed in favor of a small number of highly successful males, in spite of the leveling effects of multiple paternity within this elite group. This second interpretation is compatible with data presented by Fiumera *et al.*[44] who used SSR loci to estimate N_e in captive breeding populations of two endangered Lake Victoria species. In these populations, N_e was found to be only 1–15% of census size, although it is not clear whether these results are applicable to wild populations. A better understanding of effective population sizes will be crucial for gauging the relative influence of selection and drift in this system, and it is also needed to determine whether ephemeral populations have reached a selection–migration–drift equilibrium.

The relative importance of allopatric and sympatric speciation mechanisms remains an active theme in cichlid studies (Box 2). Because the East African lakes contain closely spaced patches of habitat (and some of the smaller lakes contain only a single kind of habitat) it could be more practical to consider allopatry and sympatry as points along a gene flow continuum, which might or might not have an easily identifiable geographic component. Allopatric populations that

Acknowledgements

We would like to thank C. Caudill, R. Clark, J. Curole, N. Garnhart, P. Hedrick, J. Hey, T. Kocher, I. Kornfield, C. Machado, P. Smouse, J. Streelman, M. Taborsky, G. Turner and three anonymous reviewers who offered valuable advice and guidance. We thank Ad Konings for generously supplying the photographs for Fig. 1.

exchange many migrants might be genetically sympatric. Alternatively, sedentary organisms could show evidence of structure within a continuous habitat. The crucial factor with respect to the evolution of phenotypic differences is the balance between the homogenizing effects of gene flow and the diversifying effect of local selective pressures (*sensu* Endler[45]). Microsatellite loci have been useful in describing the contribution of gene exchange to this balance[30]. The power of highly polymorphic DNA markers to track individuals and lineages will probably be useful in helping to quantify the strength of both natural and sexual selection. When coupled with a better understanding of the community and abiotic characteristics of each habitat patch, a more comprehensive view of cichlid evolution will emerge. Emphasis should be placed on repeatable tests of the predictions offered by speciation theories within the different trophic guilds, reproductive strategies, biogeographical settings, and genetic lineages encompassed within this extraordinary comparative evolutionary system.

References

1 Galis, F. and Metz, J.A.J. (1998) Why are there so many cichlids? *Trends Ecol. Evol.* 13, 1–2

2 Liem, K.F. (1973) Evolutionary strategies and morphological innovations: cichlid pharyngeal jaws. *Syst. Zool.* 22, 425–441

3 Seehausen, O. *et al.* (1998) The effect of male coloration on female mate choice in closely related Lake Victoria cichlids. *Behav. Ecol. Sociobiol.* 42, 1–8

4 Kondrashov, A.S. and Kondrashov, F.A. (1999) Interactions among quantitative traits in the course of sympatric speciation. *Nature* 400, 351–354

5 Dieckmann, U. and Doebeli, M. (1999) On the origin of species by sympatric speciation. *Nature* 400, 354–357

6 Kornfield, I. and Smith, P.F. (2000) African cichlid fishes: model systems for evolutionary biology. *Annu. Rev. Ecol. Syst.* 31, 163–196

7 Owen, R.B. *et al.* (1990) Major low lake levels of Lake Malawi and their implications for speciation rates in cichlid fishes. *Proc. R. Soc. London B Biol. Sci.* 240, 519–553

8 Moran, P. and Kornfield, I. (1993) Retention of an ancestral polymorphism in the Mbuna species flock (Teleostei: Cichlidae) of Lake Malawi. *Mol. Biol. Evol.* 10, 1015–1029

9 McKaye, K.R. *et al.* (1984) Genetic evidence of allopatric and sympatric differentiation among color morphs of a Lake Malawi cichlid fish. *Evolution* 38, 215–219

10 Parker, A. and Kornfield, I. (1997) Evolution of the mitochondrial DNA control region in the mbuna (Cichlidae) species flock of Lake Malawi. *J. Mol. Evol.* 45, 70–83

11 Dorit, R. (1990) Correlates of high diversity in Lake Victoria haplochromine cichlids. In *Causes of Evolution: A Paleontological Perspective* (Ross, R.M. and Allmon, W.D., eds), pp. 322–353, University of Chicago Press

12 Tautz, D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* 17, 6463–6471

13 McKaye, K.R. (1991) Sexual selection and the evolution of the cichlid fishes of Lake Malawi, Africa. In *Behaviour, Ecology, and the Evolution of the Cichlid Fishes* (Keenleyside, M., ed.), pp. 241–257, Chapman & Hall

14 Kellogg, K.A. *et al.* (1995) Microsatellite variation demonstrates multiple paternity in lekking cichlid fishes from Lake Malawi, Africa. *Proc. R. Soc. London B Biol. Sci.* 260, 79–84

15 Parker, A. and Kornfield, I. (1996) Polygynandry in *Pseudotropheus zebra*, a cichlid fish from Lake Malawi. *Environ. Biol. Fishes* 47, 345–352

16 Kellogg, K. *et al.* (1998) Intraspecific brood-mixing and reduced polyandry in a maternal mouthbrooding cichlid. *Behav. Ecol.* 9, 309–312

17 Seehausen, O. *et al.* (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* 277, 1808–1811

18 Ribbink, A.J. *et al.* (1983) A preliminary survey of the cichlid fishes of the rocky habitats in Lake Malawi. *S. Afr. J. Zool.* 18, 149–310

19 van Oppen, M.J.H. *et al.* (1998) Assortative mating among rock-dwelling cichlid fishes supports high estimates of species richness from Lake Malawi. *Mol. Ecol.* 7, 991–1001

20 Wang, R. *et al.* (1997) Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* 147, 1091–1106

21 Sato, T. (1994) Active accumulation of spawning substrate: a determinant of extreme polygyny in a shell-brooding cichlid fish. *Anim. Behav.* 48, 669–678

22 Taborsky, M. (1998) Sperm competition in fish: 'bourgeois' males and parasitic spawning. *Trends Ecol. Evol.* 13, 222–227

23 Taborsky, M. (1984) Broodcare helpers in the cichlid fish *Lamprologus brichardi*: their costs and benefits. *Anim. Behav.* 32, 1236–1252

24 Dierkes, P. *et al.* (1999) Reproductive parasitism of broodcare helpers in a cooperatively breeding fish. *Behav. Ecol.* 10, 510–515

25 Dominey, W.J. (1984) Effects of sexual selection and life history on speciation: species flocks in African cichlids and Hawaiian *Drosophila*. In *Evolution of Species Flocks* (Echelle, A.A. and Kornfield, I., eds), pp. 231–249, University of Maine at Orono Press

26 Fryer, G. (1965) Predation and its effects on migration and speciation in African fishes: a comment. *Proc Zool. Soc. London* 144, 301–322

27 Marsh, A.C. and Ribbink, A.J. (1981) A comparison of the abilities of three sympatric species of *Petrotilapia* (Cichlidae, Lake Malawi) to penetrate deep water. *Environ. Biol. Fishes* 6, 367–369

28 van Oppen, M.J.H. *et al.* (1997) Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proc. R. Soc. London B Biol. Sci.* 264, 1803–1812

29 Arnegard, M.E. *et al.* (1999) Population structure and colour variation of the cichlid fish *Labeotropheus fuelleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. *Proc. R. Soc. London B Biol. Sci.* 266, 119–130

30 Danley, P. *et al.* (2000) Divergence with gene flow in the rock-dwelling cichlids of Lake Malawi. *Evolution* 54, 1725–1737

31 Markert, J.A. *et al.* (1999) Biogeography and population genetics of the Lake Malawi cichlid *Melanochromis auratus*: habitat transience, philopatry and speciation. *Mol. Ecol.* 8, 1013–1026

32 Knight, M. *et al.* (1999) Evidence for male-biased dispersal in Lake Malawi cichlids from microsatellites. *Mol. Ecol.* 8, 1521–1527

33 Nicholson, S.E. (1998) Fluctuations of rift valley lakes Malawi and Chilwa during historical times: a synthesis of ecological, archaeological and historical information. In *Environmental Change and Response in East African Lakes* (Lehman, J.T., ed.), pp. 207–231, Kluwer Academic Press

34 Moran, P. and Kornfield, I. (1995) Were population bottlenecks associated with the radiation of the mbuna species flock (Teleostei: Cichlidae) of Lake Malawi? *Mol. Biol. Evol.* 12, 1085–1093

35 Nishida, M. (1991) Lake Tanganyika as an evolutionary reservoir of old lineages of East African cichlid fishes: inferences from allozyme data. *Experientia* 47, 974–979

36 Albertson, R.C. *et al.* (1999) Phylogeny of a rapidly evolving clade: the cichlid fishes of Lake Malawi, East Africa. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5107–5110

37 Kornfield, I. and Parker, A. (1997) Molecular systematics of a rapidly evolving species flock: the mbuna of Lake Malawi and the search for phylogenetic signals. In *Molecular Systematics of Fishes* (Kocher, T. and Stepien, C., eds), pp. 25–37, Academic Press

38 Sültmann, H. and Mayer, E.W. (1997) Reconstruction of cichlid fish phylogeny using nuclear DNA markers. In *Molecular Systematics of Fishes* (Kocher, T.D. and Stepien, C.A., eds), pp. 39–51, Academic Press

39 Estoup, A. and Cornuet, J-M. (1999) Microsatellite evolution: inferences from population data. In *Microsatellites: Evolution and Applications* (Goldstein, D.B. and Schlötterer, C., eds), pp. 49–65, Oxford University Press

40 van Oppen, M.J.H. *et al.* (2000) Extensive homoplasy, nonstepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi cichlids. *Mol. Biol. Evol.* 17, 489–498

41 Slatkin, M. (1995) Measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139, 457–462

42 Colson, I. and Goldstein, D.B. (1999) Evidence for complex mutations at microsatellite loci in *Drosophila*. *Genetics* 152, 617–627

43 Hert, E. (1991) Female choice based on egg-spots in *Pseudotropheus aurora* Burgess 1976, a rock-dwelling cichlid of Lake Malawi, Africa. *J. Fish. Biol.* 38, 951–953

44 Fiumera, A.C. *et al.* (2000) Effective population size and maintenance of genetic diversity in captive-bred populations of a Lake Victoria cichlid. *Conserv. Biol.* 14, 886–892

45 Endler, J.A. (1973) Gene flow and population differentiation. *Science* 179, 243–250

46 Barton, N.H. and Slatkin, M. (1986) A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* 56, 409–415