Using phylogenies to test evolutionary hypotheses about cichlid fishes

Nicholas B. Goodwin*, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK (e-mail: N.Goodwin@uea.ac.uk); Sigal Balshine, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK (Current address: Department of Psychology, McMaster University, Hamilton, Ontario L8S 4K1, Canada); and John D. Reynolds, School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

ABSTRACT
Cichlid fishes provide a remarkable example of explosive speciation and adaptive radiation. Consequently, the phylogenetic relationships of these fishes have attracted a great deal of attention from biologists. Over a century of morphological investigation into the evolutionary relationships of cichlids has recently been augmented and refined by a proliferation of new data based on molecular genetic techniques. In this review, we first present the key morphological and molecular studies of cichlid relationships. Second, we explain how such evolutionary frameworks can be incorporated into the study of the behaviour, ecology and evolution of these fishes. Several comparative methods are described that are used to identify independent evolutionary events and make statistically valid comparisons. Three examples show how these techniques have been used to study cichlid evolution. We argue that even if one is not interested in evolutionary relationships themselves, an understanding of such relationships is vital for testing hypotheses about morphology, life histories and behaviour.

INTRODUCTION
The Cichlidae is one of the most diverse families of vertebrates. To date approximately 1300 species have been described, and these show enormous diversity in morphology, behaviour and habitat. Cichlids are found in almost every tropical freshwater habitat ranging from the pelagic zones of the Great African Lakes, to benthic regions of small ephemeral South American streams (Nelson, 1994). Their remarkable success has been facilitated by the plasticity of their jaw structure (Casciotta and Arratia, 1993), allowing for considerable diversity in feeding habits across many microhabitats (Yamaoka, 1997; Yanagisawa et al., 1997). All Cichlid species provide parental care for eggs and larvae, with variation among species in the form (e.g. substrate guarding and mouthbrooding) and sex providing care (e.g. male-only, female-only and biparental) (Keenleyside, 1991). Thus, it is not surprising that cichlids have received a great deal of attention from researchers interested in hypotheses of life history and ecology (Keenleyside, 1991; Kawanabe et al., 1997). Cichlids have also been studied in relation to allopatric and sympatric speciation (van Oppen et al., 1998), adaptive radiation (Meyer, 1993; Meyer et al., 1994) and sexual selection (Seehausen et al., 1998).

Cichlid speciation is particularly impressive because many taxa have evolved over a comparatively short period of time, as shown by the short genetic distances between species of *Haplochromis* from Lake Victoria and East African rivers (Mayer et al., 1998) (Fig. 1). The African lakes also display high levels of local endemism, reflecting the scale of the speciation process occurring throughout the lakes (Ribbink, 1991). Extensive research into the systematics of Cichlidae has been conducted for two main
Figure 1. A phylogeny of cichlid species flocks from East and West African lakes and rivers, based on nuclear DNA markers (Mayer et al., 1998). Numbers at each node are bootstrap values, these indicate the percentage of trees constructed from random resampling of the sequence data that have the same branch structure. Genetic distance is an estimate of the proportion of substitutions in gene sequence between taxa.
reasons: (1) to understand the process and sequence of evolution, especially in the context of explosive radiations in the lakes of Africa (Meyer et al., 1994; Sturmbauer, 1998); and (2) to make comparisons among taxa, which may shed light on the adaptive context of their behaviours and morphology (Liem, 1973).

In this review we will attempt to summarize recent research into both aspects of cichlid phylogenetics. First we summarize the latest thinking about the relationships among cichlid taxa. Then we show how phylogenetic relationships have been used to answer biological questions, illustrated with some examples from published studies. A summary of some phylogenetic methods used in constructing evolutionary trees, and for testing hypotheses is described in Appendix I. We aim to show that even if one isn’t interested in the information about evolutionary relationships per se, it is impossible to test adaptive hypotheses about behaviour and morphology without it.

CICHLID RELATIONSHIPS

Since the early works of Pellegrin (1904) and Regan (1920), the systematics of Cichlidae, and their position within Perciformes, have undergone continuous revision and debate as new material and techniques have become available. Today, after decades of research there is some consensus over the relationships of many taxa, and researchers are producing ever more inclusive phylogenies, using more rigorous methods.


The molecular revolution has strongly augmented the traditional morphological approach to systematics. These newer methods have been able to resolve relationships that morphometrics could not.

For example, molecular research into cichlid phylogeny has resolved relationships for species in Lake Tanganyika (Sturmbauer et al., 1994; Nishida, 1997), the species flocks of Lakes Malawi and Victoria in East Africa (Meyer, 1993; Meyer et al., 1994; Mayer et al., 1998), tilapiine taxa (Sodsuk, 1993; Schliewen et al., 1994), and Neotropical taxa (van der Bank, 1994; Martin and Birmingham, 1998; Farias et al., 1998, 1999; Sides and Lydeard, 1999).

Inevitably, conflicts can arise between different methods and characters in the placement of taxa. For example, Stiassny (1991) placed *Heterochromis*, a west African riverine genus, as the sister group to the Indian etropelines, whereas Kullander (1998) placed it as basal to the Neotropical cichlids. Both relationships were based on morphological characters. Furthermore, a study of molecular data placed *Heterochromis* with the African *Hemichromis* (Farias et al., 1999). Such ambiguities may become resolved as more data become available. When two or more trees have been produced for the same group, the species they share can be used to combine or merge trees and provide a wider scheme of their relationships, known as ‘supertrees’ (or ‘total evidence tree’) (Sanderson et al., 1998) (see Appendix 1a). For example Kullander (1998) and Farias et al. (1999) present relationships for 41 and 48 cichlid genera respectively. If these could be combined they would include 57 genera. Unfortunately the conflicting results concerning relationships of the West African genus *Heterochromis* and the Neotropical genera *Cichla*, *Mikrogeophagus*, *Hoplochromus* and *Pterophyllum* prevent a combination method. In general, the relationships between West African and East Africa taxa from rivers and lakes require more attention. This will resolve how the ancestral riverine taxa dispersed to colonise rivers and lakes, perhaps on repeated occasions.

The Cichlidae are members of the suborder I. aequidens that also include the families Embiotocidae (surperches), Pomacentridae (damselfishes) and Labridae (wrasses) (Kaufman and Liem, 1982; Stiassny and Jensen, 1987). Kaufman and Liem (1982) combined Odacidae (cales and weed whittings) and Scaridae (parrotfishes) into the Labridae. However Nelson (1994) continues to recognize both families until more information is available. Labroid families are united by their pharyngeal jaw apparatus (PJA) (Kaufman and Liem, 1982; Stiassny and Jensen,
However, elements of the PJA are not unique to the Labroidei, which creates some doubt over whether labroid families have a common ancestor (Staatsny and Jensen, 1987; Johnson, 1993). Indeed, Streetman and Karl (1997) provide molecular evidence showing that other families from the order Perciformes fall within the Labroidei. Therefore, Cichlidae should not be considered the sister-group to the remaining Labroidei (Johnson, 1993; Nelson, 1994). Unfortunately, until more perciform families are sequenced we cannot state conclusively which families are most closely related to Cichlidae. Thus it is ironic that so much has been learnt about relationships within cichlids without having discovered where cichlids themselves have come from.

USES OF PHYLOGENIES
One may be more interested in the evolution of particular characteristics of animals than in the evolutionary relationships of the taxa themselves. Cichlid phylogenies have been used both to trace directional changes in characters through time (e.g. Goodwin et al., 1997), and also to deduce which taxa should be compared with each other in order to make evolutionary inferences about responses of traits to selection (e.g. Seehausen et al., 1999). Phylogenetic relationships are required for successful comparative analyses because closely related species share many characters which they have inherited from common ancestors (Harvey and Pagel, 1991). Recent comparative studies of other families of fishes have shown that phylogenetically-based comparisons can be more powerful at detecting relationships among characters than traditional cross-species studies (Jennings et al., 1998; Rickman et al. 2000). This is because paired comparisons among closely-related species control for many spurious traits shared by closely-related taxa, in the same way that a paired t-test is more powerful than an unpaired one.

For example, consider the evolution of different feeding modes (Fig. 2). One might be interested in whether body size differs between substrate-sifting fish and generalized predators. Treating the eleven substrate-sifting taxa and five generalized predators as independent data points might show some correlation between body size and feeding mode. However, the phylogenies show that substrate-sifting taxa has evolved from a generalized predator on only one or two independent occasions (Roe et al., 1997). These are the independent clades for comparison because the body size of taxa in the two substrate-sifting clades will reflect in part the body size of their common ancestor, and therefore cannot be considered independent.

Characters used in the study of evolutionary hypotheses can be either discrete or continuous. For each we provide a basic summary and two examples of published studies on cichlid character evolution. Appendix 1b gives further explanation of the methodology used in these examples.

(a) Discrete Traits Discrete characters are variables with binary (presence or absence) or multistate categories (e.g. different behaviours) and can be followed through time along the branches of a phylogenetic tree (Harvey and Pagel, 1991). A transition, or change in character state along a branch, might occur anywhere in the phylogeny. Unless there is good reason we make no a priori assumptions that characters evolve in a sequence, so characters can change between states, and a trait that is initially absent may evolve and later disappear. Further, we can test whether the evolutionary change in one trait is correlated with the change in a second trait, or if they evolve independently.

Figure 2. A reconstruction of the evolution of substratum-sifting for 18 species of Cichlasoma (Roe et al., 1997). The reconstruction is equivocal, requiring either (A) two independent transitions to substratum-sifting, and piscivores and detrivores evolving once each from the ancestral state of generalized predator, or (B) one transition to substratum-sifting and piscivory each from generalized predation, and from substratum-sifting, once to detrivory, and once to generalized predation.

Cichlasoma subgenera codes:

Am. = Amphiliopus; Ac. = Archocentrus;
As. = Astatheros; T. = Thorichthys; Th. = Theraps;
N. = Nandopsis; H. = Herichthys.

Journal of Aquaculture and Aquatic Sciences
Cichlid Research: State of the Art
Volume IX
Page 259
Journal of Aquaculture and Aquatic Sciences
Cichlid Research: State of the Art
Volume IX
Page 260
Figure 3. Phylogenetic distribution of the sex that provides care to eggs and/or juveniles in the family Cichlidae (Goodwin et al., 1998). Numbers in parentheses show the number of genera in the group. Taxa with (?) placed before them indicate a poor knowledge of phylogenetic relationships. Boxes next to taxa indicate that reproductive data were available. A indicates a single transition to female care that accounts for 59% of mouthbrooding genera. B indicates the equivocal clade that is resolved by the alternative positioning of the tribe Eretmodini, sister to the Lamprologini (marked with a star) (Kocher et al., 1995).

Example 1: Evolutionary transitions in parental care (Goodwin et al., 1998)
In cichlids there is striking diversity in two aspects of parental care: the form of parental care (substrate guarding and delayed and immediate mouthbrooding) and the sex that provides it (biparental, female-only and male-only). Our study used cichlids as a model group to test hypotheses for the evolution of these traits. One hypothesis is that both biparental care and substrate guarding are ancestral care in cichlids (Lowe-McConnell, 1959; Iles and Holden, 1969; Peters and Berns, 1982). Gittleman (1981) and Gross and Sargent (1985) hypothesised that female-only or male-only parental care would evolve from biparental care. A composite phylogeny for the family was constructed from several smaller published studies based on morphological and molecular methods. To follow the evolutionary transitions in parental care characters, information on the reproduction of extant species was collected, and using MacClade 3.0 (Maddison and Maddison, 1992) (see Appendix 1b) the ancestral states were reconstructed. The phylogenetic distribution of the sex of the care giver in the family Cichlidae (Fig. 3) shows that female-only care evolved from biparental care 21-30 times with 0-10 possible reversals. These transitions in parental care characters are the most numerous reported for any family of vertebrates, and support the hypothesis that biparental care gives rise to uniparental care. The study also shows that mouthbrooding evolved from substrate guarding 10-14 times with 0-3 reversals, and evolved in parallel with transitions in the parent providing care. Uniparental care has evolved many times in both substrate guarding and mouthbrooding clades, therefore it is unlikely that the form of care has constrained which sex provides it.

Example 2: Evolution of substratum sifting (Roe et al., 1997)
Roe et al. (1997) produced a systematic study of the relationships between 18 Cichlasoma species to test the adaptive radiation hypothesis, which implied that trophic specialization in cichlids arose independently from a generalized ancestor. Character tracing with MacClade (Maddison and Maddison, 1992) (see Appendix 1b) showed that either substratum sifting evolved independently twice (Fig. 2a), with piscivores and detritivores evolving once each from an ancestral generalized predator, or that substratum-sifting and piscivory evolved once each from a generalized predator, and detritivory and generalized predation each evolved once from substratum-sifting (Fig. 2b). Each scenario is equally parsimonious, requiring four steps. This study demonstrates that even with a small number of taxa where there is character diversity we can test evolutionary hypotheses with phylogenetic techniques.

(b) Continuous Traits. Continuous traits may include aspects of morphology, physiology, life histories and behaviour (Harvey and Pagel, 1991). Continuous data can be recoded so that discrete methods can be used (Maddison and Maddison, 1992), however this risks misinterpretation and methods designed specifically for these traits are more appropriate. Continuous traits can be can be plotted against each other to examine correlations, or we can examine the correlated evolution of a continuous trait with discrete characters (see Sanderson and Donoghue, 1996). In the example below Seehausen et al. (1999) included an analysis of discrete and continuous variables.
Example 3: Evolution of color patterns in East African cichlid fish (Seehausen et al., 1999)

It is generally recognized that the explosive speciation of African cichlids has been accompanied by extraordinary diversity in coloration and ecology. Seehausen et al. (1999) tested the hypothesis that sexual selection through mate choice has driven the evolution of color patterns. Data on flank coloration and striping were available for 107 taxa of East African cichlids. Data on mating system were used to measure the strength of sexual selection, and habitat and feeding information incorporated any ecological influences upon color and pattern.

Several methods of phylogenetic analysis were used, including phylogenetic regression (Grafen, 1989) (see Appendix 1b) to test for correlated evolution of markings. Ancestral character states were reconstructed using parsimony (Maddison and Maddison, 1992), and sister species contrasts were used to test whether characters evolved more quickly in promiscuous than in pair bonding/harem forming mating systems (see Appendix 1b).

Seehausen et al. (1999) showed that sexual selection is the most likely cause of the rapid evolution of nuptial coloration in fishes with a promiscuous mating system (a significant regression coefficient of 0.69). However, stripe patterns are constrained ecologically, with vertical bars and supraorbital stripes associated with complex habitats, such as rocky and vegetated lake shores and rivers (sig. regression coefficients range between 0.26 to 0.77). The evolution of the midlateral stripe was associated with piscivorous feeding, and longitudinal stripes with shoaling behaviour (sig. coefficients range between 0.19 to 0.3).

CONCLUSIONS

New information about phylogenetic relationships of cichlid fishes has been instrumental in tests of a variety of hypotheses about the evolution of morphology and behaviour. It should be remembered that these tests are based on relationships that are hypotheses themselves, and these are continually refined as new species are described, and when one uses new characters or methods of analysis. Therefore, it is inevitable that some statistical comparisons become invalid because they had been made among the wrong taxa. However, this may still be preferable to cross-species comparisons that ignore phylogenetic relationships altogether, because the latter inevitably contain inflated sample sizes (Harvey and Rambaut, 1998). Fortunately, new solutions for these problems are being developed. Many phylogeneticists now strive to use information from several genetic regions and sources (e.g. nuclear and ribosomal DNA), and models are currently being developed that weigh each molecule’s contribution to the tree structure appropriately. Another solution that has been proposed recently is the construction of a 'supertree' that agrees with all evolutionary scenarios (see Appendix 1a) (e.g. Sanderson et al., 1998; Seehausen et al., 1999).

Phylogenies are most useful for comparative research when they are well resolved and have included representatives from closely related taxa. Sampling programs are now becoming more ambitious by including taxa that were rarely previously sampled. We expect future cichlid research to adapt some of the successful comparative approaches embraced by recent studies on birds (Reynolds and Székely, 1997), reptiles (Shine 1994) and mammals (Purvis and Harvey, 1995). Systematic and comparative studies may also be important for future conservation issues (Solit and Gitzendanner, 1999). The wealth of information on cichlid systematics combined with the detailed research on their life histories will undoubtedly ensure that cichlids remain a model system for the study of adaptation and speciation at local and macroevolutionary scales.

ACKNOWLEDGEMENTS

This work was supported by a Biotechnology and Biological Sciences Research Council special studentship to N.B.G. and a Natural Environment Research Council grant to J. D. R. and S. B. (No. GR9/ 04061). We are grateful to Ron Coleman and David Earn for comments on the manuscript.

REFERENCES


Meyer, A., C. Montero, and A. Spreinat. 1994. Evolutionary history of the cichlid fish species flocks of the East African Great Lakes inferred from...


APPENDIX 1: METHODS

(a) Constructing Trees

An evolutionary tree or phylogeny is built as a visual representation of the evolutionary relationships among taxa. There are two popular methods for constructing phylogenetic trees. Parsimony builds relationships among taxa by minimising the number of changes in shared characters. Software packages such as PAUP ver 4.0b (Phylogenetic Analysis Using Parsimony) (Swofford, 1991) can produce a hypothesis of the relationships between taxa and some measure of confidence. Maximum likelihood is used mostly for sequence data and requires a model of evolution, with parameters that can be estimated from the data, with some assumptions (Felsenstein, 1981). The method then evaluates the likelihood for each possible phylogeny that the evolutionary model will fit the observed characters. The tree with the highest likelihood is considered the best estimate of the true phylogeny. Each method suits particular hypotheses of evolution, and since we do not know which evolutionary scenario is operating, the greatest confidence in our phylogenies comes when we obtain the same result by several methods (Kim, 1993). The University of California Museum of Paleontology, Berkeley maintains a comprehensive website that includes links to databases, software, publications, meetings and links to discussion groups and educational material on phylogenetic systematics (http://www.ucmp.berkeley.edu/subway/phylogen.html).

In most cases it is unlikely that a convenient phylogeny will always exist for all of the taxa we wish to study. When two or more trees are available which have some taxa in common we can combine datasets to generate a 'supertree', or 'total evidence tree', suitable for analyses. This method codes the topology of the trees, and when codes are combined to form a character matrix, a consensus is produced using parsimony (see above) (Sanderson et al., 1998).

(b) Testing Hypotheses of Character Evolution

Two computer-aided methods are popular for testing evolutionary hypotheses. Examples of published studies that have used these methods are given in the section 'Uses of Phylogenies' in this paper.

(1) MacClade ver 3.05 (Maddison and Maddison, 1992) is often used for tracing discrete characters along branches (e.g. Székely and Reynolds, 1995; Goodwin et al., 1998). If more than one route of character change is possible, the trace is equivocal and the program can display all the possible evolutionary routes. On small datasets MacClade can be used to test whether changes in one discrete trait are more concentrated than expected by chance in regions of the phylogeny (Maddison and Maddison, 1992; Lorch and Eadie, 1999).

(2) CAIC (Comparative Analysis for Independent Contrasts) (Purvis and Rambaut, 1995) is used to test the correlated evolution of two continuous characters, as well as the correlation of continuous data with discrete characters. Grafen's (1989) phylogenetic regression follows a similar method to CAIC. Purvis and Rambaut (1995) discuss how the methods differ. As well as testing for correlated evolution among characters, CAIC can be used to compare rates of evolution among clades or characters (Martins and Garland, 1991; Garland, 1992).