

## 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 (1999)

Niels Bouton

Section Integrative Zoology,  
Institute of Evolutionary and  
Ecological Sciences, University of  
Leiden, PO Box 9516, 2300 RA  
Leiden, the Netherlands. E-mail:  
bouton@rulsfb.leidenuniv.nl

### Abstract

I show that distribution patterns of rock-dwelling haplochromine cichlids in the south-eastern part of Lake Victoria can be explained by allopatric speciation alone. Observed patterns are consistent with a progressive invasion of different lineages from older (= deeper) parts of the lake, after which some lineages speciate more frequently than others. Based on a quantitative analysis of distribution data, Seehausen and van Alphen conclude that a high percentage of sympatric speciation (17.5%) is essential to explain the patterns. In my view this conclusion has its origins in an unjustifiable assumption. Seehausen and van Alphen assume that the unresolved phylogeny is balanced. However, allopatric speciation would typically result in imbalanced phylogenies.

### Keywords

Allopatric speciation, cichlids, distribution patterns, haplochromines, Lake Victoria, sympatric speciation, tree balance.

*Ecology Letters* (2000) 3: 166–171

Seehausen & van Alphen (1999) claim that distribution patterns of rock-dwelling haplochromine cichlids from Lake Victoria provide evidence for sympatric and allopatric speciation. However, it can be shown that allopatric speciation alone can also explain these distribution patterns.

Seehausen & van Alphen (1999) test the hypothesis that sympatric speciation by disruptive selection on polymorphic colouration has frequently occurred. They infer modes of speciation and their association with colour variation from patterns of geographical distribution in relation to phylogeny. They state that if dispersal was entirely unrestricted within the lake, all intralacustrine speciation would be sympatric. They continue that this is unlikely for rock-dwelling cichlids, because these are stenotopic and patches of soft or sand bottom often bar gene flow between their populations (Dorit 1990; van Oppen *et al.* 1997). Hence, they expect allopatric speciation in this group. They state that if sympatric speciation by disruptive sexual selection on polymorphic colouration contributed to the origin of species diversity in the group too, it should be traceable in the phylogeographic record: (a) the frequency distribution of the number of closely related species pairs plotted

against the degree of overlap of their geographical distribution should be bimodal, with modes on the fully allopatric and the fully sympatric; (b) the proportion of sister species pairs that are heteromorphic for one of the two colour polymorphisms associated with disruptive sexual selection should be higher in sympatric pairs than in allopatric pairs. As Seehausen & van Alphen (1999) point out, their second prediction is conclusive only if the first prediction is true. Alone the second prediction is inconclusive, because allopatric divergence followed by dispersal with selective local extinction, due to introgression of populations with the same body colouration, would result in similar distribution patterns.

In this paper we will reconsider the first prediction. Seehausen & van Alphen (1999) analyse the distribution patterns of 41 species of four genera over 47 localities. The four genera are *Pundamilia* (18 species), *Lithochromis* (10 species), *Neochromis* (10 species) and *Mbipia* (three species). The phylogeny within these genera is unresolved. Seehausen & van Alphen (1999) assume that each species is as good a candidate to be the closest relative of any other species and that the phylogenetic tree is balanced, i.e. nodes of the tree (points where branches meet) define subgroups of equal sizes. Balanced trees have

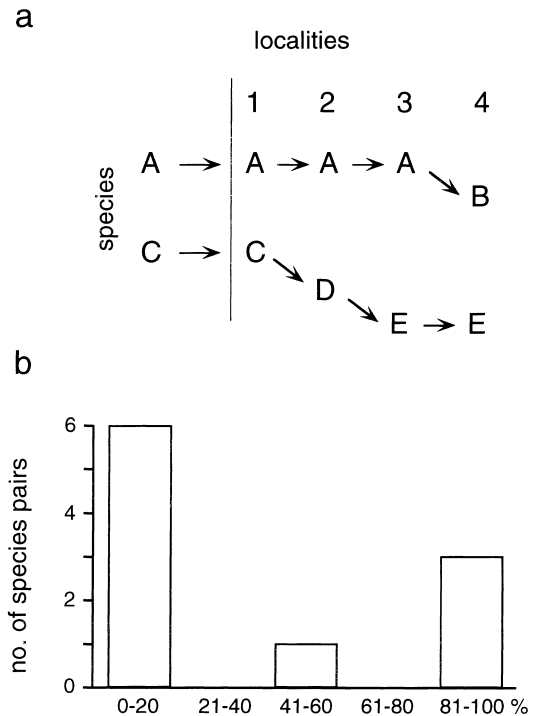
also been called symmetrical, while imbalanced trees have been called pectinate, comblike or “unbalanced” (Moore & Heard 1997).

Seehausen & van Alphen (1999) calculate the percentage of geographical overlap between all possible species pairs using the formula:

$$O_{ab} = \frac{P_{ab}}{\min(n_a, n_b)} \times 100$$

where  $O_{ab}$  is the percentage of geographical overlap between species a and b,  $P_{ab}$  is the number of localities where both species occur,  $n_a$  is the number of populations of species a, and  $\min(x, y)$  is the minimum of  $x$  and  $y$ . Percentage overlaps are calculated between all possible species pairs within a genus and classified into five categories: 0% up to and including 20%, 20%–40%, 40%–60%, 60%–80%, 80%–100%. Seehausen & van Alphen (1999) expected bimodality in the distribution of these categories as a result from allopatric (0%–20%) and sympatric (80%–100%) speciation. They consider three out of the four frequency distributions to be bimodal, namely those of *Pundamilia*, *Neochromis* and *Mbipia* (but not that of *Lithochromis*). They also show their second prediction to be true—colour polymorphisms associated with disruptive sexual selection occur more often in sympatric than in allopatric pairs of species—and conclude that sympatric speciation contributes to species richness of these genera.

The three genera under consideration consist of few wide-spread species and a much larger number of species that are present only at very few localities (see Appendix of Seehausen & van Alphen 1999). I will demonstrate that such a pattern may originate from allopatric speciation only. The sample localities are situated in two elongated shallow bays of Lake Victoria (one-time valleys), about half of the lake is deeper and therefore older. Presumably, these bays have been invaded by rock-dwelling cichlids during or after a lake level-rise. Invasion would follow the shore (the sides of the former valleys), which consist of rocky outcrops alternated with sandy or muddy bays. For rock-dwelling species this resembles a more or less linear chain of rocky islands. Invasion of such a chain is likely to be progressive, the last invaded outcrop is invaded from the last but one and so on. I hypothesize that from the older parts of the lake few lineages invade and speciate in allopatry at the rocky outcrops, some more frequently than others. If this scenario is worked out for two invading lineages (Fig. 1a), a geographical distribution pattern similar to that observed by Seehausen & van Alphen (1999) is obtained. The frequency distribution of geographical overlap between all possible species pairs is bimodal (Fig. 1b), which would indicate allopatric and sympatric speciation in the view of Seehausen & van Alphen (1999).



**Figure 1** (a) Schematic representation of progressive invasion and allopatric speciation. The left side represents the older part of the lake, the right side the research area. The arrows indicate the progressively invaded localities 1, 2, 3 and 4 by lineages A and C. Lineage A has speciated once, lineage C has speciated twice. (b) Frequency distribution of geographical overlap calculated from the research area in the example, species pairs A–B, B–C, B–D, C–D, C–E and D–E are allopatric ( $\leq 20\%$  overlap); species pairs A–C, A–D and B–E are sympatric ( $> 80\%$  overlap); species A and E have 50% overlap.

A consequence of the scenario outlined here (Fig. 1) is that many of the narrowly distributed species would be allopatric sister species. This can be tested by removing the wide-spread species from the original dataset and recalculating the frequency distribution of geographical overlap. If my hypothesis is correct this would result in unimodal frequency distributions with a mode on the allopatric category. On the other hand if bimodality of the pattern remains this would lend support for the conclusions of Seehausen & van Alphen (1999). After all, if many species with few populations live in sympatry, this could not be explained by the scenario I suggested, while it could be explained by sympatric speciation.

## METHODS

I calculate the frequency distributions of geographical overlap for the narrowly distributed species (1–6 populations) of the three genera for which Seehausen & van

Alphen (1999) presume a high percentage of sympatric speciation: 15 out of 18 species in *Pundamilia*, six out of 10 in *Neochromis* and two out of three in *Mbipia*. I compare these with the frequency distributions calculated from all species within the genera.

## RESULTS

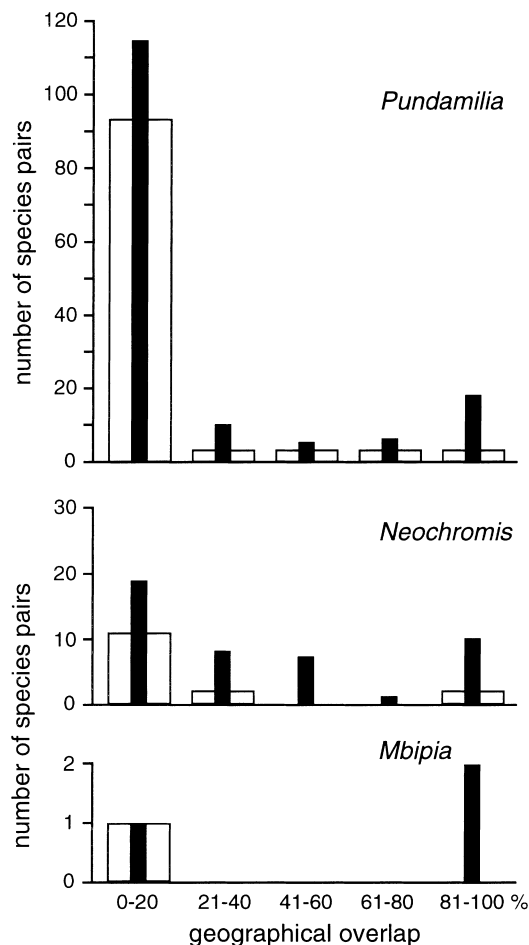
The frequency distributions of the geographical overlap based on the narrowly distributed species have a marked mode at the allopatric category, but not at the sympatric category (Fig. 2). This shows that a large number of the narrowly distributed species may be allopatric sister species. If the widely distributed species of the genera are added the bimodal patterns emerge (Fig. 2).

## DISCUSSION

Progressive invasion combined with allopatric speciation can explain the distribution patterns of rock-dwelling cichlids. Of congeneric species with 1–6 populations, a very low number lived in sympatry. Therefore, these species can be allopatric sister species. The probability of allopatric sister species on such short distances ( $\approx 500$ –2000 m) may seem small, but rock-dwelling haplochromine cichlids are stenotopic and lack a dispersal stage. Between populations of rock-dwelling haplochromine cichlids from Lakes Malawi (van Oppen *et al.* 1997) and Victoria (Dorit 1990) an exceptional fine-scale genetic structuring was found. Furthermore, it has been demonstrated that populations adapt their morphology to exploit local resource spectra (Bouton *et al.* 1999). The combination of a high degree of genetic isolation and local adaptation provides excellent opportunities for allopatric speciation.

I find the assumption of Seehausen & van Alphen (1999) that the phylogenetic tree should be balanced problematic, and this is the cause of the difference in interpretation of the dataset. The shape of a phylogenetic tree is influenced by many factors, one of which is mode of speciation (Mooers 1995). Progressive invasion of rows of islands or habitat islands in combination with allopatric speciation typically results in imbalanced phylogenetic trees (Mooers & Heard 1997). An example is the phylogenetic tree of *Tetragnatha* spiders, which progressively invaded the Hawaiian Islands (reproduced in Grant 1998). If conclusions about mode of speciation should be drawn from biogeographic data, it is unjustifiable to exclude beforehand phylogenetic trees that would result from one of these modes, namely allopatric speciation after progressive invasion.

In conclusion, either of two hypotheses, allopatric speciation after (or during) progressive invasion (this



**Figure 2** Recalculation of frequency distribution of geographical overlap for the narrowly distributed species only (□) and after inclusion of the widely distributed species (■).

manuscript), or a combination of sympatric and allopatric speciation (Seehausen & van Alphen 1999) can explain the distribution patterns of cichlid fishes in the south-eastern part of Lake Victoria. Resolving the phylogenetic tree may be the only method to confirm one of these hypotheses.

## ACKNOWLEDGEMENTS

I profited from discussions with Jurriaan de Groot and Ole Seehausen. Evert Meelis advised on statistical methods. Patsy Haccou, Frans Witte and two anonymous referees gave useful comments on earlier versions of the manuscript. Martin Brittijn is thanked for drawing the figures.

## REFERENCES

Bouton, N., Witte, F., van Alphen, J.J.M., Schenk, A. & Seehausen, O. (1999). Local adaptations in populations of

- rock-dwelling haplochromines (Pisces: Cichlidae) from southern Lake Victoria. *Proc. Roy. Soc. Lond. B*, 266, 355–360.
- Dorit, R.L. (1990). The correlates of high diversity in Lake Victoria haplochromine cichlids: a neontological perspective. In: *Causes of Evolution: a Paleontological Perspective* (eds Ross, R.M. & Allmon, W.D.). University of Chicago Press, Chicago, pp. 322–353.
- Grant, P.R. (1998). Epilogue and questions. In: *Evolution on Islands* (eds Grant, P.R.). Oxford University Press, Oxford, pp. 305–319.
- Mooers, A.O. (1995). Tree balance and tree completeness. *Evolution*, 49, 379–384.
- Mooers, A.O. & Heard, S.B. (1997). Inferring evolutionary process from phylogenetic tree shape. *Q. Rev. Biol.*, 72, 31–54.
- van Oppen, M.J.H., Turner, G.F., Rico, C., Deutsch, J.C., Ibrahim, K.M., Robinson, R.L. & Hewitt, G.M. (1997). Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proc. Roy. Soc. Lond. B*, 264, 1803–1812.
- Seehausen, O. & van Alphen, J.J.M. (1999). Can sympatric speciation by disruptive selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.*, 2, 262–271.

#### BIOSKETCH

Niels Bouton's main research interest is the process of adaptation, in particular in relation to ecological speciation. He recently finished a PhD on feeding ecology of rock-dwelling cichlids from Lake Victoria.

Editor, M. Hochberg  
 Manuscript received 22 December 1999  
 First decision made 24 January 2000  
 Manuscript accepted 9 March 2000

#### REPLY: INFERRING MODES OF SPECIATION FROM DISTRIBUTION PATTERNS

In a recent issue of *Ecology Letters* we presented distribution data for cichlid fish in Lake Victoria (presence/absence of 41 rock-shore-dwelling species of one clade distributed over 47 rocky headlands). We tested predictions from a model of sympatric speciation by disruptive sexual selection on specific and highly variable elements of colouration, evidence for which comes from theoretical (Turner & Burrows 1995; Van Doorn *et al.* 1998) and experimental behavioural and genetical work (McKaye 1991 and references therein; Seehausen 2000 and references therein). We showed that the frequency distribution of geographical range overlap between congeneric species has one mode on allopatric and a second mode on sympatric in three of four genera. We interpreted this as consistent with the hypothesis that alongside speciation by geographical isolation of populations, there is speciation

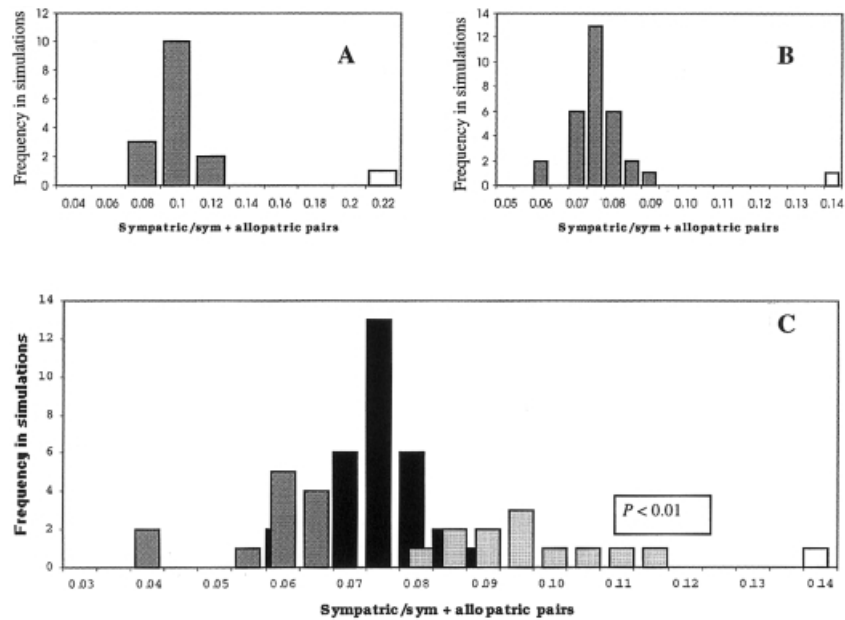
caused by the evolution of behavioural or other forms of isolation within populations. Bouton disagrees with this interpretation. Below we show that, contrary to his claims, (1) our assumptions are justified, (2) Bouton's allopatric-only model is not more parsimonious but requires substantial sympatric speciation, and (3) his proclaimed test of it is flawed.

#### THE ASSUMPTION ON PHYLOGENETIC TREE SHAPE

We made two assumptions when interpreting the observed distributions of range overlap: The speed of range extension (eg. dispersal rate) does not greatly differ between sister species, and the real but unknown internal phylogeny of monophyletic clades is not very imbalanced. Bouton argues that the latter assumption is unjustified because allopatric speciation would lead to imbalanced phylogenies. This is correct only if geographical discontinuities arise sequentially in one direction and in a spatially linear fashion. Changing water levels in a lake, by creating new habitat and breaking up old habitat ranges, make discontinuities arise in both directions even in a linear island chain. Simulating between two and six allopatric speciation events over 50 islands 30 times (Fig. 1), we obtained a mean Colless' imbalance score  $I_m$  of  $0.69 \pm 0.30$ , where maximum possible balance was 0.33 (odd number of tree tips) and complete imbalance is 1. Hence, our assumption is perfectly justified even if islands are linearly aligned, but it does not appear crucial to interpretation of the data anyway. Our two assumptions can be collapsed into one: the widely distributed species which cause most of the sympatry should not come off terminal nodes in the tree. Distribution range size should positively correlate with age of species. Phrased differently, recently evolved species should not undergo disproportionately fast range extension. We believe this is a parsimonious assumption.

#### THE ALLOPATRIC-ONLY MODEL

Bouton argues that allopatric speciation along an island chain alone would produce distributions of overlaps similar to those observed if two instead of one founder species are assumed for each genus. Many unrelated species, generated by different founders, would then be sympatric. He suggests that this hypothesis can be tested by calculating, from our data matrix, overlap frequencies exclusively among species that occur on six or fewer islands, and that absence of a mode on the sympatric side among these is evidence against sympatric speciation. The probability that sympatric speciation occurs in a species is directly proportional to the number of its populations, whereas the probability of allopatric speciation is propor-



**Figure 1** Frequency distributions of sympatry values ( $n$  sympatric species pairs/[ $n$  sympatric +  $n$  allopatric species pairs]) obtained in simulations of allopatric speciation with two founder lineages. Geographical space is represented by one dimension as a chain of 50 islands. Both founder species are initially globally distributed (this maximizes sympatry). Geographical discontinuities do then arise randomly between any two adjacent islands, and any species whose range is intercepted by a discontinuity speciates across it with a probability of 0.5. If speciation occurs, the entire population between the new discontinuity and an adjacent already existing discontinuity or the end of the island chain, is transformed into a new species. (A) Simulation runs were stopped when 10 species were generated (15 simulations in species per lineage varied between 4 and 6). (B) Simulation runs were stopped when 18 species were generated (30 simulations in species per lineage varied between 6 and 12). Range overlaps between any two species are then calculated. The observed values for the genera *Neochromis* (10 species on 47 islands) and *Pundamilia* (18 species on 47 islands) are indicated by a white bar. (C) Same as (B), but here the effects of range alterations after speciation were simulated for the extremes of the primary frequency distribution. This was done by using the distribution matrices that generated the two highest and the two lowest sympatry values, and letting each undergo five expansion/extinction cycles, where each cycle consisted of 10 randomly assigned range expansions (adding new populations at the range edge) and 10 randomly assigned extinctions of range edge populations. The distributions of sympatry values from each cycle are plotted (dark grey for the low sympatry extremes, light grey for the high sympatry extremes, black in the background the values of the primary pattern). Expansion/extinction cycles cause fluctuations around the original sympatry value. The probability of obtaining the highest of the observed sympatry values is indicated.

tional only to the number of populations at the margin of its distribution range. In Bouton's linear island chain model, the probability of allopatric speciation is exactly the same in all species, independent of the number of populations. Species range sizes in our data set vary between one and 35 islands. If Bouton's speciation model applies, omitting all ranges over six islands reduces the relative probability of retaining cases of sympatric speciation by a factor of up to 5.8 (linear island chain case). Considering this, the fact that there are sympatric pairs in two of the three genera even among these narrow endemics is support for sympatric speciation.

To test Bouton's hypothesis we have simulated purely allopatric speciation by a simple model that describes Bouton's verbal island chain model: initially two species distribute themselves over a chain of 50 islands. They, and subsequently also any of their daughter species, split up randomly across randomly arising geographical disconti-

nities (which is technically the same as having new islands arise that can be colonized by either species). When the number of species has reached the level that we observed in *Pundamilia* (18) and *Neochromis* (10), the frequencies of geographical range overlaps are calculated. In both cases, the observed sympatry value lies outside the tail of the distributions generated by simulations, suggesting that allopatric speciation starting with two founders does not parsimoniously explain the observed patterns (Fig. 1). It appears that a minimum of three founders would be required both for *Neochromis* and for *Pundamilia* to produce the observed sympatry values.

## THE PALAEOLOGICAL EVIDENCE

The real question is not how many lineages invaded the surveyed area but whether range overlap distributions observed in the area are representative for those in the

lake as a whole. Lake Victoria is saucer-shaped with a maximum depth of 69 m. It had been dry until 12500 years ago and has subsequently filled up rapidly (Johnson *et al.* 1996). The clade of cichlids that we studied appears to be monophyletic (Seehausen *et al.* 1998), possibly as is the entire species flock (Meyer *et al.* 1990). Explaining high intrageneric sympatry in a surveyed area by postulating that a substantial number of species invaded this area and underwent allopatric speciation there, is evading the issue of how these founder species originated. The surveyed area is a large chunk of the southern lake, with depths of between 4 and 50 m in the immediate vicinity (< 100 m) of the sampled rocky shores. We have good reason to believe that the distribution patterns observed in the surveyed area are representative for the entire lake. All four genera, and seven of the species that are widely distributed within that area, are widely distributed also outside the area, effectively having lake-wide distributions being recorded also from the northern shores of the lake (L. Kaufman, personal communication, personal observation). In the surveyed area the genera overlap fully (except for ecological reasons) and most of the closely related widely distributed species overlap largely (83%, 18%, 9% within *Neochromis*, 96%, 87%, 67% within *Pundamilia*, 88% within *Mbipia*). There is no evidence that they overlap less in the rest of the lake. They must have split after the lake began to fill up and before the surveyed area was flooded, hence in a proto-lake that would have been no deeper than 15 m on average. The potential for allopatric speciation must then have been much smaller than it is now, consistent with the high sympatry values observed between most of the widely distributed species. Four of the five largely overlapping pairs of these species differ in colouration in the way our sympatric speciation model predicts. All this is in line with our argument that properties that cause disruptive selection and sympatric speciation are likely to elevate speciation rates particularly in an early stage of lake and species flock formation, when distances between habitat patches are small, species are unspecialized and populations are panmictic. A property that causes sympatric speciation in that situation could jump-start an evolutionary radiation.

## CONCLUSION

At the end of the day progressive invasion with allopatric speciation is not suitable to explain the distribution patterns observed among rock-dwelling cichlids in Lake Victoria without making additional assumptions. A combination of sympatric and allopatric speciation is, hence, the more parsimonious explanation.

## ACKNOWLEDGEMENTS

We thank Mairi Knight and Andy Davey for constructive comments on the MS and George Turner for discussion.

OleSeehausen<sup>1,2</sup> and Jacques J. M. van Alphen<sup>2</sup>  
<sup>1</sup>*Biodiversity and Ecology Research Division, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, UK, E-mail: os@soton.ac*  
<sup>2</sup>*Institute of Evolutionary and Ecological Sciences, University of Leiden, PO Box 9516, 2300 RA Leiden, The Netherlands*

## REFERENCES

- Johnson, T.C., Scholz, C.A., Talbot, M.R., Kelts, K., Ricketts, R.D., Ngobi, G., Beuning, K., Ssemmanda, I. & McGill, J.W. (1996). Late pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science*, 273, 1091–1093.
- McKaye, K.R. (1991). Sexual selection and the evolution of the cichlid fishes of Lake Malawi, Africa. In: M.H.A., Keenleyside, ed. *Cichlid Fishes. Behaviour, Ecology and Evolution*. Chapman & Hall. London, pp. 241–257.
- Meyer, A., Kocher, T.D., Basasibwaki, P. & Wilson, A.C. (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, 347, 550–553.
- Seehausen, O. (2000). Explosive speciation rates and unusual species richness in haplochromine cichlid fishes: effects of sexual selection. *Adv. Ecol. Res.*, 30, in press.
- Seehausen, O., Lippitsch, E., Bouton, N. & Zwennes, H. (1998). Mbipi, the rock-dwelling cichlids of Lake Victoria: description of three new genera and 15 new species (Teleostei). *Ichthyol Exploration Freshwaters*, 9, 129–228.
- Turner, G.F. & Burrows, M.T. (1995). A model of sympatric speciation by sexual selection. *Proc. Royal Soc. London*, B260, 287–292.
- Van Doorn, G.S., Noest, A.J. & Hogeweg, P. (1998). Sympatric speciation and extinction driven by environment dependent sexual selection. *Proc. Royal Soc. London*, B265, 1915–1919.