

# Case studies and mathematical models of ecological speciation. 1. Cichlids in a crater lake

SERGEY GAVRILETS,\*† AARON VOSE,\*‡ MARTA BARLUENGA,§¶ WALTER SALZBURGER§¶ and AXEL MEYER§

Departments of \*Ecology and Evolutionary Biology, †Mathematics, and ‡Computer Science, University of Tennessee, Knoxville, TN 37996, USA, §Department of Biology, University of Konstanz, D-78457 Konstanz, Germany, ¶Department of Ecology and Evolution, University of Lausanne, UNIL Sorge, Le Biophore, CH-1015 Lausanne, Switzerland

## Abstract

A recent study of a pair of sympatric species of cichlids in Lake Apoyo in Nicaragua is viewed as providing probably one of the most convincing examples of sympatric speciation to date. Here, we describe and study a stochastic, individual-based, explicit genetic model tailored for this cichlid system. Our results show that relatively rapid (< 20 000 generations) colonization of a new ecological niche and (sympatric or parapatric) speciation via local adaptation and divergence in habitat and mating preferences are theoretically plausible if: (i) the number of loci underlying the traits controlling local adaptation, and habitat and mating preferences is small; (ii) the strength of selection for local adaptation is intermediate; (iii) the carrying capacity of the population is intermediate; and (iv) the effects of the loci influencing nonrandom mating are strong. We discuss patterns and timescales of ecological speciation identified by our model, and we highlight important parameters and features that need to be studied empirically to provide information that can be used to improve the biological realism and power of mathematical models of ecological speciation.

*Keywords:* ecological, mathematical, model, parapatric, speciation, sympatric

*Received 24 October 2006; revision accepted 22 January 2007*

## Introduction

Recent years have seen significant advances in speciation research (e.g. Howard & Berlocher 1998; Coyne & Orr 2004; Dieckmann *et al.* 2004; Gavrilets 2004). As our understanding of the processes leading to the origin of new species increases, we appreciate more and more the importance of the insight of the 20th century giants of speciation research that 'speciation can occur in different ways' (Dobzhansky *et al.* 1977) and that 'there are multiple answers to every aspect of speciation' (Mayr 1982).

Here, we are concerned with one particular way in which speciation can occur, *ecological speciation* (e.g. Mayr 1947; Schluter 2000; Rundle & Nosil 2005), which usually refers to speciation driven by ecologically-based divergent selection. Selection is divergent when it favours opposite phenotypes within different populations or the same

population. Selection is ecological when it arises as a consequence of the interaction of individuals with their abiotic and biotic environment during resource acquisition.

Ecological speciation emphasizes the importance of ecological selection over other evolutionary factors, such as geographical isolation. Consequently, ecological speciation is often discussed within the context of sympatric speciation. The latter is defined as the emergence of new species from a population where mating is random with respect to the birthplace of the mating partners (Gavrilets 2003, 2004). During sympatric speciation mating may be nonrandom with respect to, for example, genotype, phenotype and culturally-inherited traits.

Both sympatric and, to a lesser degree, ecological speciation have traditionally been controversial. These controversies have attracted the attention of many theoreticians; and by now the great majority of theoretical work on speciation concerns speciation in the presence of gene flow between diverging populations, driven by ecological selection (Kirkpatrick & Ravigné 2002; Gavrilets 2004). Most of this

Correspondence: Sergey Gavrilets, Fax: +1 865 974 3067; E-mail: sergey@tiem.utk.edu

work is represented by numerical studies but there now exist more than ten simple analytical models of sympatric speciation (Gavrilets & Waxman 2002; Gavrilets 2003, 2004, 2006). The theory of sympatric speciation is arguably the most developed part of theoretical speciation research. The general conditions for sympatric speciation as identified by theoretical research are: (i) strong combined effects of disruptive selection and nonrandom mating; (ii) strong association of the genes controlling traits subject to selection, and those underlying nonrandom mating; (iii) high levels of genetic variation; and (iv) the absence of costs from being choosy (Gavrilets 2004). Two of the most straightforward ways for sympatric speciation are provided by a 'magic trait' mechanism and a habitat selection mechanism. The former describes situations when there is a trait that is both subject to disruptive/divergent selection and simultaneously controls nonrandom mating (such as size in stickleback fish or colour in *Heliconius* butterflies). The latter corresponds to situations when organisms evolve stronger and stronger preferences for specific habitats with mating-pair formation and/or mating taking place within preferred habitats.

Mathematical models clearly show that, under certain biologically reasonable conditions, sympatric speciation is possible (Gavrilets 2004). Why then are there so few cases (reviewed in Coyne & Orr 2004) where sympatric speciation is strongly implicated in spite of the enormous interest in sympatric speciation and strong motivation to find examples? One explanation is that it is difficult to prove or to rule out alternative scenarios. Another possibility is that conditions for sympatric speciation as identified by mathematical models are rarely satisfied in natural populations. Incorporating theoretical insights into empirical work, bridging existing theory and data, and applying mathematical models to particular case studies are all crucial steps towards assessing the importance of ecological and sympatric speciation in nature.

In their review of empirical work on sympatric speciation, Coyne & Orr (2004) suggested a number of empirical criteria/tests for sympatric speciation. Recent studies (e.g. Barluenga *et al.* 2006a; Savolainen *et al.* 2006) that have attempted to apply the criteria set forth by Coyne & Orr (2004) have invigorated the interest and controversy surrounding sympatric speciation. The general goal of this and the accompanying paper (Gavrilets & Vose 2007) is to use mathematical models tailored to particular case studies to address certain questions about nonallopatric speciation in general (e.g. whether sympatric speciation is achieved easily as some claim) and in particular case studies (e.g. whether an observed pattern is a result of *in situ* speciation or double invasion, whether speciation was truly sympatric or parapatric). We will also look at specific questions, such as what does mathematical theory tell us about the plausibility/speed/patterns of (sympatric)

speciation in the case studies? What are important parameters and processes controlling the dynamics of speciation? How common are the phenomena observed in these case studies? We will also attempt to identify important parameters and features that need to be studied empirically to provide information that can be used to improve the biological realism and power of mathematical models of ecological speciation.

Here, we look at the case of cichlids in a crater lake described by Barluenga *et al.* (2006a) while the accompanying paper considers the case of palms on an oceanic island described by Savolainen *et al.* (2006).

### Empirical evidence

The study of a pair of sympatric species in Lake Apoyo in Nicaragua published by Barluenga *et al.* (2006a) (see also McKaye *et al.* 2002; Barluenga & Meyer 2004) is viewed as providing probably one of the most convincing examples of sympatric speciation to date. In describing this system we will closely follow Barluenga *et al.* (2006a); see Wilson *et al.* (2000) for some earlier work on cichlids in this area.

Crater Lake Apoyo in Nicaragua is a relatively small, filled caldera (diameter ~5 km; max. depth ~200 m) of recent volcanic origin (less than 23 000 years old). The crater lake is completely isolated from surrounding lakes, and has a homogeneous habitat. Its fauna is impoverished with respect to nearby water bodies, but contains at least one endemic cichlid fish species, the Arrow cichlid, *Amphilophus zalius*, as well as the widespread Midas cichlid, *A. citrinellus*.

Analysis of the mitochondrial DNA (mtDNA) control region showed that the two *Amphilophus* species from Lake Apoyo form a monophyletic assemblage. Notably, not a single mtDNA haplotype is shared between Lake Apoyo's *Amphilophus* species and the faunas of other water bodies in Central America. This suggests that there was a single colonization event from an ancestral lineage of *A. citrinellus*, making contemporary genetic exchange or secondary colonization highly unlikely. Equally unlikely is the extinction of a previously widespread *A. zalius* species everywhere but in Lake Apoyo. The genetic diversity and number of mtDNA haplotypes are low in Lake Apoyo, which is in-line with the young age of the species assemblage. Demographical analyses recovered a single expansion of *A. citrinellus* in Lake Apoyo about 20 000 years ago, while a more recent expansion was detected for the new species *A. zalius*. The analysis of microsatellite allele composition and phylogenies based on microsatellites and amplified fragment length polymorphisms (AFLPs) corroborates the distinctive genetic composition of Lake Apoyo's *Amphilophus* fauna, and is consistent with the evolution of *A. zalius* from Lake Apoyo's *A. citrinellus* stock.

Microsatellite-based Bayesian population assignment tests unambiguously identified two clusters within Lake Apoyo corresponding to *A. citrinellus* and *A. zalius*, which are clearly distinct from any other *Amphilophus* population in the area. Significant *F*-statistics based on all genetic markers suggest that *A. citrinellus* and *A. zalius* are reproductively isolated in Lake Apoyo. This is corroborated by mate choice experiments that demonstrated strong assortative mating in both species (Baylis 1976). However, there is no within-species genetic differentiation for either species.

The two *Amphilophus* species in Lake Apoyo are morphologically distinct. [For a detailed analysis on the morphometric differences in the *A. citrinellus* species complex see Meyer (1989, 1990a, b); Klingenberg *et al.* (2002, 2003).] A morphometric analysis uncovered two discrete body types corresponding to the two species, with body length and height explaining most of the differences. The two species also differ in the shape of a trophically-relevant structure that is tightly linked with the ability of cichlids to process alternative food types: the pharyngeal jaw. *A. zalius* has more elongated 'papilliform' pharyngeal jaws, while those of *A. citrinellus* are more variable and wider. In addition, the significant differences in the stomach contents show significant dietary differences between the two species. The trophic niche of *A. citrinellus* is wider than that of *A. zalius*. Ecomorphological inferences thus characterize *A. citrinellus* as a benthic forager with a deeper body, while *A. zalius* is a more specialized limnetic form with an elongated trunk that appears to be adapted to living in the open water column.

According to Barluenga *et al.* (2006a) the recent volcanic origin of Lake Apoyo, its small size, its degree of isolation, the homogeneous habitat, the sympatric occurrence of both species throughout the lake and the absence of genetic structure within each of the two species collected from different parts of the lake rule out the possibility of (micro) allopatric or parapatric differentiation. Barluenga *et al.* (2006a) argue that sympatric speciation is the most plausible scenario for explaining the origin of a new cichlid species from a more widespread ancestral species in Lake Apoyo in less than 10 000 years after the lake was initially colonized. Barluenga *et al.* (2006a) suggested that sympatric speciation was driven by ecological mechanisms through divergent habitat preferences and resource partitioning, and by assortative mating through behavioural isolation.

We note that Schliewen *et al.* (2006) questioned this scenario and proposed an alternative scenario of multiple colonization and hybridization. However Barluenga *et al.* (2006b) believe the alternative advanced by Schliewen *et al.* (2006) is considerably less parsimonious, contains some inconsistencies and is incompatible with the available evidence.

## Mathematical model

We will use a generalization of the model in Gavrillets & Vose (2005), which in turn generalizes and extends those in Diehl & Bush (1989), Johnson *et al.* (1996), and Fry (2003); see also Kawecki 1996, 1997. The most important extension of our previous model concerns the incorporation of three additional traits controlling (nonrandom) mating. We will use this model to study the process of invasion by a specialist species into an environment where a new ecological niche is available (Kawata 2002; Gavrillets & Vose 2005). Most modeling work on evolution in a spatially heterogeneous environment assumes *soft selection* (e.g. Kisdi & Geritz 1999; Spichtig & Kawecki 2004 and references therein; but see DeMeeus *et al.* 1993) and does not consider population densities explicitly. In contrast, selection for local adaptation in our model is both density-dependent and *hard* (*sensu* Christiansen 1975). That is, the contribution of each niche to offspring depends on the fitness of individuals in the niche. The following describes the major components of the model.

### Space and environment

In our model, space is represented by an array of 'patches' each of which can support a population of a certain size. Here, we study models with just two patches arranged one on top of another and models with a large number of patches arranged in two circles laying one on top of another. The top patches will represent the near-shore habitat and the bottom patches deep-water habitat. In the case of just two patches, offspring will disperse among both of them. When observed, speciation in these models will be sympatric. In the case with many patches, offspring will disperse across the patch of origin and five neighbouring patches (two in the same row and three in the other row directly above or below). When observed, speciation in these models will, in the strict sense, be parapatric (because of the isolation by distance along the lake shore).

### Individuals

Generations are discrete and nonoverlapping. Individuals are diploid and have discrete sexes. Each individual has a number of additive quantitative characters:

- an 'ecological' character  $x$ ;
- a 'habitat preference' character  $y$ ; and
- three 'mating compatibility' characters  $m$ ,  $f$ , and  $c$ .

Following previous models (e.g. Dieckmann & Doebeli 1999; Bolnick 2004, 2006; Doebeli *et al.* 2005) we assume that the male display trait,  $m$ , is expressed in males only, whereas female mating preference traits,  $f$  and  $c$ , are expressed in females only. All these traits are scaled to be

between 0 and 1 and are controlled by different unlinked diallelic loci with equal effects. Mutations occur at equal rates across all loci; the probabilities of forward and backward mutations are equal. In addition, there are a number of unlinked neutral loci with a large number of alleles subject to step-wise mutation (Ohta & Kimura 1973). These loci have higher mutation rates and will be used to evaluate the levels of genetic divergence within and between species that one would observe if using microsatellite markers.

### Life-cycle

The life-cycle consists of: (i) preferential dispersal of offspring among neighbouring patches (including the patch of origin); (ii) density-dependent viability selection within the patch; and (iii) nonrandom mating among individuals within the patch and offspring production. Note that our description of the third stage implies that we assume that mating pairs are formed in the feeding habitat.

### Habitat preference

The relative preference of an individual with habitat preference trait,  $y$ , for the deep-water habitat is given by a linear function of  $y$ :

$$p = \frac{1}{2} + a \left( y - \frac{1}{2} \right). \quad (1)$$

The relative preference of this individual for the near-shore habitat is  $1 - p$ . Parameter  $0 \leq a \leq 1$  measures the maximum possible preference. The relative preference for the deep-water niche changes linearly from  $(1 - a)/2$  for  $y = 0$  to  $(1 + a)/2$  for  $y = 1$ . The value of  $1 - a$  can be interpreted as the probability that an individual with the highest preference for one habitat mistakenly goes to the other habitat.

### Viability selection

Fitness of an individual with ecological trait  $x$  is:

$$w = \exp \left[ - \frac{(x - \theta)^2}{2\sigma_s^2} \right], \quad (2)$$

where  $\theta$  is the optimum phenotype, which we set to 0 in the near-shore habitat and 1 in the deep-water habitat, and  $\sigma_s$  is a parameter measuring the strength of selection. Smaller values of  $\sigma_s$  mean stronger selection. The carrying capacity associated with the phenotype is  $K = K_0 w$ , where  $K_0$  is the maximum carrying capacity. The probability that an individual survives to the age of reproduction is given by the Beverton–Holt model (e.g. Kot 2001):

$$v = \frac{1}{1 + (b - 1) \frac{N}{K}}, \quad (3)$$

**Table 1** Fitness  $w_{spec}$  of a specialist for one niche in the other niche and fitness  $w_{gen}$  of a generalist in both niches for different value of parameter  $\sigma_s$  characterizing the strength of selection for local adaptation. Larger values of  $\sigma_s$  imply larger fitnesses  $w_{spec}$  and  $w_{gen}$  and, thus, weaker selection for local adaptation

$\sigma_s$	$w_{spec}$	$w_{gen}$
0.35	0.017	0.36
0.4	0.044	0.458
0.45	0.085	0.539
0.5	0.135	0.607

where  $b > 0$  is a parameter (the average number of offspring per female; see below), and  $N$  is the number of juveniles in the patch. This function describes density-dependent viability selection. Parameter  $\sigma_s$  will play an important role in our simulations. Note that fitness of a ‘specialist’ (i.e. a genotype perfectly adapted to one niche) in the other niche is  $w_{spec} = \exp[-1/(2\sigma_s^2)]$ . Fitness of a ‘generalist’ (i.e. an individual with  $x = 1/2$ ) is  $w_{gen} = \exp[-1/(8\sigma_s^2)]$  in both niches. Table 1 will help one to understand the biological meaning of the numerical values of  $\sigma_s$  we use. These values were chosen on the basis of biological intuition and to allow for different outcomes of the evolutionary dynamics.

### Mating preferences

The relative probability of mating between a female with traits  $f$  and  $c$  and a male with trait  $m$  is:

$$\psi(m, f, c) = \begin{cases} \exp \left[ -(2c - 1)^2 \frac{(f - m)^2}{2\sigma_a^2} \right], & \text{if } c > 0.5, \\ 1, & \text{if } c = 0.5, \\ \exp \left[ -(2c - 1)^2 \frac{(f - (1 - m))^2}{2\sigma_a^2} \right], & \text{if } c < 0.5, \end{cases} \quad (4)$$

where parameter  $\sigma_a$  scales the strength of female mating preferences. Under this parameterization, females with  $c = 0.5$  mate randomly, females with  $c > 0.5$  prefer males whose trait  $m$  is close to the female’s trait  $f$  (positive assortative mating), and females with  $c < 0.5$  prefer males whose trait  $m$  is close to  $1 - f$  (negative assortative mating). This parameterization is a correction of the one used by Bolnick (2004, 2006) and Doebeli (2005), which introduced some artefacts (e.g. strong disruptive sexual selection; see Appendix 1 for discussion). Note that the absolute value  $|2c - 1|$ , which we will denote as  $C$ , characterizes the extent of deviation of the female’s mate choice from random: females with  $C = 0$  mate randomly while those with  $C = 1$  exhibit the strongest possible (negative or positive) assortative mating. Each mating results in a number of offspring drawn

**Table 2** Relative preference of a female deviating from random mating by a single substitution in a locus controlling trait  $c$  for males deviating from her most preferred type by a single substitution, for two different values of parameter  $\sigma_a$  and the number of loci per trait  $L$

$\sigma_a$	$L = 4$	$L = 8$
0.05	0.044	0.823
0.1	0.458	0.953

from a Poisson distribution with parameter  $b$ . We assume that all adult females mate. This assumption implies that any costs of mate choice, which can easily prevent divergence and speciation (Pomiankowski 1987; Bolnick 2004; Gavrillets 2004, 2005; Gourbiere 2004; Kirkpatrick & Nuismer 2004; Waxman & Gavrillets 2005a), are absent. This assumption also means that the effective population size is increased relative to the actual number of adults (Gavrillets & Vose 2005). Parameter  $\sigma_a$  characterizing the efficiency of female mate choice will play an important role in our simulations. The biological meaning of  $\sigma_a$  is clarified in Table 2. This table shows the preference  $\psi(m, f, c)$  of a female deviating from random mating by a single allelic substitution in a locus controlling female trait  $c$  (so that  $|c - 0.5| = 1/(2L)$ ) for males deviating from her most preferred type by a single allelic substitution (so that  $|f - m| = 1/(2L)$ ) or that  $|f - (1 - m)| = 1/(2L)$ ) for different numbers  $L$  of loci per trait. Note that small  $\sigma_a$  means that the mating preference loci have extremely strong effects on the probability of mating. For example, with  $\sigma_a = 0.05$  and  $L = 4$ , the corresponding preference is reduced by a factor of 22.

*Initial conditions*

The founding population is represented by a small number of individuals who are perfectly adapted (i.e.  $x = 0$ ) and have the highest preference (i.e.  $y = 0$ ) for the near-shore niche (i.e. with  $\theta = 0$ ). The mating characters are set exactly at the middle of the range of possible values (i.e.  $m, f, c = 0.5$ ) with no genetic variation present, so that all individuals mate randomly. Each microsatellite locus was heterozygous with two intermediate alleles out of  $2^8$  possible alleles.

*Parameters*

To analyze our models, we used individual-based simulations, which we ran for 20 000 generations, corresponding to 20 000 years. The results below are based on a number of runs, between 10 and 24, for each parameter combination. The following summarizes parameters that did not change: the number of founders was 10; mutation rate  $10^{-5}$  per generation for loci controlling traits  $x, y, m, f$  and  $c$ ; mutation rate  $10^{-3}$  per generation for ‘microsatellites’; the average

number of offspring per female  $b = 16$  (e.g. McKaye & Barlow 1976); and the number of ‘microsatellites’ was 8. All loci were unlinked. The number of loci per traits  $x, y, m, f$  and  $c$  was the same and denoted as  $L$  below. The population-level data were saved every 100th generation and the individual-level data were saved every 1000th generation.

The model was implemented in C. The code is available upon request.

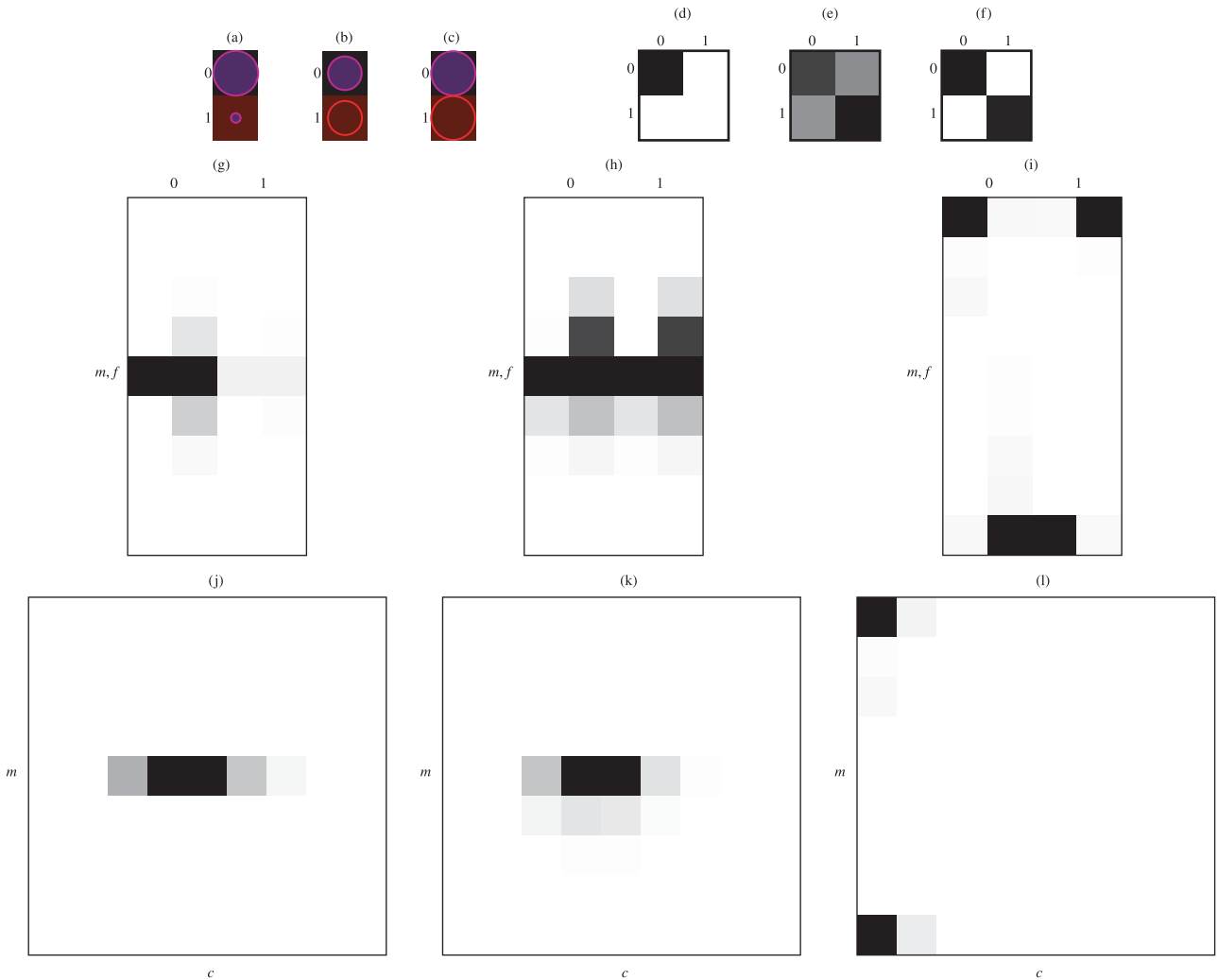
**Theoretical results**

In our simulations, we have observed three different evolutionary outcomes: (i) failure to adapt to the deep-water habitat; (ii) evolution of a generalist species utilizing both habitats relatively inefficiently; and (iii) emergence of a new species specializing on and strongly preferring the new habitat, which occurs simultaneously with strong differentiation in mating preferences between the two species. The fourth theoretically possible outcome, speciation via the evolution of strong habitat specialization and strong habitat preferences with no evolution of mating preferences (as studied by Gavrillets & Vose 2005), was not observed. In describing these states, we will focus on the distribution of individuals across space, levels of local adaptation, correlations between habitat preference and fitness, levels of genetic variation in mating characters and the degree of nonrandom mating.

Figure 1 illustrates how three different outcomes are characterized graphically in our simulations of  $2 \times 1$  systems. In the first regime, illustrated in Fig. 1(a, d, g, j), the number of individuals using the deep-water niche is relatively small (Fig. 1a), and all individuals have the highest preference for the near-shore niche and highest viability in the near-shore niche (Fig. 1d). Some genetic variation in male and female mating characters is maintained by a balance between mutation and random drift (Fig. 1g, j). Some individuals may exhibit nonrandom mating (Fig. 1g, j) which evolves by mutation and random genetic drift.

In the second regime, illustrated in Fig. 1(b, e, h, k), the individuals are present in both niches but at densities that are smaller than the maximum carrying capacity  $K_0$  (Fig. 1b). Some geographical differentiation between the two niches may be observed so that some correlation between the most preferred niche and the niche providing the best fitness may be present (Fig. 1e). Some genetic variation in male and female mating characters (Fig. 1h, k) and some nonrandom mating are present but not structured between the subpopulations (Fig. 1h, k).

In the third regime, illustrated in Fig. 1(c, f, i, l), both niches are filled to carrying capacities (Fig. 1c). There is strong differentiation in habitat preferences and fitness, and very high correlation between the most preferred niche and the niche providing the highest fitness (Fig. 1f). There is also significant genetic differentiation in male and



**Fig. 1** Three different outcomes of evolutionary dynamics in  $2 \times 1$  systems: no adaptation to the new niche (a, d, g, j), the evolution of a generalist species (b, c, h, k), and speciation (c, f, i, l). (a, b, c) Each background square represents a niche. Each circle represents a local population. The area of the circle is proportional to the population size. The colour of the circle identifies the niche preferred by the majority of individuals. (d, e, f) The  $2 \times 2$  matrix for which the brightness of the  $(i, j)$ th element is proportional to the number of individuals who have the highest fitness in niche  $i$  and the highest preference for niche  $j$  ( $i, j = 0, 1$ ). (g, h, i) The distribution of the male display ( $m$ ) and female preference ( $f$ ) traits within each of the two niches (0 and 1). The traits change between 0 (upper boundary of the graph) and 1 (lower boundary of the graph). The intensity of the black colour is proportional to the frequency of individuals with the corresponding trait value. (j, k, l) The overall distribution of male display trait ( $m$ ) vs. the overall distribution of female mating tolerance trait  $c$ . The traits change between 0 (upper left corner of the graph) and 1 (lower right corner of the graph). The intensity of the black colour is proportional to the frequency of individuals with the corresponding trait value.

female mating characters between the niches (Fig. 1i, l), and individuals exhibit strong nonrandom mating (Fig. 1i, l). In the example given in Fig. 1(c, f, i, l), females of both species exhibit negative assortative mating ( $c \approx 0$ ). In niche 0, female trait  $f$  is close to 1 whereas the male trait  $m$  is close to 0. In niche 1, the situation is reversed. In this example, speciation was sympatric.

It is not clear whether a two-deme model is an appropriate description of the Lake Apoyo cichlids system. The fish can readily swim across the lake, *A. zalius* can be found in the centre of the lake (Barlow & Munsey 1976), and there

is no geographical structuring within each of the two species (Barluenga *et al.* 2006a). However these observations are also compatible with a relatively low probability of long-distance dispersal. Most importantly, we do not know how often their ancestors (before the speciation) were involved in such long-distance movement. Therefore we have also looked at circular stepping-stone systems where demes can be viewed as arranged along the lake's shore at two different depths. Figure 2 illustrates the ways three different outcomes are characterized graphically in our simulations of  $2 \times 32$  systems. Note that in actual

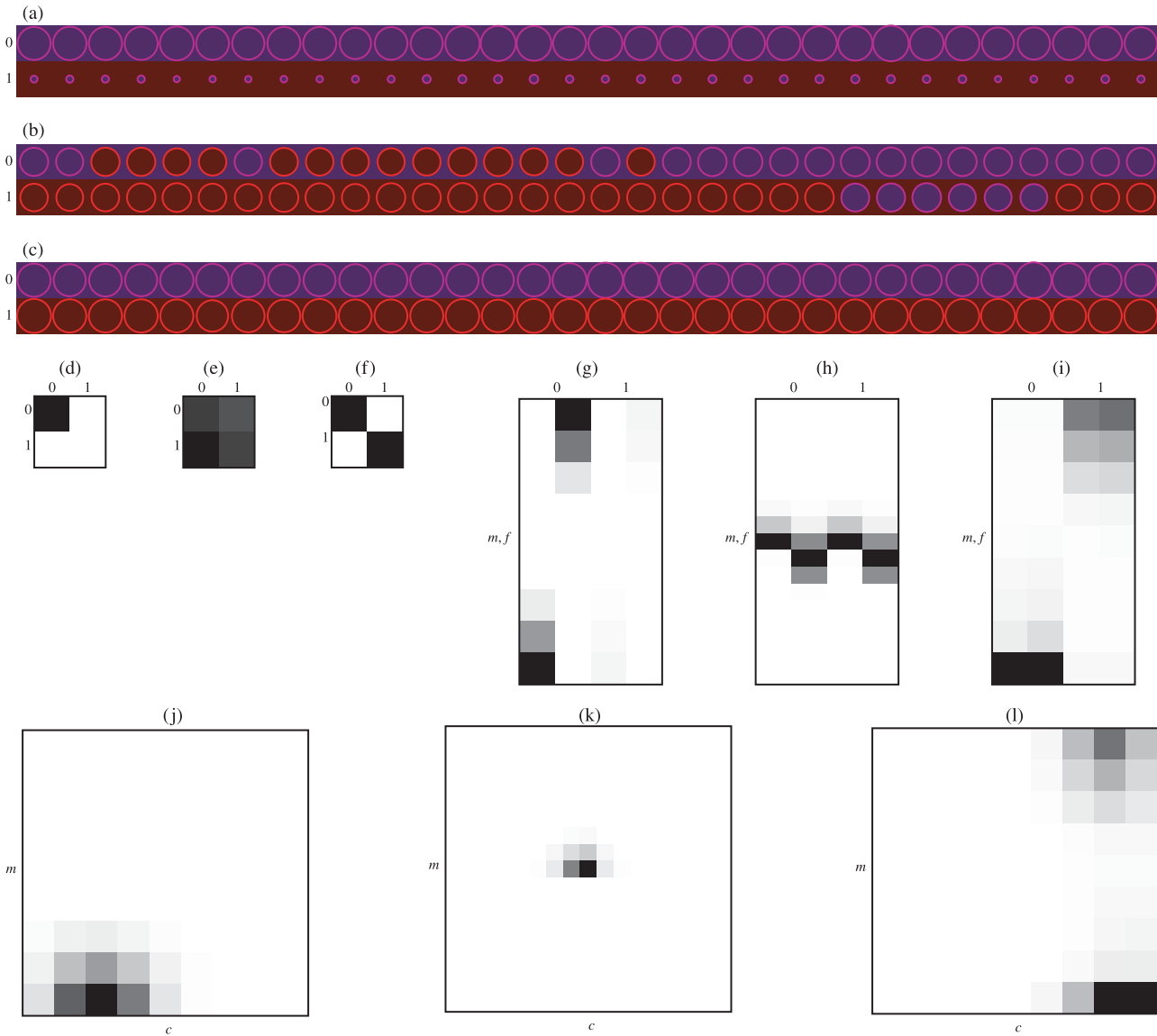


Fig. 2 Three different outcomes of evolutionary dynamics in  $2 \times 32$  systems. See the legend for Fig. 1.

simulations, the two left-most demes and the two right-most demes are viewed as neighbors exchanging migrants. Notice that in Fig. 2 in the case of no adaptation to the new niche (parts a, d, g and j), the ancestral population evolves (by random genetic drift and mutation) to a state with strong negative assortative mating. In the case of speciation (Fig. 2c, f, i, l), females of both species exhibit positive assortative mating ( $c \approx 1$ ). In niche 0, both female trait  $f$  and male trait  $m$  are close to 1 whereas in niche 1, both these traits are close to zero. In this example, speciation was parapatric.

Tables 3 and 4 show the numbers of different outcomes (no adaptation to the new niche, evolution of a generalist species, and sympatric or parapatric emergence of a new specialist species) observed in 10 preliminary runs for a

number of parameter combinations with  $L = 4$  loci per trait and  $\sigma_a = 0.05$ . One observes that: (i) the deep-water niche typically remains uncolonized if selection is too strong ( $\sigma_s = 0.35$  so that fitness of near-shore specialist in the deep-water niche is  $w_{spec} = 0.017$ ); (ii) provided adaptation to the new niche does happen, the evolution of a generalist species is more common than speciation if selection is weak ( $\sigma_s = 0.50$  so that  $w_{spec} = 0.135$ ) in both  $2 \times 1$  and  $2 \times S$  systems; and (iii) speciation is substantially more frequent if the maximum carrying capacity in  $2 \times 1$  systems is large ( $K_0 = 32\ 000$ ), or when the overall number of demes in  $2 \times S$  systems is large ( $S = 32$ ).

In effect, the  $2 \times 32$  systems produced the largest number of speciation cases. Therefore, we decided to study these systems in more detail. Table 5 summarizes our results.

**Table 3** Number of different outcomes in 10 preliminary runs in  $2 \times 1$  systems for different strengths of selection for local adaptation  $\sigma_s$  and different carrying capacities  $K_0$  with  $L = 4$  loci per trait and  $\sigma_a = 0.05$ . When observed, speciation is sympatric

$K_0$	$\sigma_s$	No		
		Adaptation	Generalist	Speciation
8000	0.35	10	—	—
	0.4	10	—	—
	0.45	10	—	—
	0.5	8	2	—
16 000	0.35	10	—	—
	0.4	9	—	1
	0.45	8	—	2
	0.5	3	6	1
32 000	0.35	10	—	—
	0.4	9	—	1
	0.45	4	4	2
	0.5	—	6	4

We observe that speciation may happen with a relatively high frequency within the span of less than 20 000 generations. However, the conditions for speciation are restrictive. If selection is strong, the empty niche remains uncolonized. If selection is weak, a generalist species evolves. Speciation, which in this case is parapatric, is observed only if the number of loci controlling the traits under selection is small ( $L = 4$ ) and the coefficient  $\sigma_a = 0.05$ . Note that under these conditions and with  $\sigma_s \geq 0.45$ , invasion of the empty niche always occurs.

Table 6 provides additional results for  $2 \times 1$  systems. In this case the maximum carrying capacity was set at  $K_0 = 51\,200$  offspring per patch. Note that  $1600 \times 32 = 51\,200$

**Table 4** Number of different outcomes in 10 preliminary runs in  $2 \times S$  systems for different strengths of selection for local adaptation  $\sigma_s$  and  $S$  with  $L = 4$  loci per trait,  $\sigma_a = 0.05$ , and carrying capacity  $K_0 = 1600$ . When observed, speciation is parapatric

size	$\sigma_s$	No		
		adaptation	Generalist	Speciation
$2 \times 10$	0.4	10	—	—
	0.45	8	—	2
	0.5	3	7	—
$2 \times 24$	0.4	10	—	—
	0.45	6	3	1
	0.5	6	4	—
$2 \times 32$	0.4	10	—	—
	0.45	2	2	3
	0.5	—	7	6

so that the overall population sizes were the same as in  $2 \times 32$  systems reported above. In these  $2 \times 1$  systems, speciation (by definition, sympatric) was observed only a few times. The data in Table 6 together with those in Table 3 show that speciation is most likely to be observed for intermediate carrying capacity  $K_0$ : with small carrying capacity (e.g.  $K_0 = 8000$ ) the niche remains uncolonized, while with large carrying capacity (e.g.  $K_0 = 51\,200$ ) a generalist species evolves. Note that speciation in  $2 \times 1$  systems can be achieved with stronger selection than in  $2 \times 32$  systems.

Tables 5 and 6 also show the average time to invasion and the average time to speciation. The former was defined as the first generation in which the population size of niche 1 is within 10% of the population size of niche 0. The latter was found by visual inspection of graphs similar to those in Figs 1 and 2 as the first moment when the average value

**Table 5** Numbers of different outcomes in 20 runs for each parameter combination in  $2 \times 32$  systems with carrying capacity  $K_0 = 1600$ . Also shown is the average time to invasion (ATI) and the average time to speciation (ATS) as well as the average strength of nonrandom mating  $C (= |2c-1|)$  in the last generation computed over the runs in which these events were observed. Parameter  $\sigma_s$  characterizes the strength of selection for local adaptation,  $L$  is the number of loci per character, and parameter  $\sigma_a$  characterizes the efficiency of female mate choice. When observed, speciation is parapatric

Parameters			Freq. of outcomes			C				
$\sigma_a$	$\sigma_s$	$L$	No	Gen	Spec	ATI	ATS	No	Gen	Spec
0.05	0.4	4	19	—	1	13 800	13 800	0.36	—	0.7
		8	20	—	—	—	—	0.19	—	—
	0.45	4	—	11	9	4 320	14 256	—	0.2	0.69
		8	20	—	—	—	—	0.2	—	—
	0.5	4	—	11	9	1 610	16 600	—	0.25	0.6
		8	9	11	—	6 663	—	0.15	0.19	—
0.1	0.4	4	20	—	—	—	—	0.19	—	—
		8	20	—	—	—	—	0.16	—	—
	0.45	4	3	17	—	6 988	—	0.19	0.14	—
		8	20	—	—	—	—	0.17	—	—
	0.5	4	—	20	—	1 580	—	—	0.22	—
		8	9	11	—	5 600	—	0.19	0.17	—

**Table 6** Numbers of different outcomes in  $2 \times 1$  systems with carrying capacity  $K_0 = 51\,200$ . Also shown is the average time to invasion (ATI) and the average time to speciation (ATS) as well as the average strength of nonrandom mating  $C (= |2c - 1|)$  in the last generation computed over the runs in which these events were observed. Parameter  $\sigma_s$  characterizes the strength of selection for local adaptation,  $L$  is the number of loci per character, and parameter  $\sigma_a$  characterizes the efficiency of female mate choice. There were 24 runs for most parameter combinations. When observed, speciation is sympatric

Parameters			Freq. of outcomes			C				
$\sigma_a$	$\sigma_s$	$L$	No	Gen	Spec	ATI	ATS	No	Gen	Spec
0.05	0.4	4	17	—	7	17 343	17 229	0.22	—	0.63
		8	23	—	—	—	—	0.16	—	—
	0.45	4	—	15	9	5 679	15 711	—	0.19	0.52
		8	24	—	—	—	—	0.23	—	—
	0.5	4	—	23	1	1 917	13 100	—	0.2	0.49
		8	18	3	—	7 875	—	0.14	0.09	—
0.1	0.4	4	22	—	—	4 443	—	0.23	—	—
		8	23	—	—	—	—	0.2	—	—
	0.45	4	3	21	—	—	—	0.34	0.19	—
		8	24	—	—	—	—	0.19	—	—
	0.5	4	1	23	—	1 165	—	0.48	0.21	—
		8	18	6	—	7 000	—	0.16	0.16	—

of female trait  $c$  deviates significantly from 0.5 and the distribution of traits  $m$  and  $f$  becomes bimodal simultaneously. As expected, the average time to invasion decreases with decreasing the strength of selection (i.e. increasing  $\sigma_s$ ) and the number of loci per trait  $L$ . The average time to speciation increases with decreasing the strength of selection.

The last three columns in Tables 5 and 6 present the average strength of nonrandom mating  $C (= |2c - 1|)$  in the last generation computed over the runs ending in a particular outcome. In runs ending up in speciation, the value of  $C$  does not approach the maximum possible value (which is one) within the time-span used in the simulations. Another notable observation is that some nonrandom mating (with  $C$  close to 0.2) always evolves even in the runs resulting in no adaptation to the new niche or in the evolution of a generalist species.

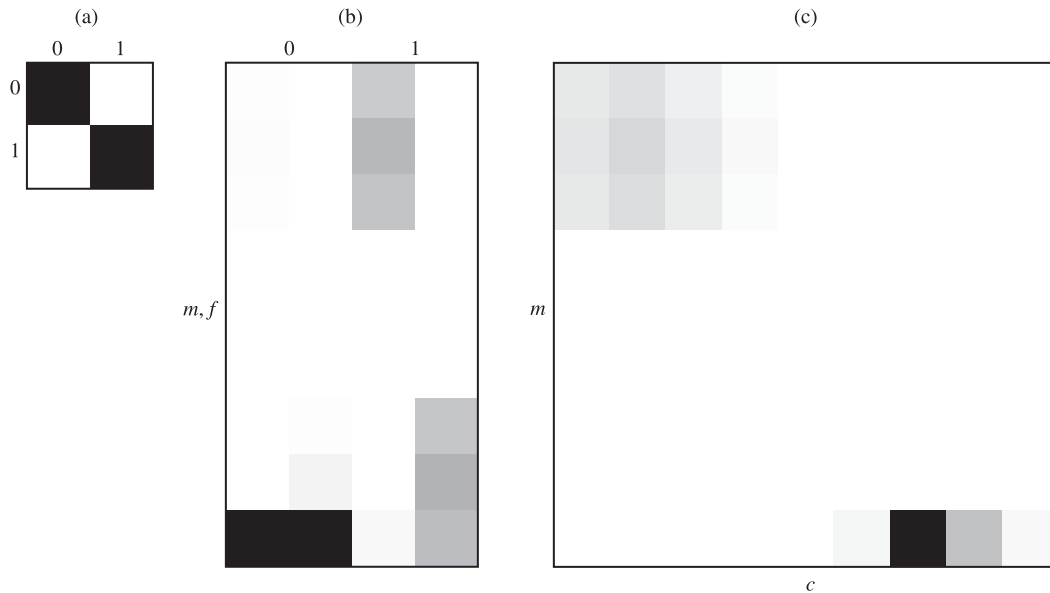
The nature of a state with a single generalist species requires some clarification. In fact, we observed two different situations (cf. Spichtig & Kawecki 2004; Gavrilets & Vose 2007). In the first, the generalist species maintains very high levels of genetic variation with allele frequencies in the loci controlling ecological and habitat preference traits close to 0.5. There is also some genetic divergence between the subpopulations in the two niches and some correlation between the most preferred niche and the niche providing the highest fitness. The high-variation generalist state was illustrated in Fig. 1(b, e, h, k). Alternatively, the generalist species is almost monomorphic with a very low level of genetic variation in the ecological and habitat preference traits. Half of the corresponding loci are almost fixed for allele 0, while the rest are almost fixed for allele 1. There is no genetic divergence between the subpopulations

inhabiting the two niches, and the correlation between the most preferred niche and the niche providing the highest fitness is absent. The low-variation generalist state was illustrated in Fig. 2(b, e, h, k).

The low-variation generalist state was always observed with  $L = 8$  and  $\sigma_s = 0.5$  in both  $2 \times 1$  and  $2 \times 32$  systems. These parameter values result in very weak selection on each individual locus. In contrast, the high-variation generalist state was almost always observed for all other parameter combinations.

The speciation examples given in Figs 1 and 2 show the evolution of either positive or negative assortative mating in both species. It is also possible that one species evolves positive assortative mating whereas the other species evolves negative assortative mating. In the example shown in Fig. 3, there is positive assortative mating in niche 0, with female traits  $c$  and  $f$  and male trait  $m$  close to one. In niche 1, mating is negative assortative with female trait  $c$  close to 0, female trait  $f$  close to 1, and male trait  $m$  close to zero. In fact, the situation with positive assortative mating in one species and negative assortative mating in another species was the most common outcome in the runs ending in speciation as summarized in Tables 5 and 6. This outcome was observed 26 times. The outcome with positive assortative mating in both species was observed seven times, and that with negative assortative mating in both species was observed only twice.

Evolution towards a state with two specialists may or may not involve an intermediate state with a single generalist species. For the data given in Tables 5 and 6, for all parameter combinations where evolution towards a generalist species was observed, evolution of two specialists



**Fig. 3** An example of speciation when one species mates positively assortatively and the other species mates negatively assortatively. See the legend for Fig. 1.

involved an intermediate high-variance generalist stage. In other situations, speciation was achieved through a particular sequence of events. First, evolution of at least moderate assortative mating within the near-shore habitat was observed. This then resulted in a polymorphism in male and female mating characters achieved by sexual selection. The polymorphism was followed by a shift to the deep-water niche accomplished by one of the ‘sexual morphs’. Finally, divergence in mating preferences between the two populations was reinforced by selection against hybridization. Under this sequence of events, the time to speciation is shorter than the time to invasion (see Tables 5 and 6). This sequence of events was actually limited to the case of  $\sigma_s = 0.40$ ,  $\sigma_c = 0.05$  and  $L = 4$  in both  $2 \times 1$  and  $2 \times 32$  systems (i.e. when selection on individual loci was the strongest).

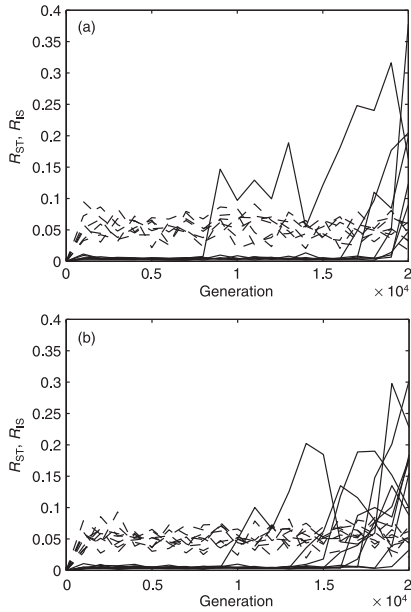
In all cases, the levels of genetic variation in female mating characters  $f$  and  $c$  were almost always significantly higher than in the male character  $m$ . This is a consequence of our assumption of the absence of any costs of being choosy for females. Since each female is guaranteed to mate no matter what her preference is, the female mating characters are not subject to any direct selection and can accumulate large variation by mutation. In contrast, males are always ‘trying’ to match the average female character  $f$ , and as a result, male trait  $m$  is subject to stabilizing selection. The fact that some genetic variance in female trait  $c$  is present implies that the female population can be highly polymorphic with regard to mating patterns. In particular, females expressing random, positive assortative, and negative assortative mating may all be present simultaneously in a single species.

The whole gallery of graphical results in the form of graphs similar to those in Figs 1, 2, and 3 for the data summarized in Tables 1 and 2 can be viewed at <http://www.tiem.utk.edu/~gavrila>.

What are the levels of divergence in neutral loci observed when speciation does happen? To estimate the levels of spatial structuring in neutral loci in  $2 \times 1$  systems we computed  $R_{ST}$  (the ratio of the genetic variance among groups and the total genetic variance) and  $R_{IS}$  (the ratio of the genetic variance among individuals within groups and the total genetic variance) using *micsatfs* (R. E. Strauss, unpublished; publicly available at <http://www.biol.ttu.edu/Strauss/Matlab/matlab.htm>). Figure 4 shows the dynamics of  $R_{ST}$  and  $R_{IS}$  statistics in most runs resulting in speciation as described in Table 6.

To estimate the levels of spatial structuring in neutral loci in  $2 \times 32$  systems we used the AMOVA framework (Excoffier *et al.* 1992; Excoffier 2001). Specifically, we computed  $R_{ST}$  (the correlations of random pair of haplotypes within demes, relative to that of random pairs of haplotypes drawn from the whole system),  $R_{SC}$  (the correlations of random pair of haplotypes within demes, relative to that of random pairs of haplotypes drawn from the demes within the same habitat), and  $R_{CT}$  (the correlations of random pairs of haplotypes within a group of demes from the same habitat, relative to that of random pairs of haplotypes drawn from the whole system). Figures 5 and 6 show the dynamics of  $R_{ST}$ ,  $R_{SC}$  and  $R_{CT}$  statistics in most runs resulting in speciation as described in Table 5.

These figures show that divergence in neutral markers does take place after speciation, with  $2 \times 1$  systems achieving higher levels of  $R_{ST}$  than the levels of  $R_{CT}$  in  $2 \times 32$ . The



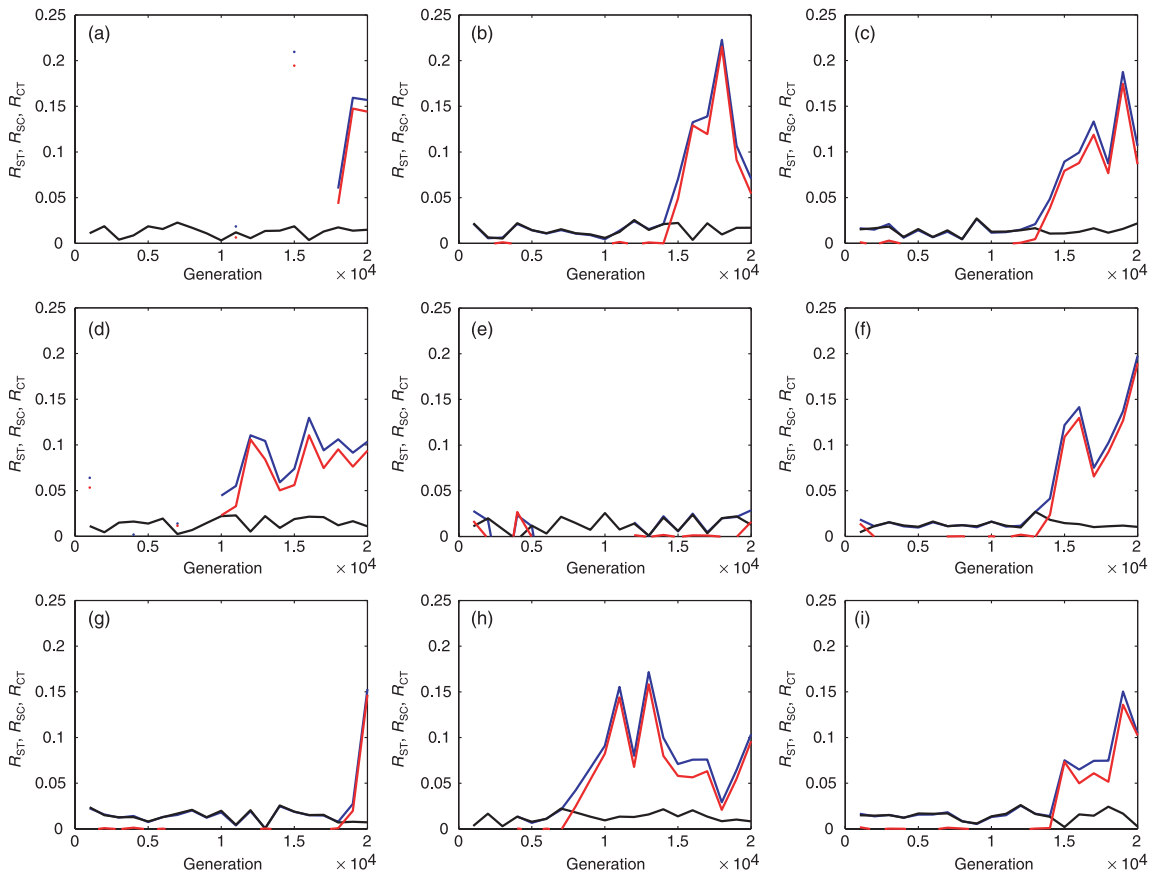
**Fig. 4** The dynamics of  $R_{ST}$  (solid line) and  $R_{IS}$  (dashed line) in  $2 \times 1$  systems with  $L = 4$  and  $\sigma_a = 0.05$ . (a) The seven runs that resulted in (sympatric) speciation with  $\sigma_a = 0.40$ . (b) The nine runs that resulted in (sympatric) speciation with  $\sigma_a = 0.45$ .

overall levels of divergence are not very high, although this is probably as a result of the relatively short time intervals used.

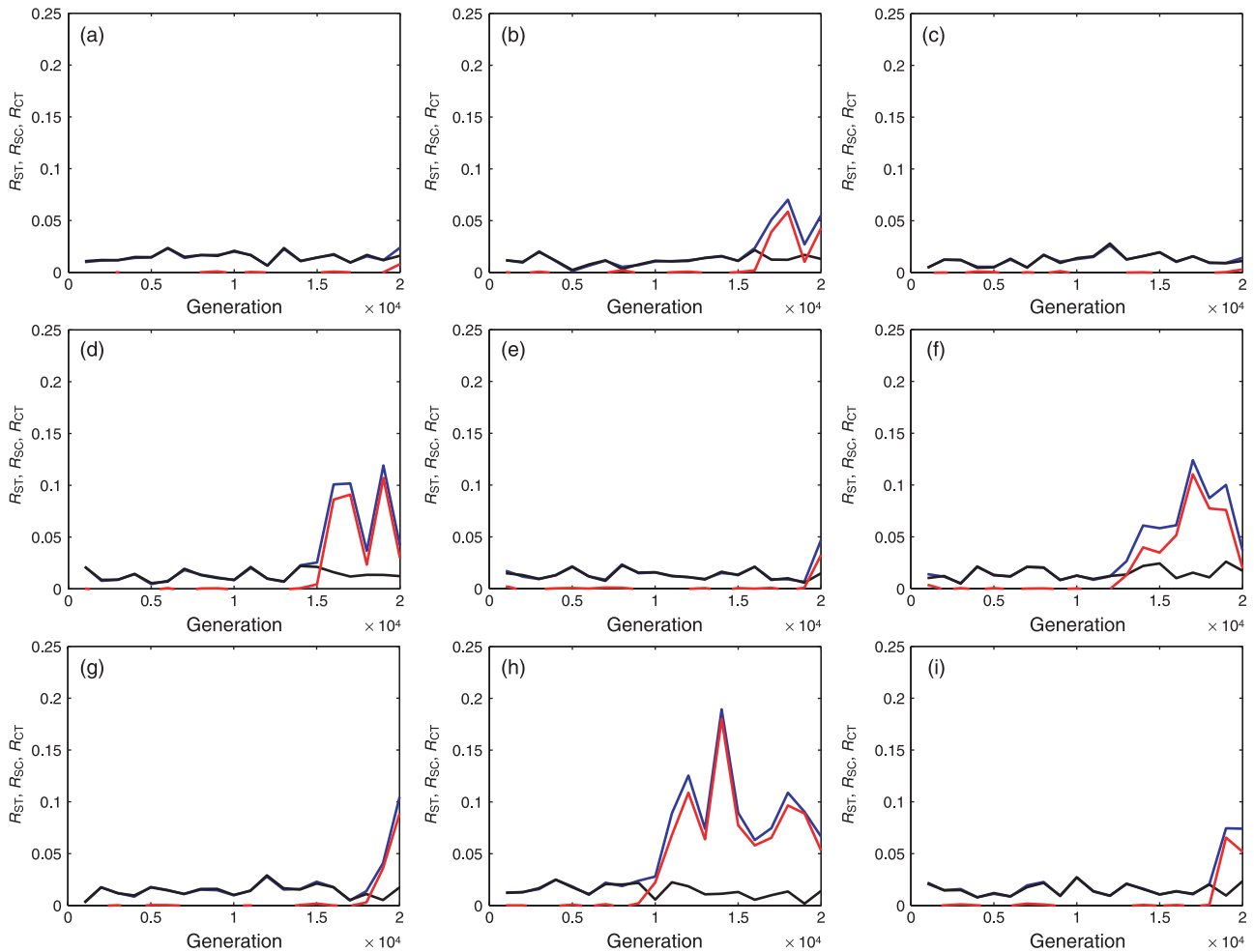
**Discussion**

Overall, our results show that relatively rapid ( $< 20\,000$  generations) colonization of a new ecological niche and speciation via local adaptation and divergence in habitat and mating preferences are theoretically plausible if: (i) the number of loci underlying the traits is small; (ii) the strength of selection for local adaptation is intermediate; (iii) the carrying capacity of the population is intermediate; and (iv) the effects of the loci influencing nonrandom mating are strong.

The observation that the number of loci underlying the traits involved in local adaptation and reproductive isolation must be small for nonallopatric speciation to occur relatively rapidly has been repeatedly made previously (e.g. Gavrilets 2004; Gavrilets & Vose 2005, 2007). The reason for this observation is simple. In the presence of gene flow between the diverging groups, disruptive selection must be strong and the advantages of



**Fig. 5** The dynamics of  $R_{ST}$  (blue line),  $R_{SC}$  (black line), and  $R_{CT}$  (red line) in the nine runs that resulted in (parapatric) speciation in  $2 \times 32$  systems with  $\sigma_s = 0.45$ ;  $\sigma_a = 0.05$ , and  $L = 4$ .



**Fig. 6** The dynamics of  $R_{ST}$  (blue line),  $R_{SC}$  (black line), and  $R_{CT}$  (red line) in the nine runs that resulted in (parapatric) speciation in  $2 \times 32$  systems with  $\sigma_s = 0.50$ ;  $\sigma_a = 0.05$ , and  $L = 4$ .

evolving reproductive isolation must be significant. A smaller number of loci imply larger locus effects, stronger selection on each individual locus, and, simultaneously, weaker effects of recombination which destroys coadapted gene complexes (and linkage disequilibrium between the loci). All this promotes nonallopatric speciation. Reversing our argument, one can claim that genetics of traits underlying extensive diversification and speciation in the presence of gene flow must be simple (Gavrilets & Vose 2005). This is a testable prediction.

For speciation to occur rapidly, the strength of selection for local adaptation and the carrying capacity must be intermediate. When selection is too strong or carrying capacity is too low, an ancestral specialist species maintains relatively low genetic variation which prevents (or, at least, significantly delays) the emergence of new genotypes that may survive in the new niche and seed a new species there. As a result, the empty niche remains unoccupied. If selection is too weak, then a generalist has relatively high fitness in both habitats. As a result, a single

generalist species evolves. It is only with intermediate strength of selection that enough of genetic variation is preserved in the ancestral specialist species for the invasion of the empty niche and, simultaneously, specialization results in higher fitness. There is also a tendency towards increasing the probability of evolving a generalist species as the carrying capacity increases, at least in  $2 \times 1$  systems.

In our simulations, speciation was only observed when parameter  $\sigma_a$ , measuring the effectiveness of female mate choice, was equal to 0.05. This value implies that a single change in a locus controlling mating can cause dramatic change in the probability of mating (see Table 2). The numerical value  $\sigma_a = 0.05$  is the same as used previously (Bolnick 2004; Doebeli *et al.* 2005). However, it is currently unknown how realistic such a small value of  $\sigma_a$  is. Based on our results, one can speculate that mate choice in Lake Apoyo cichlids is controlled by few major effect genes.

Rapid speciation in the presence of gene flow usually requires the absence of costs on being choosy (Bolnick

2004; Gavrillets 2004, 2005; Gourbiere 2004; Kirkpatrick & Nuismer 2004; Waxman & Gavrillets 2005a). Our results clearly show that in this case mutation and random genetic drift should result in the evolution of some nonrandom mating and in the within-population polymorphism with regard to female mating preferences. Therefore the presence or absence of these two effects in natural population can serve as an indicator of the absence or presence of the costs on being choosy.

Evolution towards a state with two specialists may or may not involve an intermediate state with a single generalist species. With weak selection for local adaptation, evolution of two specialists involved a generalist stage. In particular, specialization for the near-shore niche was first lost and then recreated. With strong selection for local adaptation and a small number of loci per trait, speciation was achieved through a particular sequence of events. First, evolution of at least moderate assortative mating within the near-shore niche was observed. This then resulted in a polymorphism in male and female mating characters achieved by sexual selection. The polymorphism was followed by niche shift accomplished by one of the 'sexual morphs'. Finally, the mating preferences were reinforced by selection against hybridization. That is, in this scenario, differentiation by sexual selection precedes ecological divergence. Interestingly, Wilson *et al.* (2000) found hints that assortative mating on the basis of colour may be playing a role in the divergence of populations within Nicaraguan lakes.

In our symmetric model, the coexistence of a generalist species adapted to both habitats and a specialized species was not observed. We expect that such a coexistence could happen if the distributions of the niches and the strength of selection for local adaptation are asymmetric.

In our model, individuals had to choose a niche to feed and mate in at the very beginning of the life cycle. Subsequently, individuals were given no possibility to change the niche even if they were unsuccessful in obtaining resources. Allowing for such a change is expected to accelerate speciation as low fitness individuals in the near-shore niche that will preferentially move to the deep-water niche will carry mutations advantageous in the latter niche. We also assumed that mating pairs were formed in the feeding habitat which appears to be an accurate assumption for the system studied. That mating pairs form before the nesting site is chosen, i.e. in the feeding grounds, is suggested by the observation that single individuals cannot defend and/or occupy breeding sites alone (see e.g. Barlow 2000). Also, numerous pairs moving around in search of breeding sites have been observed in *A. citrinellus* in another crater lake in Nicaragua, Lake Xiloa (McKaye 1977). Relaxing this assumption will prevent speciation (at least for parameter values and time-scales used here). We conservatively assumed that all loci were unlinked. Allowing for linkage between genes for local adaptation

and niche preference will help divergence and speciation (cf. Udovic 1980; Felsenstein 1981; Gavrillets 2004).

Speciation was occurring more rapidly, and stronger reproductive isolation (as measured by the deviation of the value of the mating tolerance trait  $c$  from 0.5) was achieved in  $2 \times 32$  systems than in  $2 \times 1$  systems. However, higher levels of differentiation in the neutral markers were observed for  $2 \times 1$  systems. In either case, the degree of divergence in neutral markers was not too high, probably because of the relatively short time interval used. To allow for direct comparisons between the modelled  $R_{ST}$  values and real data, we have recalculated  $R_{ST}$  with the initial data-set of Barluenga *et al.* (2006a). The  $R_{ST}$  between *A. citrinellus* (in Apoyo) and *A. zaliosus* is 0.13 which, of course, is highly significant ( $P = 0.001$ ). This value is comparable to those observed in the simulations. We note that these levels of genetic differentiation in microsatellites provide little evidence of reproductive isolation on their own. However, here it is much more compelling because of the supporting data on mating isolation.

Here, the time-scale for speciation under most favourable conditions was on the order of 10 000 generations. Hendry *et al.* (2007) argue that significant progress towards ecological speciation can be achieved very rapidly on the time scale of a few dozen generations. In our model and under biologically realistic conditions, achieving the speed of ecological speciation envisioned by Hendry *et al.* (2007) seems to be impossible.

The actual transition from a specialist to a generalist or from one species to two species was very rapid and took a few hundred generations. This is a general feature of models in which speciation is driven by selection in the presence of gene flow (Gavrillets 2004). This feature is compatible with patterns of 'punctuated equilibrium' (Eldredge 1971; Eldredge & Gould 1972; Gould & Eldredge 1993; Gould 2002; Eldredge *et al.* 2005).

In our model, the probability of speciation within the time frame studied was about 50% under most favourable conditions. We view these results as providing theoretical support for the original claim of Barluenga *et al.* (2006a) about *in situ* emergence of the new species of cichlids in Lake Apoyo. Whether speciation was indeed sympatric or parapatric cannot be distinguished on the basis of our model. The observable characteristics of speciation in  $2 \times 1$  and  $2 \times 32$  systems were pretty similar. Similar ecological systems with a division into a benthic and a limnetic habitat exist in other lakes (e.g. Schliewen *et al.* 1994, 2001; Gíslason *et al.* 1999; Schliewen & Klee 2004). The similarities in biological setups imply that some of our conclusions should be applicable to these other systems as well.

Our conclusions, however, should not be interpreted as the statement that sympatric (or nonallopatric) speciation occurs easily. In our model, speciation has been observed only under appropriate conditions and parameters (which

we discuss throughout the paper). As is apparent in this and many other models, genetics strongly constrains the dynamics of ecological speciation. In our opinion, the occasionally made claims about the ease with which sympatric or ecological speciation can happen (e.g. Dieckmann & Doebeli 1999; Higashi *et al.* 1999; Kondrashov & Kondrashov 1999; Doebeli 2005; Doebeli & Dieckmann 2005; Doebeli *et al.* 2005) are not justified (Gavrilets 2004, 2005; Waxman & Gavrilets 2005a, b). The results presented here and in the accompanying paper (Gavrilets & Vose 2007) only reinforce this opinion.

Schliewen *et al.* (2006) suggest the two cichlid species studied by Barluenga *et al.* (2006a) emerged as a result of double invasion where *A. zalius* resulted from the first colonization, and the Lake Apoyo population of *A. citrinellus* was the result of a 'second wave of colonization' from Lake Nicaragua that was followed by introgression from *A. zalius*. As argued by Barluenga *et al.* (2006b), this scenario is incompatible with existing molecular evidence. Moreover, it is hard to reconcile with the ecology of the fish. Indeed, why would ancestors of the first colonizers evolve adaptations to a completely new open-water environment when their ancestral near-shore niche was available? One can probably speculate that *A. zalius* switched to the new niche after the secondary invasion as a result of character displacement but this brings another problem. It is hardly possible to imagine anything resembling a 'wave of colonization' given the isolated nature of the lake. Any colonization would involve a very small number of fish which will most likely not have any significant ecological impact on the resident species. Overall the scenario advanced by Schliewen *et al.* (2006) appears much less likely than that of Barluenga *et al.* (2006a, b).

Although a lot of useful information on Lake Apoyo cichlids now exists, theoretical research would definitely benefit from additional data. In particular, it is important to have good estimates of the actual population densities and sizes, to know the extent of individual dispersal from the place of birth to the place of mating-pair formation, and to have good understanding of how and where mating pairs are formed. Of equal importance would be estimates of the strength of natural selection coming from competition for food and nesting sites, costs of choosiness, and the proportions of individuals that reproduce. Finally, the knowledge of the number of genes involved in ecological and mating traits, their mutation rates, and linkage relationship would greatly help obtaining more precise theoretical conclusions about the plausibility of different scenarios of speciation of cichlids in Lake Apoyo.

### Acknowledgements

We thank two anonymous reviewers and Loren Rieseberg for very useful comments. This work was supported by National Institutes

of Health (Grant GM56693), National Science Foundation (Grant DEB-0111613), the University of Konstanz, and the Deutsche Forschungsgemeinschaft.

### References

- Barlow GW (2000) *The Cichlid Fishes — Nature's Grand Experiment in Evolution*. Perseus Publishing, Cambridge, Massachusetts.
- Barlow GW, Munsey JW (1976) The red devil-Midas-arrow cichlid species complex in Nicaragua. In: *Investigations of the Ichthyofauna of Nicaraguan Lakes* (ed. Thorson TB), pp. 359–369. University of Nebraska Press, Lincoln, Nebraska.
- Barluenga M, Meyer A (2004) The Midas cichlid species complex: incipient sympatric speciation in Nicaraguan cichlid fishes? *Molecular Ecology*, **13**, 2061–2076.
- Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A (2006a) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature*, **439**, 719–723.
- Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A (2006b) Reply: Evidence for sympatric speciation? *Nature*, **444**, E13.
- Baylis JR (1976) A quantitative study of long-term courtship: I. Ethological isolation between sympatric populations of the Midas cichlid, *Cichlasoma citrinellum*, and the arrow cichlid, *C. zalius*. *Behaviour*, **59**, 5969.
- Bolnick D (2004) Waiting for sympatric speciation. *Evolution*, **58**, 895–899.
- Bolnick D (2006) Multi-species outcomes in a common model of sympatric speciation. *Journal of Theoretical Biology*, **241**, 734–744.
- Christiansen F (1975) Hard and soft selection in a subdivided population. *American Naturalist*, **109**, 11–16.
- Coyne J, Orr HA (2004) *Speciation*. Sinauer Associates, Inc., Sunderland, Massachusetts.
- DeMeeus T, Michalakis Y, Renaud F, Olivieri I (1993) Polymorphism in heterogeneous environments, evolution of habitat selection and sympatric speciation — soft and hard selection models. *Evolutionary Ecology*, **7**, 175–198.
- Dieckmann U, Doebeli M (1999) On the origin of species by sympatric speciation. *Nature*, **400**, 354–357.
- Dieckmann U, Doebeli M, Metz JAJ, Tautz D (2004) *Adaptive Speciation*. Cambridge University Press, Cambridge.
- Diehl SR, Bush GL (1989) The role of habitat preference in adaptation and speciation. In: *Speciation and its Consequences* (eds Otte D, Endler JA), pp. 345–365. Sinauer, Sunderland, Massachusetts.
- Dobzhansky T, Ayala FJ, Stebbins GL, Valentine JW (1977) *Evolution*. W.H. Freeman, San Francisco.
- Doebeli M (2005) Adaptive speciation when assortative mating is based on female preference for male marker traits. *Journal of Evolutionary Biology*, **18**, 1587–1600.
- Doebeli M, Dieckmann U (2005) Adaptive dynamics as a mathematical tool for studying the ecology of speciation processes. *Journal of Evolutionary Biology*, **18**, 1194–1200.
- Doebeli M, Dieckmann U, Metz JAJ, Tautz D (2005) What we have also learned: adaptive speciation is theoretically plausible. *Evolution*, **59**, 691–695.
- Eldredge N (1971) The allopatric model and phylogeny in Paleozoic invertebrates. *Evolution*, **25**, 156–167.
- Eldredge N, Gould SJ (1972) Punctuated equilibria: an alternative to phyletic gradualism. In: *Models in Paleobiology* (ed. Schopf TJ), pp. 82–115. Freeman, Cooper, San Francisco.

- Eldredge N, Thompson JN, Brakefield PM *et al.* (2005) Dynamics of evolutionary stasis. *Paleobiology*, **31**, 133–145.
- Excoffier L (2001) Analysis of population subdivision. In: *Handbook of Statistical Genetics* (eds Balding DJ, Bishop M, Cannings C), pp. 271–307. John Wiley & Sons, Ltd.
- Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Felsenstein J (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution*, **35**, 124–138.
- Fry JD (2003) Multilocus models of sympatric speciation: Bush versus Rice versus Felsenstein. *Evolution*, **57**, 1735–1746.
- Gavrilets S (2003) Models of speciation: what have we learned in 40 years? *Evolution*, **57**, 2197–2215.
- Gavrilets S (2004) *Fitness Landscapes and the Origin of Species*. Princeton University Press, Princeton, New Jersey.
- Gavrilets S (2005) Adaptive speciation: it is not that simple. *Evolution*, **59**, 696–699.
- Gavrilets S (2006) The Maynard Smith model of sympatric speciation. *Journal of Theoretical Biology*, **239**, 172–182.
- Gavrilets S, Vose A (2005) Dynamic patterns of adaptive radiation. *Proceedings of the National Academy of Sciences USA*, **102**, 18040–18 045.
- Gavrilets S, Vose A (2007) Case studies and mathematical models of ecological speciation. 2. Palms on an oceanic island. *Molecular Ecology*, **00**, 000–000.
- Gavrilets S, Waxman D (2002) Sympatric speciation by sexual conflict. *Proceedings of the National Academy of Sciences USA*, **99**, 10533–10538.
- Gíslason D, Ferguson MM, Skúlason S, Snorason SS (1999) Rapid and coupled phenotypic differentiation in Icelandic Arctic charr (*Salvelinus alpinus*). *Canadian Journal of Fishery and Aquatic Science*, **56**, 2299–2334.
- Gould SJ (2002) *The Structure of Evolutionary Theory*. Harvard University Press, Boston.
- Gould SJ, Eldredge N (1993) Punctuated equilibrium comes of age. *Nature*, **366**, 223–227.
- Gourbiere S (2004) How do natural and sexual selection contribute to sympatric speciation? *Journal of Evolutionary Biology*, **17**, 1297–1309.
- Hendry AP, Nosil P, Rieseberg LH (2007) The speed of ecological speciation. *Functional Ecology*.
- Higashi M, Takimoto G, Yamamura N (1999) Sympatric speciation by sexual selection. *Nature*, **402**, 523–526.
- Howard DJ, Berlocher SH (1998) *Endless Forms: Species and Speciation*. Oxford University Press, New York.
- Johnson PA, Hoppensteadt FC, Smith JJ, Bush GL (1996) Conditions for sympatric speciation: a diploid model incorporating habitat fidelity and non-habitat assortative mating. *Evolutionary Ecology*, **10**, 187–205.
- Kawata M (2002) Invasion of empty niches and subsequent sympatric speciation. *Proceedings of the Royal Society London B*, **269**, 55–63.
- Kawecki TJ (1996) Sympatric speciation driven by beneficial mutations. *Proceedings of the Royal Society London B*, **263**, 1515–1520.
- Kawecki TJ (1997) Sympatric speciation via habitat specialization driven by deleterious mutations. *Evolution*, **51**, 1751–1763.
- Kirkpatrick M, Nuismer SL (2004) Sexual selection can constrain sympatric speciation. *Proceedings of the Royal Society London B*, **271**, 687–693.
- Kirkpatrick M, Ravigné V (2002) Speciation by natural and sexual selection: models and experiments. *American Naturalist*, **159**, S22–S35.
- Kisdi E, Geritz SAH (1999) Adaptive dynamics in allele space: evolution of genetic polymorphism by small mutations in heterogeneous environment. *Evolution*, **53**, 993–1008.
- Klingenberg C, Barluenga M, Meyer A (2002) Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. *Evolution*, **56**, 1909–1920.
- Klingenberg C, Barluenga M, Meyer A (2003) Body shape variation in cichlid fishes of the *Amphilophus citrinellus* species complex. *Biological Journal of the Linnean Society*, **80**, 397–440.
- Kondrashov AS, Kondrashov FA (1999) Interactions among quantitative traits in the course of sympatric speciation. *Nature*, **400**, 351–354.
- Kot M (2001) *Elements of Mathematical Ecology*. Cambridge University Press, Cambridge.
- Mayr E (1947) Ecological factors in speciation. *Evolution*, **1**, 263–288.
- Mayr E (1982) *The Growth of Biological Thought*. Harvard University Press, Cambridge, Massachusetts.
- McKaye KR (1977) Competition for breeding sites between the cichlid fishes of Lake Jilóá. *Nicaragua. Ecology*, **58**, 291–302.
- McKaye KR, Barlow GW (1976) Competition between color morphs of the Midas cichlid, *Cichlasoma citrinellum*, in Lake Jilóá, Nicaragua. In: *Investigations of the Ichthyofauna of Nicaraguan Lakes* (ed. Thorson TB), pp. 333–358. University of Nebraska Press, Lincoln, Nebraska.
- McKaye KR, Stauffer JR, van der Berghe EP *et al.* (2002) Behavioral, morphological and genetic evidence of divergence of the Midas cichlid species complex in two Nicaraguan crater lakes. *Cuadernos de Investigaci de la Universidad Centroamericana*, **12**, 19–47.
- Meyer A (1989) Costs and benefits of morphological specialization: feeding performance in the trophically polymorphic Neotropical cichlid fish, *Cichlasoma citrinellum*. *Oecologia*, **80**, 431–436.
- Meyer A (1990a) Ecological and evolutionary aspects of the trophic polymorphism in *Cichlasoma citrinellum* (Pisces: Cichlidae). *Biological Journal of the Linnean Society*, **39**, 279–299.
- Meyer A (1990b) Morphometrics and allometry of the trophically polymorphic cichlid fish, *Cichlasoma citrinellum*: Alternative adaptations and ontogenetic changes in shape. *Journal of Zoology*, **221**, 237–260.
- Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research*, **22**, 201–204.
- Pomiankowski A (1987) The costs of choice in sexual selection. *Journal of Theoretical Biology*, **128**, 195–218.
- Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336–352.
- Savolainen V, Anstett M-C, Lexer C *et al.* (2006) Sympatric speciation in palms on an oceanic island. *Nature*, **441**, 210–213.
- Schliewen U, Klee B (2004) Reticulate sympatric speciation in Cameroonian crater lake cichlids. *Frontiers in Zoology*, **1**, 5.
- Schliewen U, Tautz D, Pääbo S (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature*, **368**, 629–632.
- Schliewen U, Rassmann K, Markmann M, Markert J, Kocher T, Tautz D (2001) Genetic and ecological divergence of a monophyletic cichlid species pair under fully sympatric conditions in Lake Ejagham, Cameroon. *Molecular Ecology*, **10**, 1471–1481.

- Schliewen U, Kocher T, McKaye KR, Seehausen O, Tautz D (2006) Evidence for sympatric speciation? *Nature*, **444**, E12–E13.
- Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.
- Spichtig M, Kawecki T (2004) The maintenance (or not) of polygenic variation by soft selection in heterogeneous environment. *American Naturalist*, **164**, 70–84.
- Udovic D (1980) Frequency-dependent selection, disruptive selection, and the evolution of reproductive isolation. *American Naturalist*, **116**, 621–641.
- Waxman D, Gavrilets S (2005a) Issues of terminology, gradient dynamics, and the ease of sympatric speciation in Adaptive Dynamics. *Journal of Evolutionary Biology*, **18**, 1214–1219.
- Waxman D, Gavrilets S (2005b) Target review: 20 questions on adaptive dynamics. *Journal of Evolutionary Biology*, **18**, 1139–1154.
- Wilson AB, Noack-Kunmann K, Meyer A (2000) Incipient speciation in sympatric Nicaraguan crater lake cichlid fishes:

sexual selection versus ecological diversification. *Proceedings of the Royal Society London B*, **267**, 2133–2141.

---

Sergey Gavrilets is a theoretical evolutionary biologist currently focusing on the dynamics of speciation, coevolution, sexual conflict, and cultural and social evolution. Aaron Vose is a student research assistant interested in the synergy between hardware and software and its implications for the production of fast and efficient implementations designed specifically for particular hardware configurations. Marta Barluenga is a postdoctoral fellow interested in speciation and the origin of adaptive radiations. The research of Walter Salzburger and his groups focuses on the understanding of adaptation, speciation and diversification (particularly in cichlids). Axel Meyer is an evolutionary geneticist interested in the ecological, developmental and genomic aspects of speciation.

---

**Appendix Specifying the mating preference function**

**Here we justify our preference function (4) which is a modification of function**

$$\psi(m, f, c) = \begin{cases} \exp\left[-\frac{1}{2}(2c - 1)^4 \frac{(f - m)^2}{2\sigma_a^2}\right], & \text{if } c > 0.5, \\ 1, & \text{if } c = 0.5, \\ \exp\left[-\frac{1}{2}(2c - 1)^4 \frac{(1 - (f - m))^2}{2\sigma_a^2}\right], & \text{if } c < 0.5, \end{cases} \quad (5)$$

used by Bolnick (2004, 2006) and Doebeli (2005). First, one notices that function (5) uses the 4<sup>th</sup> power of the term  $(2c - 1)$  but the 2<sup>nd</sup> power of the term  $(f - m)$  inside the exponential. In the absence of any biological justification for this asymmetry and to avoid its possible artifacts, we

prefer to use the second order terms in both cases. Second, we have discovered that function (5) produces an artifact of strong sexual disruptive selection in males. To see this, assume that male and female traits are similar (specifically, that  $|m - f| = \epsilon$ , where  $\epsilon$  is small). Note that these conditions are close to the initial conditions in our simulations when mutation has introduced only limited variation. Then, under Bolnick-Doebeli parameterization, the ratio inside the exponential for females with positive assortative mating is equal to  $\epsilon^2/(2\sigma_a^2)$ , while that for females with negative assortative mating is approximately  $1/(2\sigma_a^2)$ . Because  $\epsilon$  is small, the latter term is much larger than the former term implying strong selection in males induced by females characterized by negatively assortative mating who prefer males at the edges of the male trait distribution. This selection dramatically increases genetic variance in males and artificially simplifies conditions for sympatric speciation. Our preference function (4) is symmetric and does not introduce such artifacts.