

# CASE STUDIES AND MATHEMATICAL MODELS OF ECOLOGICAL SPECIATION. 4. HYBRID SPECIATION IN BUTTERFLIES IN A JUNGLE

Edgar A. Duenez-Guzman,<sup>1,2</sup> Jesus Mavárez,<sup>3</sup> Michael D. Vose,<sup>2</sup> and Sergey Gavrillets<sup>1,4,5,6</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996

<sup>2</sup>Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, Tennessee 37996

<sup>3</sup>Centro de Ecología, Instituto Venezolano de Investigaciones Científicas, Apartado 20632, Caracas 1020-A, Venezuela

<sup>4</sup>Department of Mathematics, University of Tennessee, Knoxville, TN 37996

<sup>5</sup>National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, Tennessee 37996

<sup>6</sup>E-mail: sergey@tiem.utk.edu

Received October 28, 2008

Accepted May 14, 2009

We build a spatial individual-based multilocus model of homoploid hybrid speciation tailored for a tentative case of hybrid origin of *Heliconius heurippa* from *H. melpomene* and *H. cydno* in South America. Our model attempts to account for empirical patterns and data on genetic incompatibility, mating preferences and selection by predation (both based on coloration patterns), habitat preference, and local adaptation for all three *Heliconius* species. Using this model, we study the likelihood of recombinational speciation and identify the effects of various ecological and genetic parameters on the dynamics, patterns, and consequences of hybrid ecological speciation. Overall, our model supports the possibility of hybrid origin of *H. heurippa* under certain conditions. The most plausible scenario would include hybridization between *H. melpomene* and *H. cydno* in an area geographically isolated from the rest of both parental species with subsequent long-lasting geographic isolation of the new hybrid species, followed by changes in the species ranges, the secondary contact, and disappearance of *H. melpomene*-type ecomorph in the hybrid species. However, much more work (both empirical and theoretical) is necessary to be able to make more definite conclusions on the importance of homoploid hybrid speciation in animals.

**KEY WORDS:** Hybridization, mating systems, predation, speciation, selection—sexual.

Our understanding of the processes leading to the origin of new species has been steadily increasing both from empirical and theoretical perspectives (e.g., Coyne and Orr 2004; Gavrillets 2004; Dieckmann et al. 2004; Mallet 2007; Nosil 2007; Seehausen 2006; Seehausen 2007; Bolnick and Fitzpatrick 2007; Gavrillets and Losos 2009). One of the lessons of recent work is a renewed appreciation of the role of ecology in speciation, which has led to a new focus on ecological speciation (e.g., Mayr 1947; Schluter

2000; Rundle and Nosil 2005; Thibert-Plante and Hendry 2009), that is, speciation driven by ecologically based divergent selection. Selection is divergent when opposite phenotypes are favored 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. Another novel development is a resurrection of arguments about the role of hybridization in speciation and adaptive radiation (Bullini 1994;

Arnold 1997; Seehausen 2004; Gompert et al. 2006; Mallet 2007). In particular, the role of homoploid hybrid speciation, i.e., hybrid speciation without change in ploidy level (also referred to as recombinational speciation, Grant 1985), is an issue receiving renewed interest and new empirical and theoretical support (Nice et al. 2002; Salzburger and Meyer 2004; Seehausen 2004; Gross and Rieseberg 2005; Gompert et al. 2006; Mavárez et al. 2006; Mavárez and Linares 2008).

An important question is what is really meant by hybrid speciation. We will call the result of introgressive hybridization a hybrid species if resulting hybrid traits directly and significantly contribute to the survival and reproductive isolation of the species (Mallet 2007; Mallet 2009). Note that we require neither that genomes of the parental species are represented in the hybrid species at equal frequencies nor that the hybrid species stably coexist in sympatry with the parental species. We find this relaxed view of hybrid speciation more useful than alternatives as it fits better the patterns observed in nature. Indeed, most suggested cases of homoploid hybrid speciation among animals appear to have involved a certain degree of backcrossing to one parental species (Mavárez and Linares 2008). Moreover, these authors also noted that about half of the homoploid hybrid taxa considered are allopatric with at least one parental species.

Although the support for the importance of ecological speciation and homoploid hybrid speciation in nature is growing, many questions remain. These concern the conditions for speciation, its time-scales, driving forces, the importance of ecological and genetic details, the role of geography, and so on. Answering these questions requires more detailed data on already studied cases, new empirical studies, and extensions of the quantitative theory of speciation.

One recent trend in theoretical research is to use complex numerical simulations tailored to particular case studies to supplement and provide additional insights into those that have emerged from relatively simple analytical models (Gavrilets 2004). For example, recently Gavrillets et al. (2007), Gavrillets and Vose (2007), and Sadedin et al. (2009) used detailed simulation models incorporating relevant ecological, behavioral, spatial, and genetic data to examine putative cases of ecological speciation of cichlids in a crater lake (Barluenga et al. 2006), of palms on an oceanic island (Savolainen et al. 2006), and of snails on sea shores (Hollander et al. 2005; Hollander et al. 2006). By simulating these real systems, they were able to address certain questions about ecological speciation in general (e.g., whether sympatric speciation is achieved easily as it is claimed occasionally) 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). They also looked at specific questions such as: what does mathematical theory tell us about the plausibility, speed, and patterns of (sympatric) speciation in the case studies?

What are the important parameters and processes controlling the dynamics of speciation? How common are the phenomena observed in these case studies? They also identified some 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 and to make the interpretation of empirical findings less speculative.

In contrast to a significant effort to understand ecological speciation, recombinational speciation has so far received only very limited attention from theoreticians. Two previous theoretical papers by McCarthy et al. (1995) and Buerkle et al. (2000) have utilized mathematical models to help understand the dynamics of hybrid speciation in plants. McCarthy et al. (1995) used spatially explicit individually based simulations to study the consequences of hybridization between two parental forms that differed in two chromosomal rearrangements. These authors assumed that one homozygous hybrid genotype (“hybrid species”) had higher fertility whereas all heterozygous hybrid genotypes had smaller fertility than both parental genotypes. McCarthy et al. (1995) analyzed how the waiting time until the advantageous recombinant type is fixed in the population depended on parameters (such as the strength of selection, the rate of selfing, and the size of the area where hybridization occurs). McCarthy et al. (1995) noted that even when the relative fertility of  $F_1$  hybrids was very low (e.g., less than 2%), the new species was established in just a few hundred generations. Decreasing the hybrid species advantage markedly increased the amount of selfing required for rapid speciation, but when the advantage was sufficiently large, speciation was observed even for obligate outcrossers. Increasing the size of the area in which hybridization occurs significantly reduced the time to speciation. In the simulations of McCarthy et al. (1995), the environment was spatially homogeneous and the new recombinant species, once emerged, always replaced the parental forms. Buerkle et al. (2000) considered a more complex situation with a spatially heterogeneous environment and an additional “ecological” trait controlled by two unlinked additive diallelic loci. Specifically, they assumed a gradient in viability selection so that small values of the ecological trait (which were characteristic of one parental form) were advantageous in one habitat whereas large values of the trait (which were the characteristic of another parental form) were advantageous in the other habitat. The intermediate values of the ecological trait resulted in higher viability in a relatively narrow intermediate zone positioned between the parental habitats. Buerkle et al. (2000) showed that under certain conditions the hybrid species can appear and be stably maintained in the intermediate zone simultaneously with the parental forms being preserved in their respective habitats. The time scale for speciation was on the order of several hundred to several thousand generations. An analysis of neutral markers showed a significant

gene flow and loss of differentiation in neutral loci between the species.

Here, we build on the previous theoretical studies of hybrid speciation (McCarthy et al. 1995; Buerkle et al. 2000) and ecological speciation (Gavrilets et al. 2007; Gavrilets and Vose 2007; Sadedin et al. 2009) to analyze a tentative case of ecological hybrid speciation in *Heliconius* butterflies (Salazar et al. 2005; Mavárez et al. 2006). This case has received a lot of attention after the publication of a paper by Mavárez et al. (2006) arguing in favor of hybrid speciation of *Heliconius heurippa* from *H. melpomene* and *H. cydno* in South America. Our model attempts to account for empirical patterns and data on genetic incompatibility, mating preferences and selection by predation (both based on coloration patterns), habitat preference, and local adaptation for all three *Heliconius* species. Using this model, we study the likelihood of recombinational speciation and identify the effects of various ecological and genetic parameters on the dynamics, patterns, and consequences of hybrid ecological speciation.

## Empirical Evidence

In describing empirical evidence, we will closely follow Mavárez et al. (2006) and references therein. *Heliconius cydno* and *H. melpomene* are two closely related species that overlap extensively in lower Mesoamerica and the Northern range of the Andes (i.e., Venezuela, Colombia, and Ecuador). Speciation in these butterflies does not involve any change in chromosome number (Brown et al. 1992) but it is instead associated with shifts in wing color patterns that generate assortative mating as well as postzygotic isolation due to predation-mediated selection (McMillan et al. 1997; Mallet et al. 1998; Jiggins et al. 2004). Different geographic races of *H. cydno* have black wings with white and yellow marks, whereas races of *H. melpomene* have black wings with red, yellow, and orange marks. Both species exhibit strong positive assortative mating based on their wing patterns (Jiggins et al. 2001; Mavárez et al. 2006) and also differ in habitat use (*H. cydno* in forest understory, *H. melpomene* in forest gaps/open areas) and the degree of preference for host plants in the genus *Passiflora* (low preference in *H. cydno*, high preference in *H. melpomene*). In spite of these differences, interspecific hybrids still occur in the wild throughout their sympatric range, usually at low frequencies ( $<10^{-3}$ , Mallet et al. 1998) but sometimes representing a significant fraction of the population ( $\sim 10\%$ , Mavárez et al. 2006). The tentative hybrid species, *H. heurippa*, is ecologically most similar to *H. cydno*, which it replaces geographically in the eastern Andes of Colombia. The wing pattern of *H. heurippa* has elements of those of the geographically close races of *H. cydno* and *H. melpomene*. Its hind-wing is indistinguishable from that of sympatric *H. m. melpomene*, whereas its forewing shows a

mixture of elements of both *H. m. melpomene* and of parapatric *H. cydno cordula*.

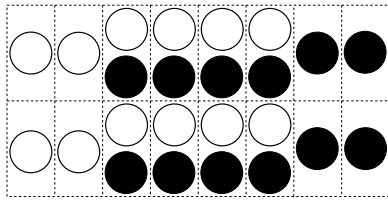
Genetic evidence for a hybrid origin of *H. heurippa* comes from analysis of polymorphism at the two nuclear genes *Invected* and *Distal-less*. These genes show no allele sharing between *H. cydno* and *H. melpomene*, whereas the *H. heurippa* genome appears as an admixture, sharing allelic variation from both putative parental species. Moreover, by performing interspecific crosses between *H. cydno* and *H. melpomene* in the laboratory, Mavárez et al. (2006) were able to recover a color pattern phenotype almost indistinguishable from that of wild *H. heurippa* in three generations, which provides some insight into the possible steps of introgressive hybridization that could have given rise to this species. Furthermore, the laboratory reconstructed pattern breeds true when crossed among themselves and with wild *H. heurippa* individuals.

Most interspecific crosses in *Heliconius* follow Haldane's rule for sterility, that is  $F_1$  females, which are the heterogametic sex in Lepidoptera, are sterile. In particular, crosses between *H. cydno* and *H. melpomene* produce  $F_1$  female sterility in both directions of the cross. On the other hand,  $F_1$  female hybrids between *H. heurippa* and *H. melpomene* show asymmetrical sterility. The female offspring of a cross between a female *H. heurippa* and male *H. melpomene* are completely sterile, either failing to lay eggs or laying eggs that never hatch. In contrast, the reciprocal cross produces female offspring that lay fertile eggs. Finally, crosses between *H. cydno* and *H. heurippa* produce fertile  $F_1$  females in both directions. These results are also compatible with the hypothesis of hybrid origin for *H. heurippa* (Salazar et al. 2005). In addition, backcrosses using  $F_1$  hybrid males provide evidence for a large Z(X)-chromosome effect on sterility and for recessive autosomal sterility factors as predicted by the dominance theory (Jiggins et al. 2001; Naisbit et al. 2002; Salazar et al. 2005).

Mavárez et al. (2006) also studied genetic isolation on several sympatric and allopatric populations of the three species using Bayesian assignment analysis performed with 12 microsatellite loci. The three species clearly represent different genetic entities, even in sympatry, which suggests the operation of some form of reproductive isolation. To further explore this idea, Mavárez et al. (2006) tested the degree to which *H. heurippa* is isolated from *H. melpomene* and *H. cydno* by assortative mating. No-choice mating experiments were conducted and showed that both males and females courted their co-specifics up to 10 times more often than individuals from other species. In mating experiments with choice, there was similarly strong assortative mating, although occasional matings between *H. heurippa* and *H. cydno* were observed.

## Model Description

The following describes the major components of our model.



**Figure 1.** Schematic representation of the spatial structure of the system. Each rectangle represents a patch. Open circles represent hosts in the open habitat; filled circles represent hosts in the closed-canopy habitat. Note the presence of both types of hosts in the central area.

*Space and environment*

Space is subdivided into an  $S_x \times S_y$  rectangular array of “patches” each representing roughly one square kilometer of forest. In Central America, *H. melpomene* are found in open habitat, *H. cydno* are found in closed-canopy forest (Smiley 1978; Estrada and Jiggins 2002), and *H. heurippa* uses the same habitat as *H. cydno* (J. Mavárez, pers. obs.). In our model, the environment changes in the horizontal direction. We assume that both habitats (denoted as 0 and 1) are present in the  $S$  central columns of patches, whereas the  $(S_x - S)/2$  leftmost columns and  $(S_x - S)/2$  rightmost columns of patches have only habitat 0 (“open”) and habitat 1 (“closed-canopy”), respectively (Fig. 1). Each habitat has a number of host plants that can be used for oviposition. Because the central area has more host plants present, the population density there will be higher as well (see below). To reduce boundary effects, we wrap the rectangle into a tube so that the top and the bottom rows become neighbors. Time is discrete and generations are nonoverlapping.

*Individuals*

Individuals are sexual and diploid and have discrete sexes. Each individual has a number of various phenotypic characters. To capture the wing color pattern differences between *H. melpomene*, *H. cydno*, and *H. heurippa* we assume, following Mavárez et al. (2006), two unlinked diallelic loci (with alleles  $A, a$  and  $B, b$ ) controlling the presence and intensity of red and yellow bands on the dorsal forewing. Let  $i$  and  $j$  be the numbers of “red” ( $A$ ) and “yellow” ( $B$ ) alleles in the genotype ( $i, j = 0, 1, \text{ or } 2$ ). Then, the intensities of red and yellow colors are characterized by color intensity matrices

$$R = \begin{pmatrix} & bb & Bb & BB \\ aa & 0 & 0 & 0 \\ Aa & 1 & 1 & 1/2 \\ AA & 1 & 1 & 1 \end{pmatrix}, \quad Y = \begin{pmatrix} & bb & Bb & BB \\ aa & 0 & 1/2 & 1 \\ Aa & 0 & 0 & 1 \\ AA & 0 & 0 & 1 \end{pmatrix}.$$

These matrices capture epistatic interactions between the loci. The two ones in the middle row of the  $R$ -matrix account for strong

expression of red color in some  $Aa$  heterozygotes. Specifically, red allele  $A$  behaves as dominant in  $bb$  and  $Bb$  individuals, but its effect is additive in  $BB$  individuals. The two zeroes in the middle column of the  $Y$ -matrix account for the absence of yellow color in some  $Bb$  heterozygotes (notice the lack of yellow band in the right box in Fig. 2A in the Mavárez et al. 2006). Specifically, yellow allele  $B$  behaves as recessive in  $Aa$  and  $AA$  individuals, but its effect is additive in  $aa$  individuals. Under this parametrization, *H. melpomene* is represented by genotype ( $i = 2, j = 0$ ) with  $R_{20} = 1, Y_{20} = 0$  (bottom left corner of the color intensity matrices), *H. cydno* by genotype ( $0, 2$ ) with  $R_{02} = 0, Y_{02} = 1$  (top right corner), and *H. heurippa* by genotype ( $2, 2$ ) with  $R_{22} = 1, Y_{22} = 1$  (bottom right corner). We note that there are several known color pattern genes that differentiate *H. melpomene* and *H. cydno*; we focus on only two of them for simplicity.

In Lepidoptera, males are the homogametic sex (and have two  $Z$  chromosomes) whereas females are the heterogametic sex (and have one  $Z$  and one  $W$  chromosome). To account for Haldane’s rule patterns observed among the three species (Salazar et al. 2005; see Introduction), we assume that hybrid female sterility is due to incompatibility between  $Z$ -linked and autosomal genes (Orr 1997). Specifically, we postulate that there are three types of sex chromosome  $Z$  ( $Z_1, Z_2, Z_3$ ) and a diallelic autosomal locus with alleles  $C_1$  and  $C_2$  such that relative female fertilities are given by a matrix

	$C_1C_1$	$C_1C_2$	$C_2C_2$
$Z_1$	1	0	0
$Z_2$	1	1	1
$Z_3$	0	0	1

Because the matrix concerns females exclusively, the  $W$  sex chromosome is omitted. *Heliconius melpomene* and *H. cydno* females are represented by the top left and bottom right corners, respectively. *Heliconius heurippa* females are represented by genotype  $Z_2C_2C_2$ . This matrix simultaneously captures the bidirectional Haldane’s rule observed in hybrids between *H. melpomene* and *H. cydno* (so that all  $F_1$  hybrid females are sterile) and the one-directional Haldane’s rule observed in hybrids between *H. melpomene* and *H. heurippa* (so that only the  $F_1$  female offspring of male *H. melpomene*  $\times$  female *H. heurippa* crossings are sterile) (Salazar et al. 2005; Mavárez et al. 2006). To keep the model’s complexity at a reasonable level, we neglect a possibility that fertilities of fertile females differ and that there may be more autosomal loci involved in reproductive isolation. All males are assumed to be fully fertile. We note that although there seems to be some sterility in males from certain hybrids crosses, it is definitely not as strong as in females.

To account for mating preferences we assume the existence of two additive quantitative characters  $p_x$  and  $p_y$  controlling males’

preference for the presence in females of the red and yellow fore-wing bands, respectively. To allow for habitat preference and local adaptation, we assume the existence of two additive quantitative characters: a “habitat preference” character  $h$  controlling the probability of choosing one or another habitat and an “ecological” character  $x$  controlling the adaptation to host. The traits  $p_r, p_y, h$ , and  $x$  are scaled between 0 and 1 and are controlled by different unlinked diallelic loci with equal effects.

In addition, there are 32 unlinked neutral loci with 256 possible alleles subject to stepwise mutation (Ohta and Kimura 1973) at rate  $10^{-3}$  per locus per generation. These loci are used to evaluate the levels of genetic divergence within and between species that one would observe if using microsatellite markers. Each of the parental species is initialized with a specific allele at these loci to allow the origin of neutral markers in the hybrids to be determined later.

### Life cycle

The life cycle consists of: (1) formation of mating pairs, (2) offspring production, (3) viability selection in eggs due to selection for local adaptation; (4) selection in adults due to predation, and (5) dispersal. In *Heliconius*, both selection by predation and dispersal occur simultaneously. However, in our simulations for computational purposes, these two processes were serialized. The results presented below were obtained assuming selection before dispersal. Limited simulations suggest that the alternative ordering leads to similar results.

### Habitat preference

Following Gavrillets and Vose (2005), the relative preference of an individual with habitat preference trait  $h$  for habitat 0 (i.e., the open habitat) is given by a linear function of  $h$ :

$$p = \varepsilon + (1 - 2\varepsilon)h, \quad (1)$$

where  $\varepsilon$  is a small positive parameter. The value of  $p$  changes linearly from  $\varepsilon$  for  $h = 0$  to  $1 - \varepsilon$  for  $h = 1$ . The value of  $\varepsilon$  can be interpreted as the probability that an individual with a perfect preference for one habitat mistakenly goes to the other habitat. The relative preference of this same individual for habitat 1 (i.e., the closed-canopy habitat) is  $1 - p$ .

### Reproduction

Mating occurs between individuals in the same “patch.” If a “patch” has only one habitat type, each male has an equal probability to encounter each female. If a “patch” has both habitat types, each male encounters a female with a probability proportional to the similarity of their preferred habitats. That is, a male and a female with habitat preference traits  $h_m$  and  $h_f$  encounter or do not encounter each other with probabilities  $p_m p_f + (1 - p_m)(1 - p_f)$  and  $p_m(1 - p_f) + p_f(1 - p_m)$ , respectively, where  $p_m$  and  $p_f$

are given by equation (1) with an appropriate value of  $h$ . Note that the rate of hybridization is an emergent character controlled by the current values of ecological and preference traits. Given an encounter, the actual mating occurs with a probability proportional to the male’s preference for the female’s color pattern. The preference of a male with preference traits  $p_r, p_y$  for a female with color pattern  $(i, j)$  is

$$\psi(i, j | p_r, p_y) = \exp[\alpha_r(p_r - 1/2)R_{ij}] \exp[\alpha_y(p_y - 1/2)Y_{ij}], \quad (2)$$

where  $R_{ij}$  and  $Y_{ij}$  are the corresponding elements of the color intensity matrices and  $\alpha_r$  and  $\alpha_y$  are positive parameters measuring the strength of preference for red and yellow, respectively. Large values of  $\alpha$ ’s imply strong preferences; small values imply weak preferences. This model represents a special case of an open-ended mating preference (Lande 1982; Gavrillets 2004, Chap. 9). Note that under our choice of the preference function (2), each of the three species has the highest mating preference for its own type (Naisbit et al. 2001; Mavárez et al. 2006; Melo et al. 2009).

Each mating results in a clutch of offspring whose size is drawn from a Poisson distribution with parameter  $b$ . We assume that all adult females mate once. This assumption implies that any costs of mate choice are absent, and that the effective population size is increased relative to the actual number of adults (Gavrillets and Vose 2005). The clutch is laid on a host with a probability proportional to the preference  $p$  for host. Mutations occur at a constant rate  $\mu$  per gene per generation in all loci except for the sex-chromosome  $Z$  and the autosomal locus  $C$ , which do not mutate.

### Adaptation to host plants

In our model, adaptation concerns the ability to grow on the host plant. The probability that the clutch laid by a mother with an ecological trait  $x$  survives on a host is

$$w = \exp\left[-\frac{(x - \theta)^2}{2\sigma_s^2}\right], \quad (3)$$

where  $\theta$  is the optimum phenotype, which is set to 0 in hosts of the open habitat, and to 1 in hosts of the closed-canopy habitat, and  $\sigma_s$  is a positive parameter measuring the strength of selection. This form of ecological selection introduces constraints on local adaptation so that offspring cannot have high fitness on both hosts. Smaller values of  $\sigma_s$  imply stronger selection for local adaptation and stronger selection against generalists (i.e., individuals with ecological trait  $x \sim 0.5$ ) and specialists for the alternative habitat. To account for the population size control by the number of host plants available, we assume that the number of adults surviving selection for local adaptation is  $K_0$  in habitat 0 and  $K_1$  in habitat 1 where the carrying capacities  $K_0$  and  $K_1$  are parameters of the model.

### Selection by predation

Adults are subject to selection by differential predation. We assume that predators (i.e., birds) stop eating butterflies (which are distasteful) once they learn that a particular color or a color pattern is associated with bad taste. The efficiency of learning process depends both on the number of butterflies eaten and the intensity of the stimuli (i.e., color and taste). We posit that all butterflies are equally distasteful.

Two approaches are used here. The first assumes that birds learn to avoid aposematic prey by separating the different color elements of butterfly wing patterns. For a bird that has eaten  $k$  butterflies, we say its “learning scores” for red and yellow colors are

$$Q_r = \sum_{j=1}^k R_j, \quad Q_y = \sum_{j=1}^k Y_j, \quad (4)$$

where  $R_j$  and  $Y_j$  are the appropriate elements (“color intensities”) of matrices  $R$  and  $Y$  for the  $j$ th butterfly. We assume that the bird stops eating the butterflies of a particular color once the corresponding learning score reaches a positive threshold  $Q$ . Small values of  $Q$  imply that a small number of butterflies will be eaten so that selection by predation is weak; large values of  $Q$  imply strong selection. The model assumes that more intense colors (i.e., with larger values of  $R$  and  $Y$ ) are easier to remember. This simple model is closely related to the ideas on the avoidance of distasteful prey going back to Müller (1879) (see Mallet 2001). Note that once one color (red or yellow) is learned, selection on the other color in mixed-color individuals disappears. We assume that parameter  $Q$  is the same for all patches. This assumption implies that the density of predators is uniform across the whole system.

The second approach assumes that color patterns (i.e., combinations of colors) are learned as a whole. By combining matrices  $R$  and  $Y$ , we get six unique patterns; a learning score  $Q$  is introduced for each of the six color patterns. The learning process happens as before with each eaten butterfly contributing 1 to the corresponding learning score of a predator for the color pattern.

We note that the birds are not modeled explicitly but only implicitly (via the learning thresholds) and that no evolution in birds was allowed.

### Dispersal

Individuals surviving predation disperse. Butterflies are highly mobile and, thus, biologically realistic values of the migration rates are large. We assume that each individual migrates to one of the eight neighboring patches or stays in its native patch with the same probability  $1/9$ . For the left-most and right-most rows of patches that have five rather than eight neighboring patches, a proportion  $3/9$  of individuals are removed (reflecting their migration outside the system).

### Initial conditions

The peripheral patches that contain a single habitat type are initially populated with individuals perfectly adapted to and with the highest preference for this habitat, and with a genetic and phenotypic compositions of either *H. melpomene* (on the left side) or *H. cydno* (on the right side). The patches in the central area with both habitats were empty. Ninety-five percent of the individuals are homozygous for all genes, but the remaining 5% have an “intermediate” sex chromosome (i.e.,  $Z_2$ ) to account for an incomplete segregation between the two parental species.

### Parameters

To analyze our model, we used individual-based simulations, which we ran for 50,000 generations, roughly corresponding to 17,000 years. We systematically varied four parameters: the carrying capacity of habitat 1 ( $K_1 = 200$  and  $400$ ), clutch size ( $b = 4, 8,$  and  $16$ ), the learning threshold ( $Q = 1, 4,$  and  $16$ ), and the strength of selection for local adaptation ( $\sigma_s = 0.1, 0.25,$  and  $0.4$ ). The following summarizes the parameters that did not change: system size  $96 \times 4$ , the width of the central area 32 patches, the carrying capacity of habitat 0  $K_0 = 200$ , mutation rate  $\mu = 10^{-5}$  per generation, number of loci per trait 8, strength of color preference  $\alpha_r = \alpha_y = 4$ , and probability of habitat choosing error  $\varepsilon_h = 0.001$ . The reason for including higher carrying capacity  $K_1 = 400$  in habitat 1 is that some observations suggest that *H. cydno* has more hosts available than *H. melpomene* (J. Mavárez, pers. obs.).

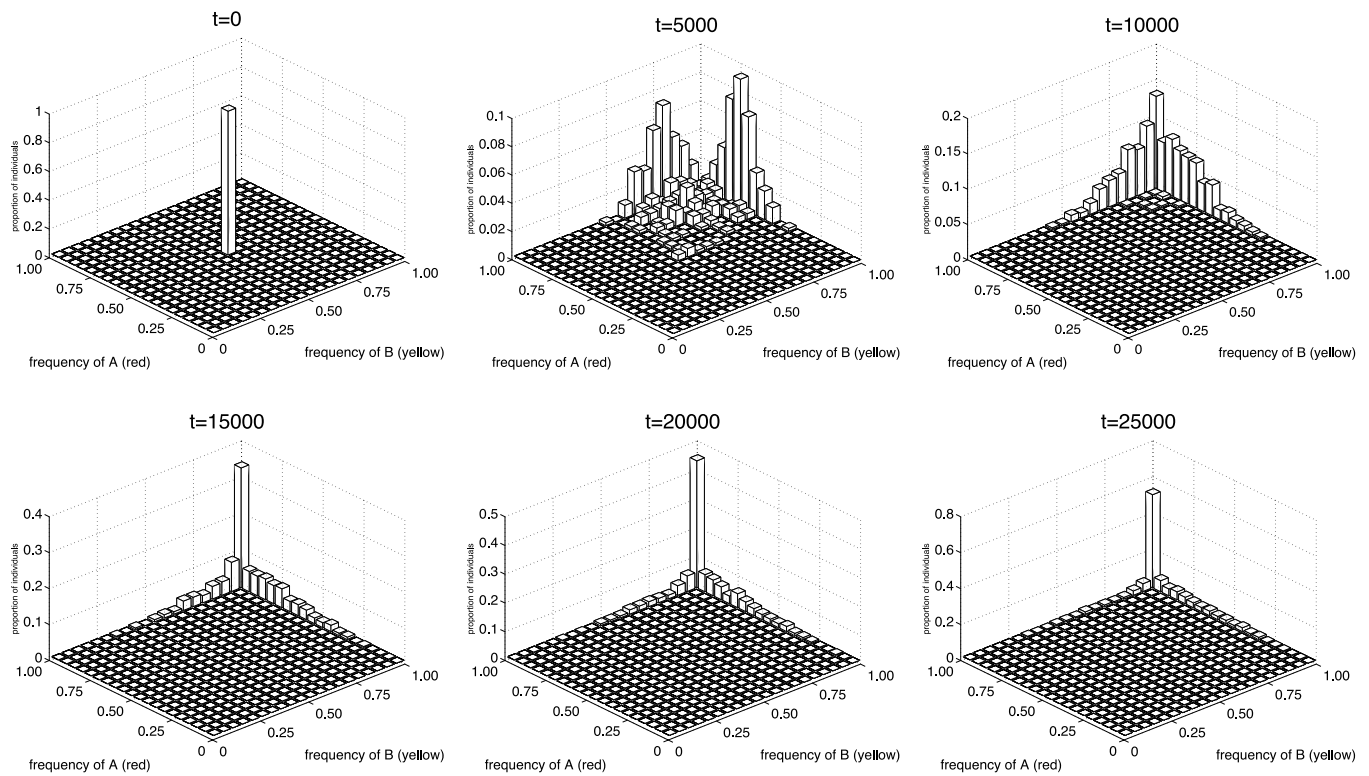
System-level data were saved every 100th generation; patch-level data were saved every 1000th generation. The results below are based on 20 runs for each parameter combination. The model was implemented in C. The code is available upon request.

## Results

After the start of each simulation run, individuals of both parental species rapidly spread through the initially uninhabited central area from the opposite sides and then start hybridizing, initially at low frequencies (The initial probability of hybridization was of order  $10^{-4}$  and the overall proportion of individuals with foreign neutral alleles during the first 1000 generations was no larger than several percent.). In the course of the simulations, we observed several different dynamic regimes, and we will discuss them according to observed changes in (1) coloration patterns, (2) habitat use and local adaptation, and (3) reproductive compatibility. First, we discuss our results corresponding to the case when birds learn different colors independently.

### The dynamics of coloration patterns

In the simulations, there was a strong tendency for fixation of both red (A) and yellow (B) alleles across the whole system so that a



**Figure 2.** The dynamics of the frequencies of A (red) and B (yellow) alleles across the whole system and all runs for the first 25,000 generations.

*Heurippa*-like color pattern eventually dominated (Fig. 2). These dynamics are explained by two factors. First, natural selection by predation favors the establishment of aposematic coloration and, thus, increases the frequencies of the warning coloration alleles. Second, because parental species already have mating preference alleles for the corresponding colors, hybrids acquire them as well. As a result, sexual selection favors the fixation of the “red” and “yellow” alleles via a Fisherian runaway process (Lande 1981; Andersson 1994). In fact, the color preference traits  $p_r$  and  $p_y$  approach the maximum possible value of one across the whole system much faster than either of the color alleles approach fixation: on average, in 3600 generations. In contrast, the fixation of a first color allele (red or yellow) takes place on average in 6300 generations. The second color locus approaches fixation much later—on average about 16,100 generations later. The reason for this discrepancy is that although the initial evolution of the color loci is driven both by natural and sexual selection, once one color locus is close to fixation, natural selection by predation is not acting on the second locus anymore and its evolution is mostly driven by mating preferences (and random genetic drift).

In the color locus that is fixed first, the polymorphism is less likely to be observed in natural populations (see Table 1 for the frequencies at which the red and yellow alleles are first to be fixed for different parameter values). If the carrying capacities of both hosts are equal (i.e.,  $K_0 = K_1 = 200$ ), fixation of the

red allele (and, consequently, long-lasting polymorphism in the yellow locus) is favored by an intermediate clutch size ( $b = 8.0$ ). It is also favored by strong selection for local adaptation (i.e., low values of  $\sigma_s$ ) coupled with extreme values of the learning threshold  $Q$  ( $Q = 1$  and 16). If the carrying capacity of habitat 1 is twice that of habitat 0 (i.e.,  $K_1 = 400$ ), the yellow allele is almost always fixed first whereas the red locus remains polymorphic. This is expected because initially the system as a whole has twice as many  $B$  alleles as  $A$  alleles.

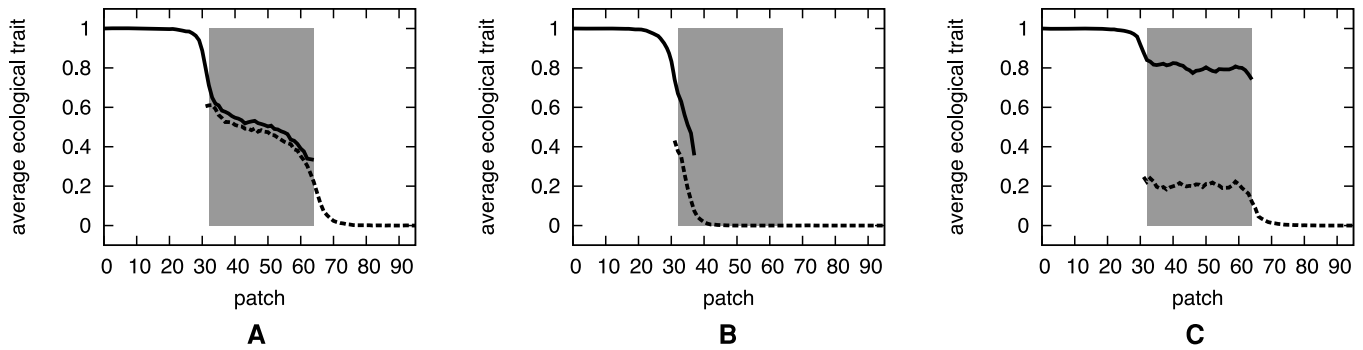
#### *The dynamics of ecological traits*

Populations inhabiting the peripheral areas with only one host type present always remain close to the original values of the ecological traits (and, correspondingly, remain locally adapted) independently of the strength of selection for local adaptation  $\sigma_s$ . The situation, however, is different in the central area, where both types of hosts are available for butterflies to mate and oviposit. Three different long-term outcomes are observed. First, most or all individuals in the central area have intermediate values of the ecological trait close to 0.5 (Fig. 3A). These are the generalists utilizing both host species. At the boundaries of the central and peripheral areas, two tension zones are formed separating the populations of generalists in the central area and specialists in the peripheral areas (A tension zone as a hybrid zone maintained by a balance of dispersal of parental forms and selection against

**Table 1.** Numbers of different outcomes in the dynamics of coloration patterns, ecological traits, and reproductive compatibility (20 runs for each parameter combinations). "Red"—allele A fixed first, "yellow"—allele B fixed first. "General," "Tens," and "Symp" correspond to the formation of the generalist ecotype, a tension zone separating two specialist ecotypes, and two sympatrically distributed specialist ecotypes, respectively.  $Z_2$ ,  $Z_1$  &  $Z_3$ , and  $Z_2$  &  $Z_3$ —signify fixation of  $Z_2$ , to the loss of  $Z_2$ , and to the fixation of  $C_2$ , respectively. The stars mark the combinations of parameters discussed later in the text.

Parameters				Color		Ecotypes			Z chromosomes			
$K_1$	$b$	$Q$	$\sigma_s$	Red	Yellow	General	Tens	Symp	$Z_2$ fixed	$Z_2$ lost	$C_2$ fixed	
200	4.0	1	0.10	13	7	0	13	7	14	6	0	
			0.25	8	12	0	20	0	16	4	0	
			0.40	5	15	0	20	0	17	3	0	
		4	0.10	10	10	0	17	3	18	2	0	
			0.25	14	6	0	20	0	15	4	1	
			0.40	9	11	0	20	0	19	0	1	
		16	0.10	11	9	0	18	2	17	2	0	
			0.25	11	9	0	20	0	18	1	0	
			0.40	15	5	0	20	0	18	1	1	
		8.0	1	0.10*	9	11	0	2	18	13	7	0
				0.25*	4	16	0	17	3	17	3	0
				0.40*	9	11	18	2	0	18	2	0
	4		0.10	14	6	0	4	16	18	2	0	
			0.25	8	12	0	20	0	20	0	0	
			0.40	12	8	20	0	0	20	0	0	
	16		0.10	16	4	0	2	18	18	2	0	
			0.25	12	8	0	19	1	18	1	0	
			0.40	14	6	19	1	0	19	1	0	
	16.0		1	0.10	11	9	0	0	20	19	1	0
				0.25	9	11	5	13	2	18	2	0
				0.40	6	14	20	0	0	16	4	0
		4	0.10	11	9	0	0	20	17	3	0	
			0.25	8	12	9	11	0	20	0	0	
			0.40	10	10	20	0	0	20	0	0	
16		0.10	13	7	0	0	20	19	1	0		
		0.25	13	7	7	13	0	20	0	0		
		0.40	9	11	20	0	0	20	0	0		
400		4.0	1	0.10	0	20	0	20	0	8	3	9
				0.25	0	20	0	20	0	11	3	6
				0.40	0	20	0	20	0	16	1	3
	4		0.10	0	20	0	20	0	5	1	14	
			0.25	0	20	0	20	0	13	1	6	
			0.40	0	20	0	20	0	20	0	0	
	16		0.10	0	20	0	20	0	2	4	14	
			0.25	0	20	0	20	0	12	0	8	
			0.40	0	20	0	20	0	20	0	0	
	8.0		1	0.10	0	20	0	20	0	17	0	3
				0.25*	0	20	0	20	0	19	0	1
				0.40	0	20	0	20	0	19	1	0
		4	0.10	0	20	0	20	0	18	0	2	
			0.25	0	20	0	20	0	20	0	0	
			0.40	0	20	0	20	0	20	0	0	
		16	0.10	0	20	0	20	0	19	0	1	
			0.25	0	20	0	20	0	20	0	0	
			0.40	1	19	0	20	0	20	0	0	
		16.0	1	0.10	3	17	0	16	4	17	3	0
				0.25	0	20	0	20	0	20	0	0
				0.40	0	20	0	20	0	16	4	0
	4		0.10	0	20	0	16	4	20	0	0	
			0.25	0	20	0	20	0	20	0	0	
			0.40	1	19	0	20	0	20	0	0	
16	0.10		0	20	0	17	3	20	0	0		
	0.25		0	20	0	20	0	20	0	0		
	0.40		1	19	0	20	0	20	0	0		





**Figure 3.** Three different outcomes in the dynamics of the average ecological character  $x$  per habitat across space. (A) Central area is inhabited by a population of generalists. (B) Two specialists separated by a tension zone at a border of the central area. (C) Two specialists coexisting throughout the central area. The gray area marks the central patches. Parameter values used to produce the figures: (A)  $K_1 = 200$ ,  $b = 16$ ,  $Q = 1$ ,  $\sigma_s = 0.4$ , (B)  $K_1 = 200$ ,  $b = 8$ ,  $Q = 4$ ,  $\sigma_s = 0.25$ , (C)  $K_1 = 200$ ,  $b = 8$ ,  $Q = 4$ ,  $\sigma_s = 0.1$ . The two lines in the central region show average trait values computed over two hosts.

hybrids.] Second, a single tension zone is formed in the central area that then moves and eventually stabilizes at a boundary of the central area. This tension zone separates two specialist populations and largely prevents the utilization of one host in the central area (Fig. 3B). Finally, two specialist populations (each adapted to and utilizing its own host) coexist sympatrically through the central area (Fig. 3C). Note that in this case the traits values in the central area are shifted away from the optimum values because of the gene flow between the specialists due to hybridization. In all three cases, the habitat preference trait  $h$  closely matches the ecological trait across the whole system.

Consider first the situation when the carrying capacities of both hosts are equal (i.e.,  $K_0 = K_1 = 200$ ). In this case, the most common outcome is the formation of a tension zone separating two specialists (see Table 1, Fig. 3B). Generalists (as in Fig. 3A) are more likely to emerge if the birth rate  $b$  and the coefficient  $\sigma_s$  are large so that selection for local adaptation is relatively weak. Sympatric coexistence of two specialists in the central area (as in Fig. 3C) was observed only few times and required selection for local adaptation to be the strongest (that is,  $\sigma_s = 0.1$ ). If the carrying capacity of habitat 1 is twice that of habitat 0 (i.e.,  $K_1 = 400$ ), generalists were never observed. Two sympatric specialists coexisting in the central area were observed only a couple of times, leaving the formation of a tension zone separating two specialists (Fig. 3B) as the most common outcome.

#### The dynamics of reproductive compatibility

Recall that the simulations started with two parental species almost fixed for the reproductively isolated combinations of sex chromosomes and autosomal genes  $Z_1C_1C_1$  and  $Z_3C_2C_2$ , respectively, with only a small proportion of the *H. heurippa*-type chromosome  $Z_2$  present in both parental species. At the end of the simulations, four situations were observed. First,  $Z_2$  can be lost, with the reproductively isolated ancestral combinations  $Z_1C_1C_1$  and

$Z_3C_2C_2$  being both present in the system. Second, *H. heurippa*-type  $Z_2$  can be fixed across the whole system, whereas the autosomal locus  $C$  can exhibit neutral polymorphism with both  $C_1$  and  $C_2$  alleles present. Third, the *H. melpomene*-type autosomal allele  $C_1$  can be fixed whereas sex chromosomes  $Z_1$  and  $Z_2$  are both maintained in a neutral fashion. Finally, the *H. cydno*-type autosomal allele  $C_2$  can be fixed whereas sex chromosomes  $Z_3$  and  $Z_2$  are both maintained in a neutral fashion. In the last three cases, all genotypes present in the system are reproductively compatible.

If the carrying capacities of both hosts are equal (i.e.,  $K_0 = K_1 = 200$ ), the most common outcome is the fixation of *H. heurippa*-type sex chromosome  $Z_2$  in the whole system (about 90% of cases). The loss of  $Z_2$  that was the second most common outcome (almost 9% of cases) was promoted by smaller birth rate  $b$  and stronger selection for local adaptation (i.e., smaller  $\sigma_s$ ). After the loss of the  $Z_2$  chromosome, a future hybrid speciation event is unlikely. The fixation of *H. melpomene*-type allele  $C_1$  or *H. cydno*-type allele  $C_2$  with a neutral polymorphism in the  $Z$  chromosome was observed only a handful of times.

If the carrying capacity of habitat 1 is twice that of habitat 0 (i.e.,  $K_1 = 400$ ), the most common outcome again was the fixation of  $Z_2$ . However, now with a small birth rate (i.e.,  $b = 4$ ), the fixation of *H. cydno*-type allele  $C_2$  with the neutral coexistence of  $Z_3$  and  $Z_2$  was frequently observed for strong and intermediate strengths of selection for local adaptation (i.e.,  $\sigma_s = 0.1$  and 0.25). This is what is expected because of higher initial number of *H. cydno*-type alleles in the system. The loss of  $Z_2$  was observed few times (mostly with  $b = 4$ ) whereas the fixation of  $C_1$  with the neutral coexistence of  $Z_1$  and  $Z_2$  was never observed.

#### Joint dynamics

The hybrid aposematic coloration (i.e., the presence of both red and yellow bands) readily spreads across the whole system in all cases considered. This spread is preceded by the establishment

of strong mating preferences for both red and yellow colors. The spread of hybrid coloration is most often accompanied by the fixation of the hybrid chromosome  $Z_2$  in the whole system so that all individuals are reproductively compatible. With regard to the ecological and habitat preference characters, the two ecotypes initially present in the parental species (which have high preference for and local adaptation on hosts in habitat 0 and 1, respectively) are preserved and either coexist through the central area or one of them occupies most of this area. It is also possible that an additional generalist ecotype utilizing both habitats can emerge and spreads across the central area excluding both specialists.

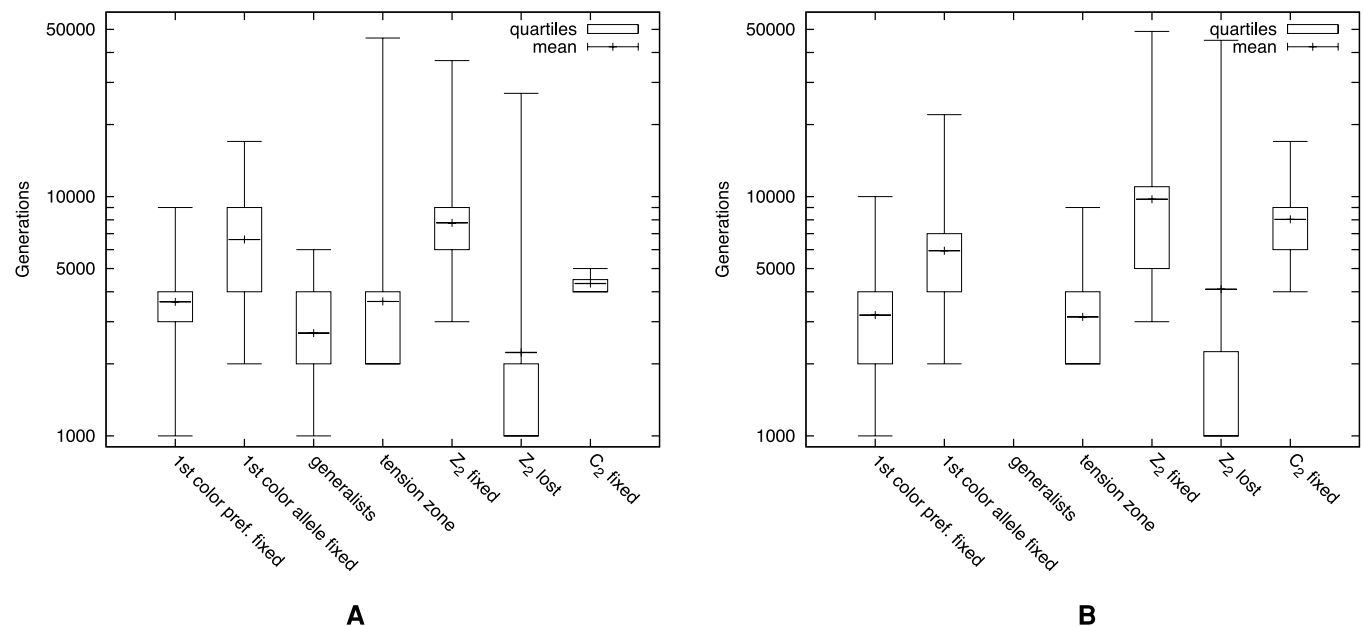
The loss of  $Z_2$  results in the presence of two genetically incompatible sets of populations with *H. melpomene*-type genotype  $Z_1C_1C_1$  and *H. cydno*-type genotype  $Z_3C_2C_2$ , both having *H. heurippa*-type coloration and mating preferences. Most commonly, these populations are separated by a tension zone located at a border of the central area but in a few cases they are sympatric in this area. In few other cases, one of the two sets of populations forms a generalist ecotype that spreads across the central area excluding populations of the other type and the specialist ecotype of its own type.

The fixation of *H. cydno*-type autosome  $C_2$  occurring simultaneously with the neutral coexistence of  $Z_2$  and  $Z_3$  was accompanied by the separation of the two specialist ecotypes by a tension zone located at a border of the central area.

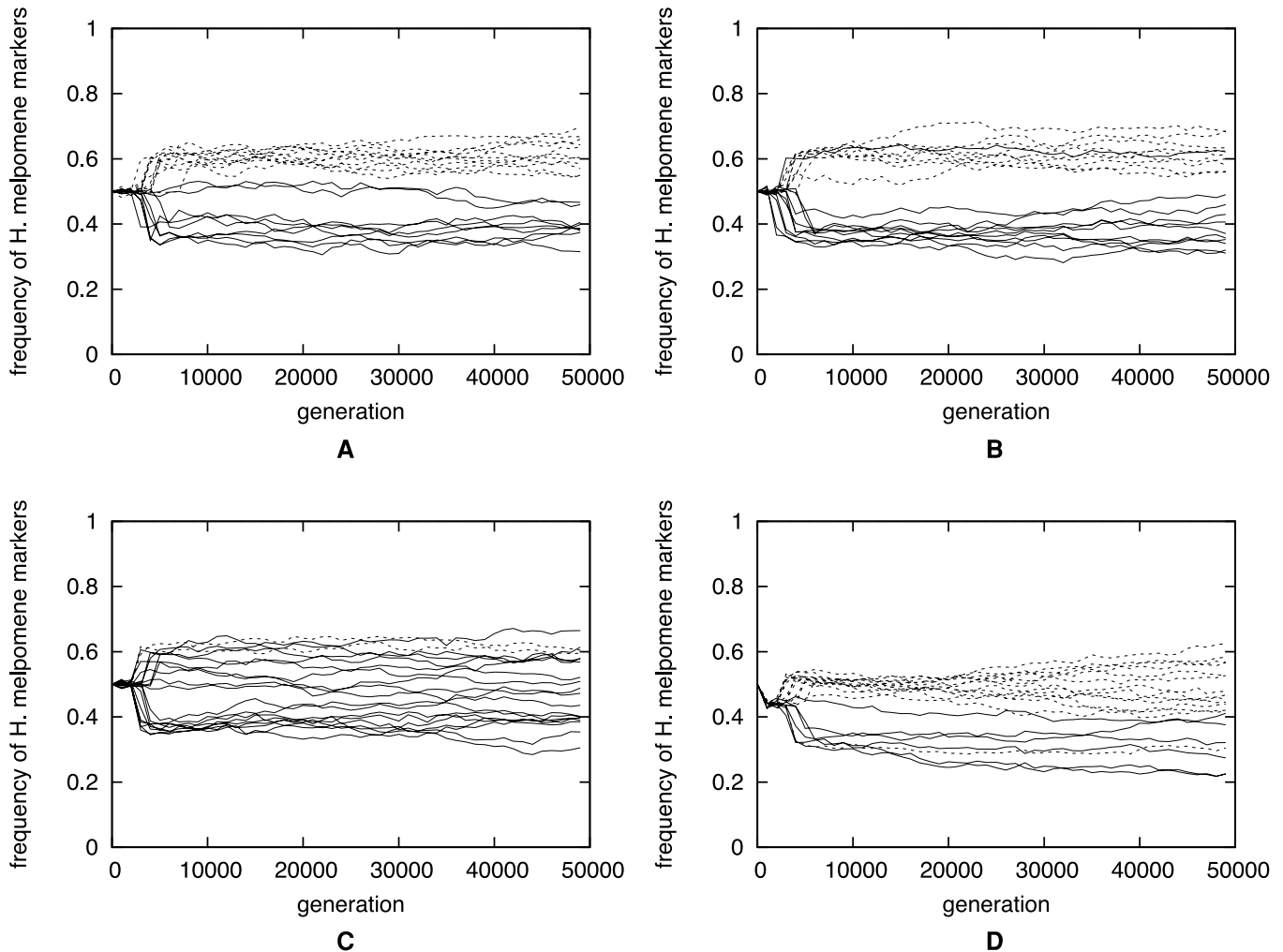
Figure 4 shows the typical time scales in our simulations. If *H. heurippa*-type chromosome  $Z_2$  is destined to be lost, this happens relatively fast, typically during the second thousand of

generations. Independently of this, within a few thousand generations after the start of the simulation both mating preference traits evolve to extreme values and the ecological and the habitat preference traits approach their equilibrium distributions. Few thousand generations later, one of the color pattern alleles approaches fixation and the *H. heurippa*-type sex chromosome  $Z_2$  achieves high frequency (if  $Z_2$  is destined to be fixed). The other color allele locus approaches fixation far more slowly (data not shown). The growth in the frequencies of the color alleles and  $Z_2$  chromosome start at the same border of the central area and then spreads throughout the whole system.

Technical comment: We note that here the time for a color allele to be “fixed” means the time for this allele to reach a frequency of 95% in the whole system (We use 95% rather than 100% because of the recurrent mutation.). The time for a particular sex chromosome to be “fixed” or “lost” means the time for the chromosome to reach a frequency of 100% or 0%. To make conclusions about ecological traits, we use a more convoluted procedure. First, we identify all patches in the central area in which each host plant contributes at least 20% of the patch’s population. We say that a “generalist” is formed if among these patches there is a band of patches of at least six-patch wide in which the trait values are close to 1/2 (specifically, the average ecological trait of the individuals born in environment type 0 is above 0.4 and for those born in type 1 is below 0.6). We say that there is sympatric coexistence of two ecomorphs if there is a band of patches at least six-patch wide in which the trait values are away from 1/2 (specifically, the average ecological trait of the individuals born



**Figure 4.** Tukey’s plots of the time scales associated with different outcomes marked on the x-axis. The y-axis represents time in generations on a logarithmic scale. The vertical lines extend from minimum to maximum observations, the middle dashed lines depict averages, and the boxes extend from lower to upper quartiles. (A)  $K_1 = 200$ . (B)  $K_1 = 400$ .



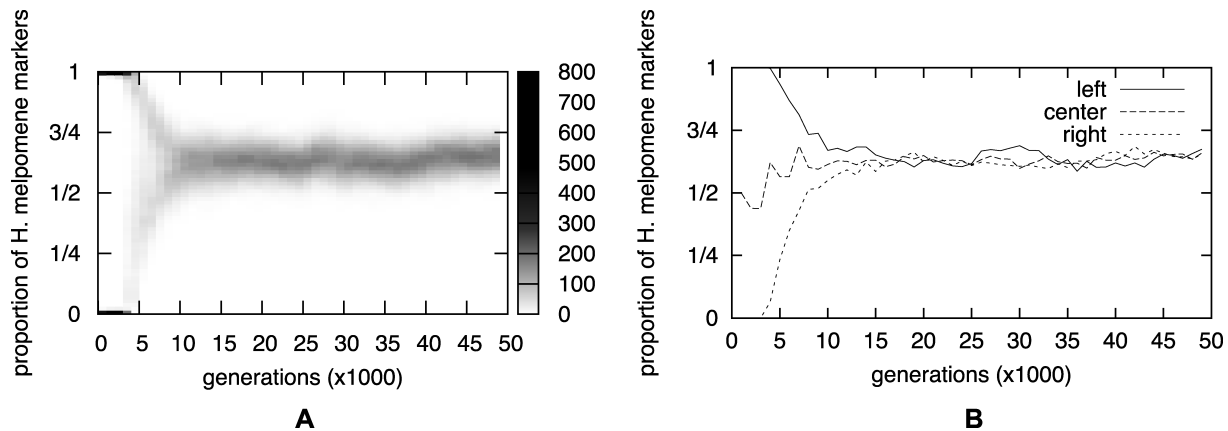
**Figure 5.** The dynamics of the frequency of neutral markers of *H. melpomene* origin for four parameter combinations (marked with a star in Table 1). In all figures,  $Q = 1$  and  $b = 8.0$ . (A)  $K_1 = 200$ ,  $\sigma_s = 0.1$  (B)  $K_1 = 200$ ,  $\sigma_s = 0.25$  (C)  $K_1 = 200$ ,  $\sigma_s = 0.4$  (D)  $K_1 = 400$ ,  $\sigma_s = 0.25$ . Twenty runs for each combination of parameters. Solid lines: runs in which the preference for yellow color trait  $p_y$  reached 0.99 first. Dashed lines: runs in which the preference for red color trait  $p_r$  reached 0.99 first.

on host 0 is below 0.4 and for those born on host 1 is above 0.6). We say there is a “tension zone” otherwise. That is, either the width of the band of patches with two populations is five patches or less, or there are demes with high average ecological character (on one side of the tension zone) and demes with low average value (on the other side).

#### The dynamics of neutral markers

The above dynamics were accompanied by the introgression in neutral markers. Figure 5 shows the frequency of neutral markers of *H. melpomene* origin across the whole system estimated by randomly sampling 16 individuals from each of the 96 patches in the first row of patches. We used four parameter combinations that showed a diversity of outcomes in the dynamics of coloration, habitat usage, and reproductive compatibility (see Table 1 where the combinations of parameters used are marked with a star).

One can see that at the onset of hybridization the frequency of *H. melpomene* markers experiences a rapid shift away from its initial value of 0.5. This shift is apparently explained by a strong sexual selection operating during a relatively short time interval when mating preference for one color has already reached the highest value whereas the preference for the second color is still intermediate. During this time interval, individuals with the maximum expression of the second color have mating advantage that results in increasing the frequency of their neutral markers in the population. The explanation just given works for the data shown in Figure 5A,B and D but not for those in Figures 5C. This suggests that some additional factors are also in place. Note that the downward shift in the frequency of neutral markers of *H. melpomene* origin observed in Figures 5D is explained by the fact that for this parameter combination initially there are twice as many *H. cydno* individuals as *H. melpomene* individuals. After



**Figure 6.** The dynamics of neutral markers of *H. melpomene* origin in a run with  $K_1 = 200$ ,  $Q = 1$ ,  $b = 8.0$ , and  $\sigma_s = 0.1$ . (A) The distribution of the frequency of neutral markers of *H. melpomene* origin over time. The intensity of the black color is proportional to the number of individuals with the corresponding frequency (the scale is given by the color bar on the right). (B) The average frequencies of neutral markers of *H. melpomene* origin in four leftmost, four centrally located, and four rightmost demes.

both mating preference traits have reached the maximum value of one, the dynamics of neutral markers appear to be driven mostly by random genetic drift.

Figure 6A gives more details on the dynamics of neutral markers using a numerical run with parameter values as in Figure 5A. Shown is the whole distribution of the frequency of the neutral markers of *H. melpomene* origin (rather than just its average value as Fig. 5). The distribution becomes very wide soon after the onset of hybridization. Figure 6B shows the corresponding dynamics of the frequencies of the neutral markers of *H. melpomene* origin estimated in three groups of four demes each of the first row: four leftmost demes, four centrally located demes, and four rightmost demes. This figure shows that it takes a long time for neutral markers to spread across the central area and then to reach the most peripheral patches.

#### Some transient dynamics

In no case did we observe a stable coexistence of the three color patterns of parental and hybrid species. Was there some kind of transient coexistence that did not persist to the end of our simulations? To answer this question, we identified time intervals and demes in which the two parental and a hybrid phenotypes were present at frequencies higher than 10% (chosen to be larger than the maximum observed proportion of natural hybrids of 8%). Across all runs, this transient period starts on average around generation 2200 (SD = 1200) and no later than generation 5000. It lasts on average 900 generations (SD = 800), with a maximum of 5000 generations. An analysis of variances shows that both the starting generation and duration of the transient period decrease with increasing the strength of predation (i.e., decreasing the value of  $Q$ ). The duration of the transient period is further affected by the clutch size; extreme values of  $b$  (i.e., 4 and 16) promote longer transients.

#### Color patterns learned as a whole

In this case, novel combinations of colors are under strong negative selection by predation. As a result, hybrids never achieve high frequencies,  $Z_2$  chromosome disappears from the system, and both parental species are preserved and coexist sympatrically in the central area (see Fig. 3C) for all combinations of parameters considered (data not shown).

## Discussion

In this article, we used a spatial individual-based multilocus model to study a proposed case of homoploid hybrid speciation involving two sister species, *H. cydno* and *H. melpomene*, and a putative hybrid species *H. heurippa* (Mavárez et al. 2006). Assessing the mechanisms potentially driving hybrid speciation and the plausibility of its different scenarios can help us better understand both the origins and evolution of *Heliconius* butterflies and the challenges faced by empiricists studying homoploid hybrid speciation in animals. Our results show that if birds learn different colors independently, then *H. heurippa*-type coloration pattern and the mating preferences for this pattern as well as *H. heurippa*-type sex chromosome can readily spread across the whole system within ten thousand generations after the initial contact between the parental species. The resulting hybrid species retains the two ecotypes of the parental species using two different hosts and can also develop an additional generalist ecotype utilizing both hosts. The emerging hybrid species has neutral markers from both parental species at substantial frequencies. It is also possible that the two parental species converge on the same hybrid coloration and mating preferences and undergo substantial exchange in neutral markers while retaining their original sex chromosomes and ecotypes. This outcome can be interpreted as the evolution of Mullerian mimicry as a result of hybridization. However, empirical

data provide little support for this alternative outcome given that most *Heliconius* sister species pairs (i.e., able to hybridize) tend to belong to different mimicry rings. On the other hand, species in the same mimicry ring are usually very distantly related and unable to hybridize (Beltrán et al. 2000).

The spread of hybrid coloration is largely driven by selection for warning coloration by predation. It is also helped by the preexisting (in the parental species) mating preferences for bright colors that pass to the hybrids (Melo et al. 2009) and then coevolve with coloration patterns in a process analogous to Fisherian runaway. In the *Heliconius* butterflies, coloration plays the role of a “magic trait” (Gavrilets 2004; Kronforst et al. 2006; Puebla et al. 2007; Jiggins 2008; Jiggins et al. 2008; Salazar et al. 2008) that controls both viability and mating. The existence of magic traits greatly simplifies nonallopatric speciation in general (Gavrilets 2004; Puebla et al. 2007) and in butterflies in particular (Mallet 2007; Jiggins 2008; Jiggins et al. 2008; Salazar et al. 2008). The evolution of ecotypes is driven by selection for local adaptation. If this selection is weak, a generalist ecotype utilizing both host can emerge in the central area and expel the specialist ecotypes from this area (cf. Gavrilets et al. 2007; Gavrilets and Vose 2007). If selection is strong, the gene flow from one specialist to another is small and does not prevent the colonization of the whole central area by both ecotypes. For intermediate strengths of selection, the formation of the generalist ecomorph is thwarted and simultaneously the gene flow between the two specialist ecomorphs is strong enough to prevent their coexistence in the patches located in the central area. As a result, one specialist ecomorphs largely excludes the other one from the central area. This happens as the tension zone moves stochastically across the central area until it gets trapped at its boundary where selection regime changes (Barton 1983; Barton 1999). The explanation of why *H. heurippa*-type sex chromosome  $Z_2$  spreads through the system is a bit more convoluted. After the mating preferences for both red and yellow colors have spread through the system, mating becomes random with respect to sex chromosomes. In this case, while individuals with *H. melpomene*-type chromosome  $Z_1$  and *H. cydno*-type chromosome  $Z_3$  are mutually incompatible (in the sense that female hybrid offspring is infertile), those with  $Z_2$  are partially compatible with both parental types. As a result, individuals with  $Z_2$  chromosome have on average more fertile offspring than those with either  $Z_1$  or  $Z_3$ . This ultimately drives  $Z_2$  chromosome toward fixation.

In contrast, if different color patterns are learned as a whole, then selection by predation prevents any significant accumulation of hybrids in the system and hybrid speciation does not occur.

Recently, Salazar et al. (2008) and Jiggins et al. (2008) have contrasted two modes of homoploid hybrid speciation that they called “hybrid trait speciation” and “mosaic genome hybrid speciation.” In the former mode, speciation occurs through the es-

tablishment via hybridization of a novel adaptive trait. The novel trait must also confer a degree of reproductive isolation from the parental lineages. In the later mode, speciation involves stabilization of a hybrid genome that initially contains a large number of intrinsically incompatible genes. The two modes, thus, mostly differ in the type of selection involved (i.e., ecological vs. intrinsic incompatibility). Jiggins et al. (2008) also expected them to differ in the proportion of the genome involved with hybrid trait speciation resulting in introgression of a handful of genes and in mosaic genome hybrid speciation resulting in massive introgression (as in the case of sunflowers, e.g., Ungerer et al. 1998). According to Salazar et al. (2008) and Jiggins et al. (2008), who analyzed genetic divergence between *H. melpomene*, *H. cydno*, and *H. heurippa* in several loci, this case fits the hybrid trait speciation mode. Our data do not support the argument of Salazar et al. (2008) and Jiggins et al. (2008). In our model, hybrid speciation is indeed mostly driven by evolutionary processes involving a single trait—coloration. However, the extent of hybridization in neutral loci is very extensive and is more compatible with that expected under mosaic genome hybrid speciation. This suggests that discriminating between hybrid trait speciation and mosaic genome hybrid speciation may not be easy.

A question of the speed of ecological speciation is of great importance and interest (Hendry et al. 2007; Gavrilets et al. 2007; Gavrilets and Vose 2007; Thibert-Plante and Hendry 2009) with most models suggesting that speciation (i.e., evolution of significant reproductive isolation) could happen on the time scale of a few hundred to a few thousand generations. The data in Mavárez et al. (2006) have sometimes been interpreted as suggesting that hybrid speciation could happen within a few generations. Our model does not support this interpretation. Indeed in our simulations, the typical time scales of the evolution of Mullerian mimicry and of hybrid speciation are on the order of several thousand generations. In our simulations, the initial rate of hybridization is very low. If it were higher, the hybrid species would evolve faster. The rate of hybrid speciation is expected to increase if the hybrids utilize a novel habitat (Nolte et al. 2005; Schwarz et al. 2007) that is free of the parental forms.

In our simulations, we never observed the coexistence of the two parental and the hybrid species. The same mechanisms that facilitated the emergence and establishment of the hybrid species also lead to the spread of the hybrid species traits across the whole system. As a result, the parental combinations of traits disappear (colors always,  $Z$  chromosomes most of the time). The same outcome is expected after a secondary contact between the hybrid and a parental species. Therefore, our model suggests that the coexistence of the three *Heliconius* species in South America may be transient rather than stable. An alternative, and likely more reasonable, explanation is that our model lacks certain components crucial for the coexistence of the three species. One possibility

is spatial and temporal heterogeneity of the biotic environment that can result in a partial fragmentation of the species ranges (e.g., if the density of hosts varies dramatically in space) and/or spatial variation of selection (both for adaptation to hosts and by predation). This heterogeneity will reduce the opportunity for gene flow and might ensure coexistence. Alternatively, selection by predation can act via some other mechanism that would increase geographic variation in the system. For example, it is possible that parental colorations are protected by selection if both parental species have some additional Mullerian mimics, which is indeed the case in Heliconiinae butterflies. Evolutionary dynamics of Mullerian mimicry can be complex especially if different members of a Mullerian ring differ in the degree of unpalatability (Gavrilets and Hastings 1998). This may further complicate the dynamics of hybrid speciation and make predictions difficult. Other factors that may affect the dynamics are selection for local adaptation and the genetic architecture of the traits considered. In the model we made the simplest assumptions of additive genetics and unlinked genes. More complex schemes can of course alter the dynamics.

Our model also predicts the appearance of *H. heurippa*-type individuals utilizing *H. melpomene*-type hosts and of a generalist ecotype utilizing both hosts. Apparently, neither of these two outcomes has been observed in nature. Although evolution of the generalist is predicted to happen only under some conditions (specifically, under weak selection for local adaptation), some utilization of the open habitat initially utilized by *H. melpomene* occurs in the model always. It is well known that although tension zones represent a barrier to the neutral genes (e.g., Bengtsson 1985; Barton and Bengtsson 1986; Gavrilets 1997; Gavrilets and Hastings 1998), advantageous genes will pass through the tension zones easily. What factors would prevent the spread of *H. heurippa* coloration across *H. melpomene*-like ecotype once hybridization starts is unclear but one possibility is that the signaling quality of aposematic color patterns varies with the habitat type or the environmental background (Sweeney et al. 2003). The *H. heurippa* color pattern might therefore be less fit to the open habitat areas typically occupied by *H. melpomene*.

Unfortunately, the lack of data does not yet allow one to come up with a mathematical model that would have a better explanatory power. More and better data are needed to that end. From the theoretical perspective, at this stage the most crucial would be to understand better how selection by predation operates in the system, how birds learn the colors, and to what extent they distinguish different components of color patterns. Of course any data on genetics of the traits considered including those controlling reproductive compatibility would greatly increase the realism of the model. The same effect will be achieved by more precise information of the spatial distributions and densities of the parental

species and their co-mimics, and on the degree of temporal stability of *H. heurippa* geographic distribution. Without such data it will remain difficult to make definite conclusions on the origins of *H. heurippa*.

However, one can attempt to identify possible scenarios from both empirical data and mathematical models. *Heliconius heurippa* is well differentiated from both *H. melpomene* and *H. cydno* in some microsatellite loci (J. Mavarez, unpubl. data). In addition, the current range of *H. heurippa* is well outside of the area of sympatry of *H. melpomene* and *H. cydno*. If one accepts the hybrid origin of *H. heurippa*, these two observations suggest both long-lasting geographic isolation of *H. heurippa* from its parental species and significant changes in the species ranges since speciation. Overall, our model supports the possibility of hybrid origin of *H. heurippa*. The most plausible scenario would include hybridization between *H. melpomene* and *H. cydno* in an area geographically isolated from the rest of both parental species with subsequent long-lasting geographic isolation of the new hybrid species, followed by changes in the species ranges, the secondary contact, and the disappearance of *H. melpomene*-like ecomorph in the hybrid species. Our model does not rule out completely an alternative scenario of the appearance of the red allele A in a geographically isolated *H. cydno* population by mutation, with the subsequent fixation of the *H. heurippa*-type color. However, this scenario appears less likely at least for two reasons. First, it would require a very rare mutational event that would bring the red allele. Second, it would require an additional sequence of unlikely events that would fix this red allele by drift in the presence of positive frequency-dependent selection due to predation acting to reduce the frequency of this initially rare allele (Mallet and Barton 1989). Although such stochastic peak shifts can occur, their probability is very small (Barton and Charlesworth 1984; Gavrilets 2004).

Our model highlights two particular outcomes of hybridization that can be of general importance. The first concerns the evolution of Mullerian mimicry when a common hybrid coloration pattern spreads across a system of sympatric or parapatric species as a result of hybridization. In our simulations, this outcome corresponds to the loss of  $Z_2$  chromosome. The second is that the deleterious gene flow resulting from hybridization can prevent the spread of an ecomorph into a suitable area and, thus, can limit the species range. In our simulations, this outcome corresponds to the formation of a tension zone separating two specialist ecotypes at a border of the central area. The question of whether deleterious gene flow can restrict species ranges is currently of great interest (e.g., Kirkpatrick and Barton 1997; Case and Taper 2000). In contrast to current discussions, in our model the source of the deleterious gene flow is not the central populations of the same species but rather a different species or ecotype occupying adjacent areas.

Our model was intentionally tailored for a particular case study. Some features of the *Heliconious* system are probably unique (e.g., specific genetics of the color patterns). Others may not be very common in nature (e.g., the fact that color patterns play an important role simultaneously in survival and mating [Mallet and Barton 1989; Jiggins et al. 2001; Mavárez et al. 2006], and thus represent “magic traits” in the terminology of Gavrillets [2004]). At the same time, our model points to some general features of hybrid speciation, that is speciation via introgressive hybridization with hybrid traits directly and significantly contributing to the survival and reproductive isolation. Hybrid speciation can be triggered by a number of factors increasing the likelihood of hybridization and can be driven by strong ecological selection promoting the spread of a particular advantageous combination of hybrid traits. The model suggests that stable sympatric coexistence of the hybrid and parental forms should not be generally expected. Indeed, if a particular hybrid combination of traits is advantageous, it is expected to spread after a secondary contact between a hybrid and parental species (if hybridization between them is possible). In other words, the same process that leads to the creation of a hybrid species may prevent its stable sympatric coexistence with the parental forms. The model also shows that equal representation of parental neutral markers in the hybrid species is unlikely. In fact, even parental traits of ecological importance can be inherited in a hybrid species in a relatively random fashion. The model highlights the importance of preexisting assortative mating and habitat segregation between hybrids and at least one of the parental forms in simplifying the conditions for hybrid speciation. It also shows that recombinational hybrid speciation is not an instantaneous process but rather can take hundreds and thousands of generations.

In contrast to many other speciation scenarios (Gavrillets 2004), theoretical work on hybrid speciation has been very limited (McCarthy et al. 1995; Buerkle et al. 2000 and this article) so it is difficult to identify the most important evolutionary factors and forces controlling its dynamics. Still existing work suggests that the questions of spatial heterogeneity in selection, coexistence of hybrid species with their parental species, and of the stability of species ranges are of great importance for developing an adequate theory. Currently, mathematical models support the belief that hybrid speciation in animals is plausible under certain conditions. However much more work (both empirical and theoretical) is necessary to be able to make more definite conclusions on its importance in nature.

#### ACKNOWLEDGMENTS

We thank B. M. Fitzpatrick, J. A. Fordyce, A. Hendry, J. Mallet, and M. Servedio for useful comments on the manuscript. This work is supported by National Institutes of Health grant GM56693.

#### LITERATURE CITED

- Andersson, M. B. 1994. Sexual selection. Princeton Univ. Press, Princeton, NJ.
- Arnold, M. L. 1997. Natural hybridization and evolution. Oxford Univ. Press, Oxford.
- Barluenga, M., K. N. Stoltz, W. Salzburger, M. Muschick, and A. Meyer. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439:719–723.
- Barton, N. H. 1983. Multilocus clines. *Evolution* 37:454–471.
- . 1999. Clines in polygenic traits. *Genet. Res.* 74:223–236.
- Barton, N. H., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridizing populations. *Heredity* 56:357–376.
- Barton, N. H., and B. Charlesworth. 1984. Genetic revolutions, founder effects, and speciation. *Annu. Rev. Ecol. Syst.* 15:133–164.
- Beltrán, M., C. D. Jiggins, V. Bull, M. Linares, J. Mallet, W. O. McMillan, and E. Bermingham. 2000. Phylogenetic discordance at the species boundary: comparative gene genealogies among rapidly radiating *Heliconius* butterflies. *Mol. Biol. Evol.* 19:2176–2190.
- Bengtsson, B. O. 1985. The flow of genes through a genetic barrier. Pp. 31–42 in J. J. Greenwood, P. H. Harvey, and M. Slatkin, eds. *Evolution essays in honor of John Maynard Smith*. Cambridge Univ. Press, Cambridge.
- Bolnick, D. I., and B. M. Fitzpatrick. 2007. Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Syst.* 38:459–487.
- Brown, K. S., T. C. Emmel, P. J. Eliazar, and E. Suomalainen. 1992. Evolutionary patterns in chromosome numbers in neotropical lepidoptera. *Heredity* 117:109–125.
- Buerkle, C. A., R. J. Morris, M. A. Asmussen, and L. H. Rieseberg. 2000. The likelihood of homoploid hybrid speciation. *Heredity* 84:441–451.
- Bullini, L. 1994. Origin and evolution of animal hybrid species. *Trends Ecol. Evol.* 9:422–426.
- Case, T., and M. Taper. 2000. Interspecific competition, environmental gradients, gene flow, and the coevolution of species’ borders. *Am. Nat.* 155:583–605.
- Coyne, J., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Inc., Sunderland, MA.
- Dieckmann, U., M. Doebeli, J. A. J. Metz, and D. Tautz. 2004. *Adaptive speciation*. Cambridge Univ. Press, Cambridge.
- Estrada, C., and C. D. Jiggins. 2002. Patterns of pollen feeding and habitat preference among *Heliconius* species. *Ecol. Entomol.* 27:448–456.
- Gavrillets, S. 1997. Hybrid zones with epistatic selection of Dobzhansky type. *Evolution* 51:1027–1035.
- . 2004. *Fitness landscapes and the origin of species*. Princeton Univ. Press, Princeton, NJ.
- Gavrillets, S., and A. Hastings. 1998. Coevolutionary chase in two-species systems with applications to mimicry. *J. Theor. Biol.* 191:415–427.
- Gavrillets, S., and J. Losos. 2009. Adaptive radiation: contrasting theory with data. *Science* 323:732–737.
- Gavrillets, S., and A. Vose. 2005. Dynamic patterns of adaptive radiation. *Proc. Natl. Acad. Sci. USA* 102:18040–18045.
- . 2007. Case studies and mathematical models of ecological speciation. 2. Palms on an oceanic island. *Mol. Ecol.* 16:2910–2921.
- Gavrillets, S., A. Vose, M. Barluenga, W. Salzburger, and A. Meyer. 2007. Case studies and mathematical models of ecological speciation. 1. Cichlids in a crater lake. *Mol. Ecol.* 16:2893–2909.
- Gompert, Z., J. A. Fordyce, M. L. Froister, A. M. Shapiro, and C. C. Nice. 2006. Homoploid hybrid speciation in an extreme habitat. *Science* 314:1923–1925.
- Grant, V. 1985. *The evolutionary Process*. Columbia Univ. Press, New York.
- Gross, B. L., and L. H. Rieseberg. 2005. The ecological genetics of homoploid hybrid speciation. *J. Heredity* 96:241–252.

- Hendry, A. P., P. Nosil, and L. H. Rieseberg. 2007. The speed of ecological speciation. *Funct. Ecol.* 21:455–464.
- Hollander, J., M. Lindgarth, and K. Johannesson. 2005. Local adaptation but not geographical separation promotes assortative mating in a snail. *Anim. Behav.* 70:1209–1219.
- Hollander, J., D. C., Adams, and K. Johannesson. 2006. Evolution of adaptation through allometric shifts in a marine snail. *Evolution* 60:2490–2497.
- Jiggins, C. D. 2008. Ecological speciation in mimetic butterflies. *BioScience* 58:541–548.
- Jiggins, C. D., R. E. Naisbit, R. L. Coe, and J. Mallet. 2001. Reproductive isolation caused by colour pattern mimicry. *Nature* 411:302–305.
- Jiggins, C. D., C. Estrada, and A. Rodriguez. 2004. Mimicry and the evolution of premating isolation in *Heliconius melpomene* Linnaeus. *J. Evol. Biol.* 17:680–691.
- Jiggins, C. D., C. Salazar, M. Linares, and J. Mavarez. 2008. Hybrid trait speciation in *heliconius* butterflies. *Philos. Trans. R. Soc. Lond. B* 363:3047–2054.
- Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species' range. *Am. Nat.* 150:1–23.
- Kronforst, M. R., L. G. Young, D. D. Kapan, C. McNeely, R. J. O'Neill, and L. Gilbert. 2006. Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proc. Natl. Acad. Sci. USA* 103:6575–6580.
- Lande, R. 1981. Models of speciation by sexual selection on polygenic characters. *Proc. Natl. Acad. Sci. USA* 78:3721–3725.
- . 1982. Rapid origin of sexual isolation and character divergence in a cline. *Evolution* 36:213–223.
- Mallet, J. 2001. Mimicry: an interface between psychology and evolution. *Proc. Natl. Acad. Sci. USA* 98:8928–8930.
- . 2007. Hybrid speciation. *Nature* 446:279–283.
- . 2009. Rapid speciation, hybridization and adaptive radiation in the *Heliconius melpomene* group. Cambridge Univ. Press, pp. 177–194.
- Mallet, J., and N. H. Barton. 1989. Inference from clines stabilized by frequency-dependent selection. *Genetics* 122:967–976.
- Mallet, J., W. O. McMillan, and C. Jiggins. 1998. Mimicry and warning color in the boundary between races and species. Pp. 390–403 in D. J. Howard, and S. H. Berlocher, eds. *Endless forms: species and speciation*, Oxford Univ. Press, New York.
- Mavárez, J., and M. Linares. 2008. Homoploid hybrid speciation in animals. *Mol. Ecol.* 17:4181–4185.
- Mavárez, J., C. A. Salazar, E. Bermingham, C., Salcedo, C. D. Jiggins, and M. Linares. 2006. Speciation by hybridization in *Heliconius* butterflies. *Nature* 441:868–871.
- Mayr, E. 1947. Ecological factors in speciation. *Evolution* 1:263–288.
- McCarthy, E. M., M. A. Asmussen, and W. W. Anderson. 1995. A theoretical assessment of recombinational speciation. *Heredity* 74:502–509.
- McMillan, W. O., C. D. Jiggins, and J. Mallet. 1997. What initiates speciation in passion-vine butterflies? *Proc. Natl. Acad. Sci. USA* 94:8628–8633.
- Melo, M. C., C. Salazar, C. D. Jiggins, and M. Linares. 2009. Assortative mating preferences among hybrids offers a route to hybrid speciation. *Evolution* 63:1660–1665.
- Müller, F. 1879. Ituna and Thyridia: a remarkable case of mimicry in butterflies. *Proc. Entomol. Soc. Lond.*, XX–XXIX.
- Naisbit, R. E., C. D. Jiggins, and J. Mallet. 2001. Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. *Proc. R. Soc. Lond. B* 268:1–6.
- Naisbit, R. E., C. D. Jiggins, M. Linares, C. Salazar, and J. Mallet. 2002. Hybrid sterility, Haldane's rule and speciation in *Heliconius cydno* and *H. melpomene*. *Genetics* 161:1517–1526.
- Nice, C. C., J. A. Fordyce, A. M. Shapiro, and R. French-Constant. 2002. Lack of evidence for reproductive isolation among ecologically specialized lycaenid butterflies. *Ecol. Entomol.* 27:702–712.
- Nolte, A. W., J. Freyhof, K. C. Stemshorn, and D. Tautz. 2005. An invasive lineage of sculpins, *Cottus sp (Pisces, Teleostei)* in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. *Proc. R. Soc. Lond. B* 272:2379–2387.
- Nosil, P. 2007. Divergent host plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking sticks. *Am. Nat.* 169:151–162.
- Ohta, T., and M. Kimura. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genet. Res.* 22:201–204.
- Orr, H. A. 1997. Dobzhansky, Bateson, and the genetics of speciation. *Genetics* 144:1331–1335.
- Puebla, O., E. Bermingham, F. Guichard, and E. Whiteman. 2007. Colour pattern as a single trait driving speciation in *Hypoplectrus* coral reef fishes? *Proc. R. Soc. Lond. B* 274:1265–1271.
- Rundle, H. D., and P. Nosil. 2005. Ecological speciation. *Ecol. Lett.* 8:336–352.
- Sadedin, S., J. Hollander, M. Panova, K. Johannesson, and S. Gavrillets. 2009. Case studies and mathematical models of ecological speciation 3: ecotype formation in a Swedish snail. *Mol. Ecol.*, *In press*.
- Salazar, C. A., C. D. Jiggins, C. A. Arias, A. Tobler, E. Bermingham, and M. Linares. 2005. Hybrid incompatibility is consistent with hybrid origin of *Heliconius heurippa* Hewitson from its close relatives, *Heliconius cydno* Doubleday and *Heliconius melpomene* Linnaeus. *J. Evol. Biol.* 18:247–256.
- Salazar, C. A., C. D. Jiggins, J. Taylor, M. R. Kronforst, and M. Linares. 2008. Gene flow and the genealogical history of *Heliconius heurippa*. *BMC Evol. Biol.* 8:Article 132.
- Salzburger, W., and A. Meyer. 2004. The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften* 91:277–290.
- Savolainen, V., M.-C. Anstett, C. Lexer, I. Hutton, J. J. Clarkson, M. V. Norup, M. P. Powell, D. Springate, N. Salamin, and W. J. Baker. 2006. Sympatric speciation in palms on an oceanic island. *Nature* 441:210–213.
- Schluter, D. 2000. *The ecology of adaptive radiation*. Oxford Univ. Press, Oxford.
- Seehausen, O. 2004. Hybridization and adaptive radiation. *Trends in Ecology and Evolution* 19:198–207.
- . 2006. African cichlid fish: a model system in adaptive radiation research. *Proc. R. Soc. Lond. B* 273:1987–1998.
- . 2007. Chance, historical contingency and ecological determinism jointly determine the rate of adaptive radiation. *Heredity* 99:361–363.
- Schwarz, D., B. M. Matta, N. L. Shakir-Butteri, and B. McPherson. 2007. Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature* 436:546–549.
- Smiley, J. 1978. Plant chemistry and evolution of host specificity—new evidence from *Heliconius* and *Passiflora*. *Science* 201:745–747.
- Sweeney, A., C. D. Jiggins, and S. Johnsen. 2003. Polarized light as a mating signal in a butterfly. *Nature* 423:31–32.
- Thibert-Plante, X., and A. Hendry. 2009. Five questions on ecological speciation addressed with individual-based simulations. *J. Evol. Biol.* 22:109–123.
- Ungerer, M. C., S. J. E. Baird, J. Pan, and L. H. Riesberg. 1998. Rapid hybrid speciation in wild sunflowers. *Proc. Natl. Acad. Sci. USA* 95:11757–11762.

Associate Editor: M. Servedio