

Evol Ecol (2014) 28:591–609
DOI 10.1007/s10682-014-9697-8

ORIGINAL PAPER

Sensory drive speciation and patterns of variation at selectively neutral genes

Tomotaka Matsumoto · Yohey Terai · Norihiro Okada ·
Hidenori Tachida

Received: 9 December 2013 / Accepted: 20 March 2014 / Published online: 1 April 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract Speciation by sensory drive can occur if divergent adaptation of sensory systems causes rapid evolution of mating traits and the resulting development of assortative mating. Previous theoretical studies have shown that sensory drive can cause rapid divergent adaptive evolution from one to two phenotypes. In this study, we examined two topics: the possibility of adaptive radiation by sensory drive from one to more than two phenotypes and the relationships of patterns of variation at selectively neutral genes to levels of viability selection, habitat and mating preferences and migration. We conducted individual-based simulations assuming a sensory trait and a mating trait controlled by a small number of loci. We found that adaptive radiation is possible when the number of loci controlling the sensory trait is small; the levels of viability selection, habitat and mating preferences are intermediate; and the emigration rate is high. We also found that emigration rates as well as the levels of habitat and mating preferences are related to F_{ST}

Electronic supplementary material The online version of this article (doi:[10.1007/s10682-014-9697-8](https://doi.org/10.1007/s10682-014-9697-8)) contains supplementary material, which is available to authorized users.

T. Matsumoto (✉)
Graduate School of Systems Life Sciences, Kyushu University, Fukuoka, Japan
e-mail: tomotakamatsumoto2@gmail.com

Present Address:

T. Matsumoto
Division of Evolutionary Genetics, National Institute of Genetics, Shizuoka, Japan

Y. Terai · N. Okada
Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Kanagawa, Japan

Present Address:

Y. Terai
Department of Biosystems Science, The Graduate University for Advanced Studies (Sokendai), Yokohama, Kanagawa, Japan

H. Tachida
Department of Biology, Faculty of Sciences, Kyushu University, Fukuoka, Japan

values at neutral loci, but F_{ST} proved to be insensitive to a small change in the number of loci controlling the mating trait. This suggests that an estimation of the past population history is possible without an accurate genetic model.

Keywords Theoretical study · Sensory drive speciation · Genetic differentiation between species · Cichlid · Adaptive radiation · Selectively neutral gene

Introduction

Understanding the mechanisms of speciation is one of the most important research areas in evolutionary biology. In particular, speciation driven by ecological adaptation (ecological speciation) has attracted the attention of many researchers because it can cause rapid speciation even with gene flow, and many studies have reported empirical examples of and theoretical possibilities for ecological speciation (Dieckmann and Doebeli 1999; Coyne and Orr 2004; Doebeli and Dieckmann 2005; Rundle and Nosil 2005; Gavrillets and Losos 2009; Schluter 2009).

“Sensory drive” speciation is one example of ecological speciation in which sensory traits that are directly adapted to specific habitats induce the establishment of assortative mating (Boughman 2002). Under sensory drive, a female is more likely to choose a male whose mating signature can be more easily detected by their sensory apparatus. For example, visual systems adapted to different light environments recognize different coloration patterns as signals for choosing mating partners. Such visually mediated sensory drive is well documented in fishes such as guppies (Endler and Houde 1995), sticklebacks (Boughman 2001) and cichlids (Terai et al. 2006; Seehausen et al. 2008).

Among these, the cichlids of Lake Victoria in East Africa, which consists of approximately 500 species (Seehausen 1996; Turner et al. 2001), are of particular interest. These species are monophyletic (Meyer et al. 1990; Meyer 1993; Verheyen et al. 2003), and they are thought to have undergone speciation within the last 10,000–15,000 years based on the geological history of the lake (Johnson et al. 1996, 2000; Beuning et al. 1997; Scholz et al. 1998; Seehausen 2002). Sensory drive has been hypothesized to be one of the important mechanisms of this rapid adaptive radiation (Terai et al. 2006; Seehausen et al. 2008).

Kawata et al. (2007) have shown that ecological speciation caused by sensory drive can theoretically occur rapidly if the selection pressure on the sensory trait is neither too strong nor too weak. Their results were supported empirically by Seehausen et al. (2008), who investigated cichlids in several regions of Lake Victoria, and found that divergent adaptation of opsins (or visual perception) to the different light environments in regions with intermediate light gradients across depths induced divergence of nuptial coloration in males, which caused reproductive isolation. In contrast, no such divergence was found in regions with steep or shallow light gradients. These studies shed light on the mechanism of speciation by visually mediated sensory drive.

In this study, we theoretically examine two questions not investigated by Kawata et al. (2007). First, although Kawata et al. (2007) found that a single species/type can rapidly diverge into two descent species/types, divergence into several types would be necessary to explain the adaptive radiation of the Lake Victoria cichlids. In fact, many more than two cichlid species have been found in continuous habitats (Seehausen and Bouton 1997), but it remains unknown whether they could have diverged by adaptive radiation on a short time scale. One might think that multiple locally adapted phenotypes could emerge by repeated

bifurcation of lineages, but the increased number of phenotypes would increase the migration load through the immigration of non-adapted phenotypes and make it more difficult for the adaptation to occur (Butlin et al. 2009, pp 81–85). Therefore, we hypothesized that the adaptive evolution of more than two phenotypes was much more difficult than that of two phenotypes, and we examined whether this can be caused by sensory drive. Second, we monitored the patterns of variation at selectively neutral genes in addition to the evolution of the genes involved in adaptation. Although Gavrilets et al. (2007) also examined variation at selectively neutral genes during ecological speciation theoretically, they considered a model in which the traits determining the viability and mating success were completely different. Therefore, their model is not considered to be that of sensory drive. In addition, we investigated the variation at selectively neutral loci in more detail by examining its correlation with model parameters such as past selection pressures or levels of migration. If the patterns of variation at selectively neutral genes change depending on past selection pressures or levels of migration, we may be able to infer these parameters by comparing the observed variation at selectively neutral genes with the expected variation obtained theoretically. Thus, observing the patterns of neutral variation would provide valuable information for understanding the process of adaptive radiation by sensory drive in nature.

To answer these two questions, we conducted individual-based simulations. While the previous theoretical studies assumed models with patches arranged in two circles laying one on top of another (Gavrilets et al. 2007) or two-dimensional structure (Kawata et al. 2007), we assumed a three-dimensional habitat representing one part of a lake, where individuals tend to migrate to their favorable patches with fitness dependent migration. Furthermore, we assumed that one sensory trait determines the viability and habitat preference of each individual and that a male whose mating trait is detectable by the sensory trait of females is more likely to mate with those females. We considered three factors related to fitness—viability selection, habitat preference and mating preference—and examined their effects and the migration effects on adaptive radiation and the pattern of variation at neutral genes.

Model

Habitat

We considered a cube-shaped habitat corresponding to a part of a lake's water column and divided it into $5 \times 5 \times 5$ patches. Each individual lives in one of these 125 patches. We assumed that the light environment changes gradually with depth. Consequently, the optimal phenotype for the sensory trait (the absorption spectrum of opsin in cichlids) also changes gradually. A detailed description of the fitness function for the sensory trait is given in the section on *viability selection*.

Individuals

Each individual is diploid and dioecious and is assumed to have L_S loci controlling an equal number of sensory traits and L_M loci controlling a mating trait (nuptial color in cichlids). Each locus has an infinite number of alleles, and the phenotypic contribution of each allele is represented by an integer ($\dots, -2, -1, 0, 1, 2, \dots$). We assumed a stepwise mutation model (Ohta and Kimura 1973) with mutation rates u_S and u_M at the sensory and

mating trait loci, respectively. When a mutation occurs, alleles change their value by +1 or −1 with a probability of one half. For the sensory traits, each individual has $2 \times L_S$ sensory trait values because each of the $2 \times L_S$ alleles is assumed to express an integer-valued phenotype independently. This modeling was chosen because of the biology of the opsin genes of East African cichlids and animal olfactory receptor genes (Carleton and Kocher 2001; Buck and Axel 1991; Ota et al. 2012, see Discussion). In contrast, a mating trait such as the body color of East African cichlids is determined by one phenotypic value. For example, female cichlids prefer reddish color in deep areas but bluish color in shallow areas (Seehausen et al. 2008). Thus, we assumed that the mating trait of an individual is determined by the sum of the integer-valued phenotypic contributions of its $2 \times L_M$ alleles. In addition, each individual has 50 selectively neutral loci. Each neutral gene/allele consists of 100 nucleotides, which evolve according to the Jukes-Cantor mutation model (Jukes and Cantor 1969, pp 21–132) with a per-site mutation rate of u_N . Finally, we assume free recombination between any pair of loci.

Generations are discrete and non-overlapping, and allele frequencies change during a given generation as a consequence of viability selection, habitat choice and mating choice, in this order. These three factors are explained in detail below.

Viability selection

The viability of each individual is determined by its sensory trait values and the population density in the patch. As explained above, we assume five different depths. Let θ_d be the optimal sensory trait value at depth d , and assume $\theta_d = d$. We define d_T for an individual living at depth d as

$$d_T = \sum_k^{2L_s} \left(\frac{|d_k - \theta_d|}{5 \times 2L_s} \right), \tag{1}$$

where d_k is the sensory trait value of the k th gene. Thus, d_T is a standardized sum of the absolute differences of the trait values from the optimal value. Using d_T , the viability of the individual is assumed to be proportional to

$$W = \exp \left(\frac{-(d_T)^2}{\sigma_s^2} \right), \tag{2}$$

where σ_s^2 quantifies the strength of viability selection for the sensory traits. W is used in the Beverton-Holt equation (Kot 2001; Gavrillets et al. 2007) to incorporate the effect of population density, and the viability, V , is given by

$$V = \frac{1}{1 + (b - 1) \left(\frac{N}{KW} \right)}, \tag{3}$$

where N is the population size in the patch, K is the carrying capacity of the patch and b is the number of offspring per female.

Habitat preference

After viability selection, individuals can emigrate to one of the neighboring patches. In the present model, emigration occurs in three steps, which are defined as movement along the x -axis, y -axis and z -axis within the habitable space. In each step, an individual emigrates

with a probability m (the emigration rate). The emigrating individual may change its patch coordinate by one along the given axis depending on its habitat preference. We assumed that each individual prefers the patch where it has a higher viability [W in Eq. (2)]. Thus, we calculated its preference for either of the two neighboring patches along a given axis and its current patch, and we used that preference to determine the actual emigration probability of the individual as explained next. Emigration along the x - or y -axes does not change viability. Thus, the patch coordinate of the individual increases or decreases by one with a probability of $m/3$, and the patch coordinate does not change when the individual does not move to another patch with a probability of $(1-m)$ or when it moves but then returns to the current patch with a probability of $m/3$. On the other hand, for the z -axis, each individual has a preference because the light environment differs at various depths. We defined the preference for depth as

$$P_H = \exp\left(\frac{-(d_T)^2}{\sigma_H^2}\right), \tag{4}$$

where d_T is defined as in Eq. (1) and σ_H^2 determines the strength of habitat preference. We calculated P_H for the upper neighboring depth (P_{Hu}), the lower neighboring depth (P_{Hl}) and the current depth (P_{Hc}). Their values of each determine the probabilities with which the individual emigrates to a given depth: the patch coordinate of the individual along the z -axis increases by one with a probability of $[mP_{Hu}/(P_{Hu} + P_{Hl} + P_{Hc})]$ and decreases by one with a probability of $[mP_{Hl}/(P_{Hu} + P_{Hl} + P_{Hc})]$; it does not change when the individual does not move (probability of $1-m$) or when it moves but then returns to the current depth with a probability of $[mP_{Hc}/(P_{Hu} + P_{Hl} + P_{Hc})]$.

Mating preference

After migration, females mate with males in the same patch. In the present model, we assumed that each female can mate only once and chooses its mating partner depending on the total difference, D_T , between her sensory trait values, d_k , and the male mating trait values j ,

$$D_T = \sum_k^{2L_o} \left(\frac{|d_k - j|}{5 \times 2L_o}\right) \tag{6}$$

Then, the mating probability is given by

$$P_M = \exp\left(\frac{-(D_T)^2}{\sigma_M^2}\right). \tag{7}$$

The males in the patch are presented to the female in a random order without replacement until she chooses one as a mate [with the probability given by Eq. (7)]. If $P_M < 1$ for all males in the patch, the female may not be able to find a mate.

Next generation

After mating, each fertilized female produces b offspring in the patch. The genotype of each offspring is determined by a random sampling of gametes from its female and male parents. The sex of the offspring is determined randomly.

Initial conditions

The initial population is founded in the center patch of the habitat space ($x = 3$, $y = 3$, $z = 3$) and consists of 10 males and 10 females perfectly adapted to that patch (having allele 3 at all sensory and mating trait loci). All neutral loci are monomorphic with a sequence of nucleotide A.

Parameters

We investigated the effects of viability selection, habitat preference, mating preference and emigration rate by varying the parameters σ_S^2 , σ_H^2 , σ_M^2 and m , respectively. For σ_S^2 , σ_H^2 and σ_M^2 , we examined cases with four values (1, 0.1, 0.025 and 0.01). For m , we examined cases with three values (0.1, 0.05 and 0.01). We also examined the effects of the carrying capacity, K , with three values (500, 250 and 125). All combinations of these parameter values were investigated assuming three different genetic models: (1) 1 locus ($L_S = L_M = 1$), (2) 4 L_M ($L_S = 1$, $L_M = 4$) and (3) 4 L_S ($L_S = 4$, $L_M = 1$). Additionally, we briefly considered cases with more loci. For illustrative purposes, Table 1 shows viability selection, habitat preference, mating preference and mating probabilities when $L_S = 1$, $L_S = 4$ and the sensory trait value deviates from the optimal value by one mutation. For the other parameters, we used the fixed values of $b = 10$, $u_S = u_M = 10^{-5}$ and $u_N = 10^{-6}$. All parameters used in the model are listed in Table 2.

Results

We observed the population state in the simulations after 10,000 generations because the rapid adaptive radiation of the cichlids in Lake Victoria is believed to have occurred within approximately 10,000 years and because the generation time for Lake Victoria cichlids is 0.5–1 year/generation (Terai, personal communication). After 10,000 generations, the number of phenotypes adapted to different habitats seemed to be stable and did not change even after a total of 20,000 generations. We defined adaptation of an individual to the current patch using two criteria: (1) more than 75 % of the $2 \times L_S$ sensory trait values are optimal in the patch and (2) the difference between the male mating trait and the optimal sensory trait values in the patch is less than 0.25 (for example, at $z = 3$, $2.75 \leq$ the mating trait value ≤ 3.25). Note that we can also determine adaptation in the females based on their mating trait genes even though the genes are not expressed. If these criteria are satisfied by more than 50 % of the individuals in the patch, we considered adaptation to have occurred at that depth for that patch. Because we considered five different depths, there are at most five different adapted phenotypic clusters. We used the number of clusters as the number of adaptations in the following models.

Adaptive radiation in the three genetic models

We first examined whether adaptive radiation into more than two phenotypes occurred in the three genetic models. We replicated the simulation five times for each parameter set and calculated the average number of phenotypic clusters.

Figure 1 (for $K = 500$) and Supplementary Fig. 1 (for $K = 250$ and $K = 125$) show the results for the 1-locus model. Evolution of more than two adapted phenotypes occurred but did so under more restricted conditions than evolution of only two adapted phenotypes. We

Table 1 Fitness of the phenotype that deviates from the optimal by one mutation

$\sigma_S^2, \sigma_H^2, \sigma_M^2$	$L_S = 1$	$L_S = 4$
1	0.99	1
0.1	0.90	0.99
0.025	0.67	0.98
0.01	0.37	0.94

Table 2 List of the parameters used in the model

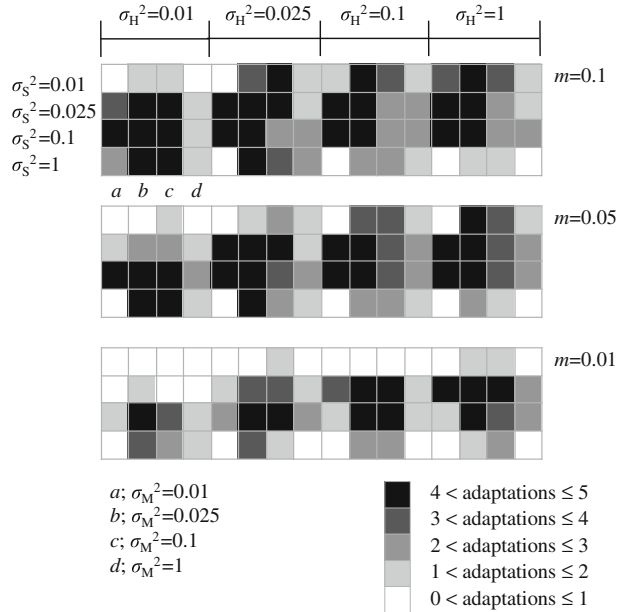
M	Migration rate
b	Number of offspring per female
K	Carrying capacity of one patch
N	Population size in one patch
L_S	Number of loci regulating the sensory trait
L_M	Number of loci regulating the mating trait
u_S	Mutation rate of sensory trait
u_M	Mutation rate of mating trait
u_N	Mutation rate of selectively neutral genes
σ_S	Standard deviation of viability function
σ_H	Standard deviation of habitat preference function
σ_M	Standard deviation of mating preference function

observed that divergent adaptation occurred by successive branching rather than simultaneously. Because the state did not change after 20,000 generations, this result suggests that simple repetitions of lineage splitting from 1 to 2 phenotypes did not always result in the evolution of a large number of phenotypes and that more stringent conditions were required for adaptive radiation to occur, in particular, with high emigration rates and intermediate strengths of viability selection and mating preference. On the other hand, the effect of habitat preferences seems weak. The carrying capacity also had a strong effect, with an increase in K leading to an increase in the number of adapted phenotypes. However, K did not qualitatively alter the effects of the other parameters (Supplementary Fig. 1).

Figure 2 ($K = 500$) and Supplementary Fig. 2 ($K = 250, 125$) show the results for the 4- L_M model. Although the number of adapted phenotypes was smaller than in the 1-locus model, this difference was very small and the effects of the three factors related to fitness, the emigration rate and the carrying capacity were similar to those in the 1-locus model. These results suggest that small differences in the number of mating trait loci do not affect adaptation to new habitats.

Figure 3 ($K = 500$) and Supplementary Fig. 3 ($K = 250, 125$) show the results for the 4- L_S model. In contrast to the mating-trait model, an increase in the number of the sensory trait loci greatly decreases the number of adaptive phenotypes. Intermediate strengths of viability selection and mating preference were more important for allowing adaptation to occur compared to the cases shown in Figs. 1 and 2. These results suggest strong effects of the number of sensory trait loci on the probability of adaptation. The sensory traits might have a stronger effect than the mating trait because they affect both natural and sexual selection. As for the two genetic models mentioned above, decreasing the carrying capacity made adaptation/divergence more difficult, and when $K = 125$, adaptation never occurred (Supplementary Fig. 3b).

Fig. 1 The average number of adapted phenotypes for various levels of viability selection (σ_S^2), habitat preference (σ_H^2), sexual selection (σ_M^2) and emigration rates in the 1-locus model. The case for $K = 500$ is shown. We ran the simulation five times for each parameter set and calculated the average number of adaptations from these five replicates. The average number of adapted phenotypes is shown in *gray scale*



Finally, the effect of isolation by distance seemed weak because, in almost all cases, adaptive phenotypes spread to all patches at the same depth, and there was no phenotypic differentiation along the horizontal axis (data not shown).

F_{ST} under different selection pressures and emigration rates in the 1-locus model

We calculated F_{ST} values (Hudson et al. 1992) at the neutral loci to determine the extent of gene flow between divergent phenotypes. First, we calculated F_{ST} assuming the 1-locus model in 276 out of the 576 parameter sets in which the average number of adapted phenotypes exceeded 1. We used a generalized linear model (Nelder and Robert 1972) with the strength of viability selection (σ_S^2), habitat preference (σ_H^2), mating preference (σ_M^2) and emigration rate (m) as the explanatory variables, and we assumed that the response variable, F_{ST} , followed a Gamma distribution based on the Akaike Information Criterion (Akaike 1974). Table 3 shows the estimates of the coefficients of the explanatory variables in the linear model. Among the four factors, habitat preference (σ_H^2), mating preference (σ_M^2) and the emigration rate (m) had significant negative correlations with F_{ST} , which shows that strong habitat and mating preferences and a low emigration rate effectively prevented gene flow between phenotypes. In contrast, however, the strength of viability selection (σ_S^2) did not have a significant effect on F_{ST} . This observation was somewhat counterintuitive and will be discussed later.

Differences in F_{ST} values between the three genetic models

Next, we examined how differences between the three genetic models affected F_{ST} . We chose 30 parameter sets for which adaptation occurred in all three models and compared the F_{ST} values between them (Fig. 4). Because the parameter range in which adaptation occurred in the 4- L_S model was narrower than and mostly included in the other two models

Fig. 2 The average number of adapted phenotypes for various levels of viability selection (σ_S^2), habitat preference (σ_H^2), sexual selection (σ_M^2) and emigration rates in the 4- L_M model. The case for $K = 500$ is shown. As in Fig. 1, we calculated the average number of adaptations from five replicates and represented the average in *gray scale*

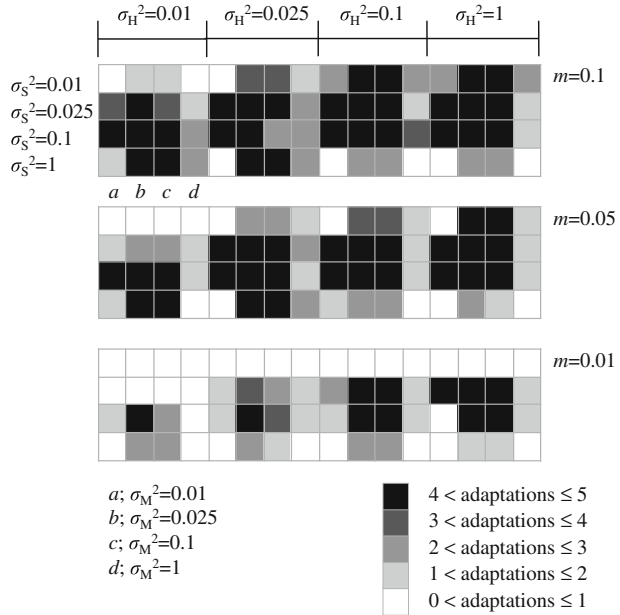
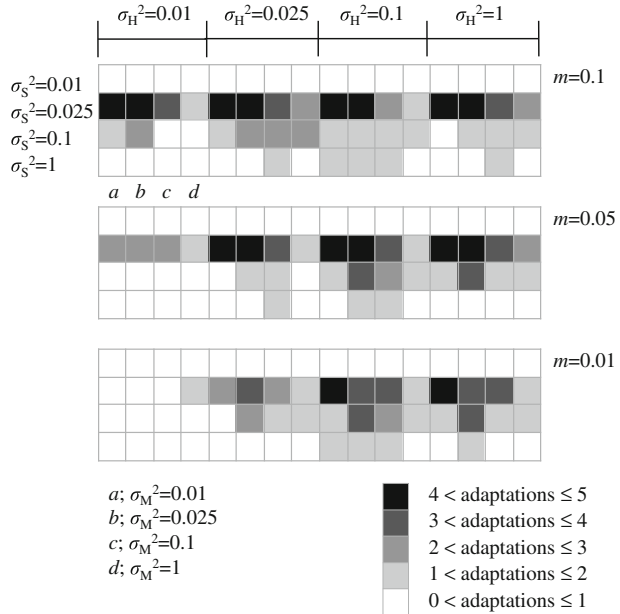


Fig. 3 The average number of adapted phenotypes for various levels of viability selection (σ_S^2), habitat preference (σ_H^2), sexual selection (σ_M^2) and emigration rates in the 4- L_S model. The case for $K = 500$ is shown. As in Fig. 1, we calculated the average number of adaptations from five replicates and represented the average in *gray scale*



(Figs. 1, 2, 3), the 30 parameter sets correspond to the parameter sets of divergent adaptation in the 4- L_S model. As shown in Fig. 4, the effects of the number of loci seemed to be small. In fact, the sign of the correlation between F_{ST} and σ_S^2 , σ_H^2 , σ_M^2 or m was the same in all three genetic models (data not shown). However, the F_{ST} values themselves were

significantly different between the models in some parameter sets. While F_{ST} differed significantly between the 1-locus and 4- L_M models in only 3 parameter sets, it differed significantly between the 1-locus and 4- L_S models and between the 4- L_M and 4- L_S models in about half of the 30 parameter sets. These results suggest that the number of sensory trait loci affects F_{ST} more than the number of mating trait loci.

More loci

We also investigated cases with more loci and describe the results briefly here. We considered two models—(1) 8 L_M ($L_S = 1$, $L_M = 8$) and (2) 8 L_S ($L_S = 8$, $L_M = 1$)—and calculated the average number of adapted phenotypes as described in the previous models (Fig. 5). In accordance with our previous results, an increase in the number of loci controlling the mating trait did not affect the results very much. In contrast, increasing the number of loci controlling the sensory trait greatly decreased the number of adapted phenotypes. We also found that the number of loci controlling the mating trait did not greatly affect F_{ST} as in Fig. 4 (data not shown).

Phylogenetic relationships among phenotypes

Finally, we examined the phylogenetic relationships among individuals of the adapted phenotypes. We chose four parameter sets that showed different F_{ST} values and reconstructed Neighbor-Joining trees using MEGA version 4 (Tamura et al. 2007) for the sequences of the 50 neutral loci (Supplementary Fig. 4). Individuals with each phenotype were never monophyletic, which suggests continued gene flow. In particular, when F_{ST} was low, individuals of all phenotypes seemed completely intermingled.

Discussion

Previously Kocher (2004) proposed that the speciation process of East African cichlids can be divided into three phases: (1) adaptation to rocky or sandy habitat, (2) adaptation to different feeding and (3) diversification of color patterns. Sensory drive accelerates the third phase through which populations in similar habitats can speciate even under gene flow; therefore, this mechanism would make a large contribution to the adaptive radiation of East African cichlids (Terai et al. 2006; Seehausen et al. 2008).

In this study, we developed a model of sensory drive, which allowed us to study the subsequent divergent adaptation of a mating trait and the establishment of assortative mating. Such sensory-driven speciation has been reported in fishes living in different light environments (Endler and Houde 1995; Boughman 2001; 2002; Seehausen et al. 2008). In a lake, the light environment varies gradually by depth (Seehausen et al. 2008), and so conditions such as those simulated in our model are likely to be found in nature. Using this model, we investigated two questions that have not been considered in previous studies: (1) the possibility of adaptive radiation (i.e., the emergence of more than two adapted phenotypic clusters) caused by sensory drive in the presence of gene flow and (2) the effects of viability selection, habitat preference, mating preference and emigration rate on the patterns of variation at selectively neutral loci.

We found that adaptive radiation was possible for specific values of the parameters representing the number of loci regulating the sensory trait, the selection pressures, the habitat preference, the emigration rate and the carrying capacity (Figs. 1, 2, 3 and

Table 3 The results of the generalized linear model analysis between the F_{ST} values and the strengths of viability selection, habitat and mating preference and emigration rate

	Coefficient	Significance
σ_S^2	0.0184 ± 0.0126	n.s.
σ_H^2	-0.0221 ± 0.0090	*
σ_M^2	-0.0515 ± 0.0091	***
m	-0.6497 ± 0.1420	***

* $0.01 \leq p \leq 0.05$ *** $0 \leq p \leq 0.001$

Supplementary Figs 1, 2, 3). Because we did not observe differentiation along the horizontal axis, the radiation along the vertical axis under these parameter sets seemed to be caused by adaptive evolution and not by isolation by distance. The number of adapted phenotypes was five in these parameter sets, which was the maximum set by the number of ecological niches in the simulation. In the other sets, however, the number was less than five. Some of the possible reasons why the maximum was not achieved, including the increase of the migration load (Butlin et al. 2009), will be discussed later. Below, we first discuss the effects of the parameters in more detail.

Importance of a small number of loci regulating the sensory trait(s)

Figures 1, 2 and 3 suggest the importance of a small number of loci regulating the sensory trait. Generally, the probability of ecological speciation increases as the number of loci regulating ecological traits decreases because selection on individual loci becomes stronger (Gavrilets and Vose 2005). In sensory drive, the evolution of a sensory trait occurs first, and once a sensory trait has diverged, a mating trait is under strong sexual selection and will diverge itself under almost any circumstance. This is most likely why the number of loci regulating the sensory trait(s) has a larger impact on ecological speciation than does the number of loci regulating the mating trait(s) (Figs. 2, 3). Importantly, one opsin locus, the long-wavelength-sensitive opsin gene (*LWS*), is thought to play an important role in speciation of the cichlids in Lake Victoria (Terai et al. 2002, 2006; Seehausen et al. 2008). In addition, Miyagi et al. (2012) found remarkably higher between-species divergence at the *LWS* gene than at the other cone opsin genes (e.g., the blue-sensitive *SWS* opsin gene and green-sensitive *RH2* opsin gene). This is in accordance with our model, which suggests that divergence occurs most easily when it is based on a single sensory locus ($L_S = 1$).

Importance of intermediate levels of selection

Figures 1, 2 and 3 also suggest the importance of intermediate levels of viability selection and mating preference. The importance of intermediate selection pressures for ecological speciation has been repeatedly stressed in previous theoretical studies (Doebeli and Dieckmann 2003; Gavrilets and Vose 2007; Gavrilets et al. 2007; Kawata et al. 2007), and our results are consistent with those results. If natural or sexual selection is too weak, gene flow between different adapted phenotypes occurs and causes their admixture or extinction of one of them by genetic drift. In addition, in our model, the migration load becomes strong when sexual selection is weak because emigrants generate hybrids that have low viability and reduce of the population size. As shown in Figs 1, 2 and 3, the average number of

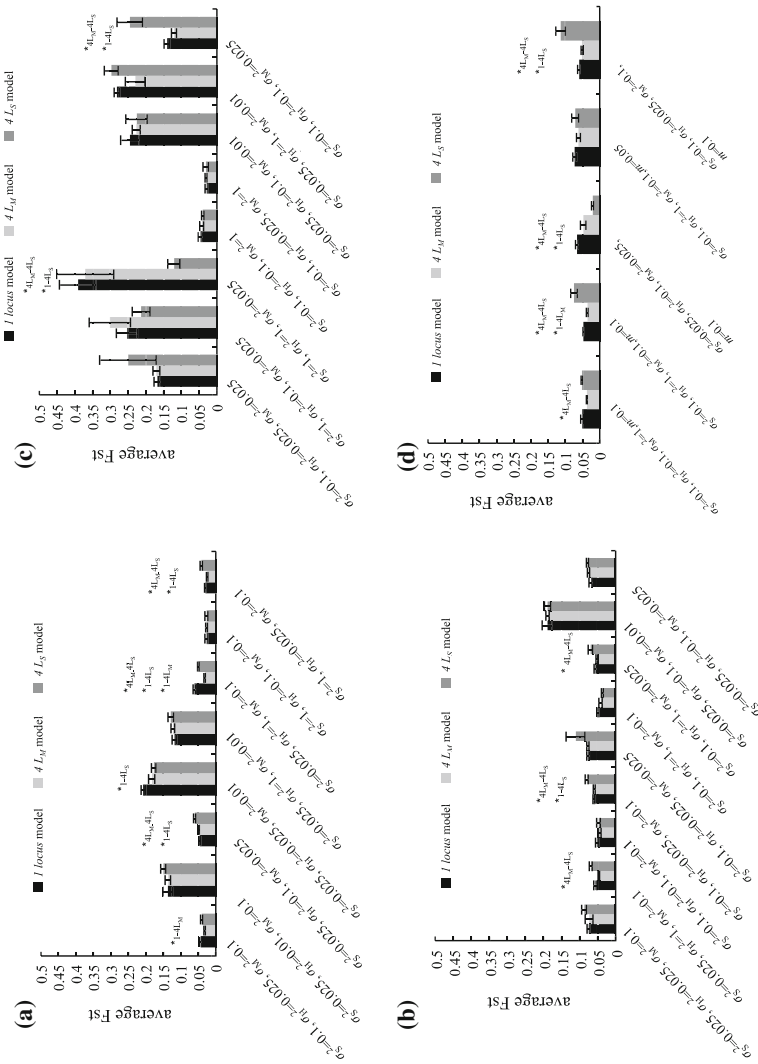
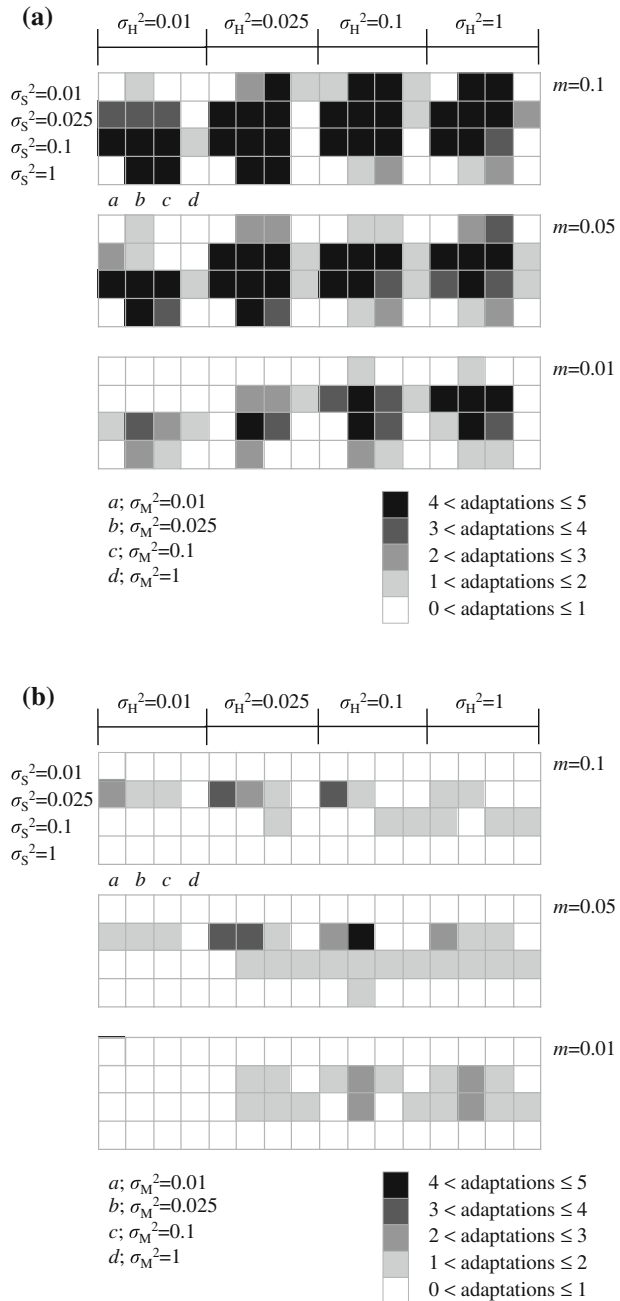


Fig. 4 The average and standard error for the F_{ST} values for the three genetic models. We show the results for 30 parameter sets where more than one phenotype appeared in all three of the genetic models; **a** $K = 500, m = 0.1$; **b** $K = 500, m = 0.05$; **c** $K = 500, m = 0.01$ and **d** $K = 250, m = 0.1$ and 0.05 . * $X_{i,j}$ indicates that the F_{ST} value was significantly different ($p \leq 0.05$) between the genetic models X and Y

Fig. 5 The average number of adapted phenotypes for various levels of viability selection (σ_S^2), habitat preference (σ_H^2), sexual selection (σ_M^2) and emigration rates in (a) the 8- L_M model and (b) the 8- L_S model. The cases for $K = 500$ are shown. As in Fig. 1, we calculated the average number of adaptations from five replicates and represented the average in *gray scale*



phenotypes was always less than 3 when $\sigma_M^2 = 1$, which suggests that the migration load makes it more difficult for evolution of the maximum number of phenotypes to occur. On the other hand, if natural or sexual selection is too strong, an invasion of new habitats or mating of individuals with new sensory or mating traits with the original phenotype

becomes more difficult. Therefore, intermediate levels of selection pressure seem important for ecological speciation. Figures 1, 2 and 3 suggest that habitat preference has a smaller effect than do viability selection and mating preference and that adaptive radiation is possible in a wide range of habitat preference. Therefore, although Gavrilets (2004, pp 350–352) stated that habitat preference/selection is important for speciation to occur with gene flow, our results suggest that sensory drive can cause adaptive radiation even if habitat preference is very weak with almost no preference between two neighboring patches (when $\sigma_H^2 = 1$, preference of an individual that has an optimal sensory trait in the current patch with that in the neighbor patch is 0.99 in $L_S = 1$, see Table 1).

Importance of a high emigration rate and large carrying capacity

We also found that the number of adapted phenotypes is positively correlated with the emigration rate and the carrying capacity (Figs. 1, 2, 3 and Supplementary Figs. 1, 2, 3). Both parameters increase the number of emigrants, some of whom may carry adaptive mutations into the new habitat, and thus increase the opportunities for adaptation to occur. Of course, higher emigration rates also increase gene flow between the differentially adapted phenotypes living at different depths and are expected to increase the migration load, which would lead to negative effects on adaptation and differentiation. However, in sensory drive, assortative mating is established rapidly and reproductive isolation will develop before substantial mixing of the differentially adapted phenotypes occurs unless sexual selection is too weak.

Here, because we started the simulations from a monomorphic population completely adapted to a single patch, divergence strongly depended on quick invasion of and adaptation to new habitats, which is facilitated by high emigration rates. Additionally, a larger carrying capacity may decrease the possibility that the invaders will become extinct. Because we did not examine other initial conditions, we do not know whether the effects of the emigration rate or the carrying capacity change when, for example, the initial population is widely distributed over the habitat space or when initial variation is present in the sensory traits or the mating trait.

Neutral variation and differentiation between phenotypes

We found that habitat preference, mating preference and the emigration rate significantly affected F_{ST} between phenotypes. Therefore, we may be able to use F_{ST} to infer the parameter ranges for the habitat preference, mating preference and emigration rates in natural populations, but an accurate estimation of each parameter value would not be possible from the F_{ST} values alone. Another important result is that F_{ST} values are robust to differences in the number of the mating trait loci. Therefore, inferring the parameters for the past population history (e.g., habitat preference, mating preference and the emigration rate) may be possible without knowing the detailed genetic architecture of the mating trait. In contrast, an increase in the number of sensory trait loci significantly changed the F_{ST} values over a large range of parameters, and so identifying an accurate genetic model for the sensory trait is crucial for inferring the population history. As mentioned above, one opsin locus, *LWS*, is thought to play an important role in the sensory-drive speciation of the cichlids in Lake Victoria. Therefore, we can make some tentative inferences about this case using the results from the 1-locus model.

Previously, Samonte et al. (2007) concluded that populations of the same species genetically intermingle with those of the other species in the African cichlids of Lake

Victoria based on an analysis of genetic variation at 11 loci. Moreover, Wagner et al. (2013) have recently shown that individuals of each species become monophyletic only when more than thousands of marker loci are used to infer their phylogeny based on a Genome-wide RAD analysis. The phylogenetic trees reconstructed in our study also showed a lack of monophyletic relationships among individuals with the same phenotype (Supplementary Fig. 4). Thus, the adaptive radiation of cichlids in Lake Victoria, which shows polyphyletic relationships among individuals belonging to a single species when a small number of loci are used for the inference, can be explained by our sensory drive model using parameter values such as those used in Supplementary Fig. 4. For example, Seehausen et al. (2008) estimated the F_{ST} between two cichlid species, *Pundamilia pundamilia* and *P. nyererei*, to be 0.01–0.025 using data from 11 microsatellite loci. In our model, such low F_{ST} values were found in only 11 parameter sets, 9 of which satisfied $m = 0.1$ and all of which satisfied $\sigma_M^2 \geq 0.1$. In addition, we may be able to use the number of adapted phenotypes for the parameter inference. In the case of *Pundamilia*, there are two types of *LWS* in the habitat (Seehausen et al. 2008). Of the above parameter sets, 9 of 11 showed two adapted phenotypes, 8 of which satisfied $m = 1$ and 7 of which satisfied $\sigma_M^2 = 1$. This might suggest that *P. pundamilia* and *P. nyererei* had a high emigration rate and weak mating preference. Under such conditions, these two species should generate hybrids, and indeed, individuals with intermediate body color have been observed in natural populations (Seehausen et al. 2008). Terai et al. (2006) also estimated F_{ST} between three populations of *Neochromis greenwoodi*, two populations (including its offshore incipient species *N. omnicaeruleus*) from one habitat and one population from another habitat. Using markers unlinked to *LWS*, the average F_{ST} between populations with different habitats was estimated to be 0.032 and 0.043. Although these two habitats are separated by long distance and we may not be able to apply our results directly to infer the parameter values, the number of parameter sets under which F_{ST} values were within 0.032–0.043 and only two phenotypes evolved was 38, 35 of which satisfied $\sigma_M^2 \geq 0.1$ and 29 of which satisfied $\sigma_H^2 \geq 0.1$. This might suggest that *N. omnicaeruleus* and *N. greenwoodi* had a weak mating preference and habitat preference.

In this study, we considered only selectively neutral genes unlinked to the selected loci to infer the population structure. However, F_{ST} at selectively neutral genes may increase as linkage to the selected loci becomes stronger (Feder and Nosil 2010; Feder et al. 2012). Therefore, if we find an F_{ST} value at a locus much larger than those at the other loci, the observation indicates that the locus is closely linked to or itself a selected locus.

No correlation between F_{ST} and strength of viability selection

In general, strong selection should result in strong reproductive isolation and is thus expected to increase F_{ST} between divergently adapted phenotypes. However, in our study, F_{ST} values did not increase as viability selection became stronger (Table 3). This means that although sexual selection affects the divergence at neutral loci, viability selection on the sensory trait does not. In sensory drive, evolution in the sensory trait is soon followed by evolution in the mating trait causing assortative mating. Then, the strength of viability selection does not matter much for the accumulation of the genetic differentiation at the neutral loci. In addition, the density-dependent viability selection (see Eq. 3) allows individuals with non-adaptive mutations to survive when the population density is low even if the viability selection is strong. These surviving individuals may migrate to other favorable patches. For these reasons, the effect of the level of viability selection on population differentiation would be smaller than those of the other factors.

Application to other organisms

As mentioned earlier, our result in a three-dimensional habitat would be applicable to fishes that speciated by sensory drive. Previous studies suggested that male coloration and female light sensitivity have caused speciation by sensory drive in guppies (Endler and Houde 1995; Hoffman et al. 2007) and sticklebacks (Boughman 2001). Therefore, we may be able to apply our model and results directly for these species. For example, if we know the number of loci encoding the opsins that caused the speciation, we may be able to infer the range of the model parameters such as mating preference and the emigration rate. Moreover, in East African cichlids, mating sound and olfaction are also known to affect the mate choice (Amorim et al. 2004; Plenderleith et al. 2005). As reviewed in Salzburger (2009), there is some evidence showing that these traits are ecologically adaptive. Therefore, if we know the genetic architecture of these traits, our theoretical result may help reveal the role of sensory drive in sound or olfaction in East African cichlids. Finally, a recent study suggested that Amazonian bird species have speciated by acoustic adaptation (Tobias et al. 2010). Because the sound transmission property is different in different habitats (bamboo forest and terra firme forest in this case), adaptation of the male song might contribute to the establishment of reproductive isolation. Again, this case might be analyzed using our theoretical results, but some modification such as an exchange of the horizontal and vertical axes and an adjustment of the number of environments may be necessary.

Finally, we briefly discuss the validity and effect of changing some of the assumptions we made in our model.

First, we assumed that emigration occurs after viability selection. If emigration occurs before viability selection, individuals with a new mutation in the sensory trait can move to patches in which they have higher fitness. This would increase the viability of new phenotypes produced by mutations and thus make adaptive evolution more likely to occur.

Second, in our model, a female might not mate if the probability of mating with this female for any male in the patch is less than one. Thus, if the mating preference is strong, a female with a new phenotype for the sensory trait may not be able to produce offspring. If we assume that all females can mate as assumed by Kawata et al. (2007) and Gavrilets et al. (2007), then adaptive evolution would be more likely to occur because a female having a new phenotype for the sensory trait can mate with the probability of one and produce offspring.

Third, we assumed that each individual had its habitat preference determined by its sensory trait and environment. Although the cichlids in Lake Victoria usually inhabit the optimal light environment, it is not clear whether this is because of habitat preference or the result of selection. However, as mentioned in *Importance of intermediate levels of selection*, adaptive radiation was possible even if habitat preference was very weak. Thus, we conclude that sensory drive can cause adaptive radiation without a strict qualification about habitat preference.

Fourth, we assumed that each gene independently expresses an integer-valued phenotype in the sensory trait [see eq. (1)]. We might have modeled the sensory trait in such a way that selection operates on the sum of the effects of the $2 \times L_S$ sensory trait alleles as in the mating trait rather than independently on the effect of each allele. However, it is known that African cichlids have multiple opsin genes and that each of them expresses a specific cone opsin with a different absorption wavelength (Carleton and Kocher 2001). Animals are also known to have many olfactory receptor genes and can distinguish many types of odor (Buck and Axel 1991). A recent study has shown that African cichlids also have and

express many different olfactory receptor genes within one individual (Ota et al. 2012). Thus, our assumption that each allele independently expresses a different sensory trait would be more realistic than the model that assumes the additive determination of one sensory trait.

Fifth, we assumed free recombination between loci. Tight linkage between loci would accelerate the accumulation of good combinations of alleles and would make adaptation more likely to occur.

In summary, some of our assumptions are based on the nature of cichlids in Lake Victoria. On the other hand, the other assumptions were made because of lack of knowledge. However, all these assumptions make adaptive evolution more difficult. Thus, if one or more of them are violated, adaptive radiation would occur over an even wider parameter range.

Conclusion

In this study, we found that sensory drive leads to adaptive radiation even when gene flow occurs between the adapted phenotypes. This shows that sensory drive is one of the important factors that caused the adaptive radiation observed in the cichlids of Lake Victoria. We also found that the F_{ST} values between rapidly diversifying phenotypes at neutral loci varied considerably depending on habitat preferences, mating preferences and emigration rates. Interestingly, the F_{ST} values were almost the same regardless of small differences in the number of loci controlling the mating trait. These results suggest that F_{ST} values can be used to estimate the past population history without a fully accurate genetic model for the mating trait.

Acknowledgments This study was partially supported by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (nos. 21370013, and 20248017) and by the Environment Research and Technology Development Fund (S-9-2) of the Ministry of the Environment, Japan.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19:716–723
- Amorim MCP, Knight ME, Stratoudakis Y et al (2004) Differences in sounds made by courting males of three closely related Lake Malawi cichlid species. *J Fish Biol* 65:1358–1371
- Beuning KRM, Kelts K, Ito E (1997) Palaeohydrology of Lake Victoria, East Africa, inferred from 18O/16O ratios in sediment cellulose. *Geology* 25:1083–1086
- Boughman JW (2001) Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* 411:944–947
- Boughman JW (2002) How sensory drive can promote speciation. *TREE* 17:571–577
- Buck L, Axel R (1991) A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* 65:175–183
- Butlin R, Bridle J, Schluter D (2009) *Speciation and patterns of diversity*. Cambridge University Press, Cambridge
- Carleton KL, Kocher TD (2001) Cone opsin genes of African cichlid fishes: tuning spectral sensitivity by differential gene expression. *Mol Biol Evol* 18:1540–1550

- Coyne J, Orr HA (2004) Speciation. Sinauer Associates Inc., Sunderland
- Dieckmann U, Doebeli M (1999) On the origin of species by sympatric speciation. *Nature* 400:354–357
- Doebeli M, Dieckmann U (2003) Speciation along environmental gradients. *Nature* 421:259–264
- Doebeli M, Dieckmann U (2005) Adaptive dynamics as a mathematical tool for studying the ecology of speciation process. *J Evol Biol* 18:1194–1200
- Endler JA, Houde AE (1995) Geographic-variation in female preferences for male traits in *Poecilia-reticulata*. *Evolution* 49:456–468
- Feder JL, Nosil P (2010) The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution* 64:1729–1747
- Feder JL, Egan SP, Nosil P (2012) The genomics of speciation-with-gene-flow. *Trends Genet* 28:342–350
- Gavrilets S (2004) Fitness landscapes and the origin of species. Princeton University Press, Princeton
- Gavrilets S, Losos JB (2009) Adaptive radiation: contrasting theory with data. *Science* 323:732–737
- Gavrilets S, Vose A (2005) Dynamic patterns of adaptive radiation. *PNAS* 102:18040–18045
- Gavrilets S, Vose A (2007) Case studies and mathematical models of ecological speciation. 2. Palms on an oceanic island. *Mol Ecol* 16:2910–2921
- Gavrilets S, Vose A, Barluenga M et al (2007) Case studies and mathematical model of ecological speciation. 1. Cichlids in a crater lake. *Mol Ecol* 16:2893–2909
- Hoffman M, Tripathi N, Henz SR (2007) Opsin gene duplication and diversification in the guppy, a model for sexual selection. *Proc R Soc B* 274:33–42
- Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics* 132:583–589
- Johnson TC, Scholz CA, Talbot MR et al (1996) Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science* 273:1091–1093
- Johnson TC, Kelts K, Odada E (2000) The Holocene history of Lake Victoria. *Ambio* 29:2–11
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. Academic Press, New York
- Kawata M, Shoji A, Kawamura S et al (2007) A genetically explicit model of speciation by sensory drive within a continuous population in aquatic environments. *BMC Evol Biol* 7:99
- Kocher TD (2004) Adaptive evolution and explosive speciation: the Cichlid fish model. *Nat Rev Genet* 5:288–298
- Kot M (2001) Elements of mathematical ecology. Cambridge University Press, Cambridge
- Meyer A (1993) Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *TREE* 8:279–284
- Meyer A, Kocher TD, Basasibwaki P et al (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequence. *Nature* 347:550–553
- Miyagi R, Terai Y, Aibara M et al (2012) Correlation between nuptial colors and visual sensitivities tuned by opsins leads to species richness in sympatric Lake Victoria cichlid fishes. *Mol Biol Evol* 29:3281–3296
- Nelder J, Robert W (1972) Generalized linear models. *J R Stat Soc* 135:370–384
- Ohta T, Kimura M (1973) Model of mutation appropriate to estimate number of electrophoretically detectable alleles in a finite population. *Genet Res* 22:201–204
- Ota T, Nikaido M, Suzuki H et al (2012) Characterization of V1R receptor (ora) genes in Lake Victoria cichlids. *Gene* 499:273–279
- Plenderleith M, van Oosterhout C, Robinson RL et al (2005) Female preference for conspecific males based on olfactory cues in a Lake Malawi cichlid fish. *Biol Lett* 1:411–414
- Rundle HD, Nosil P (2005) Ecological speciation. *Ecol Lett* 8:336–352
- Salzburger W (2009) The interaction of sexually and naturally selected traits in the adaptive radiations of cichlid fishes. *Mol Ecol* 18:169–185
- Samonte IE, Satta Y, Sato A et al (2007) Gene flow between species of Lake Victoria haplochromine fishes. *Mol Biol Evol* 23:2069–2080
- Schluter D (2009) Evidence for ecological speciation and its alternative. *Science* 323:737–741
- Scholz CA, Johnson TC, Cattaneo P et al (1998) Initial results of 1995 IDEAL seismic reflection survey of Lake Victoria, Uganda and Tanzania. In environmental change and response in East African lakes. Kluwer, Dordrecht
- Seehausen O (1996) Lake Victoria rock cichlids. *Verduijn Cichlids, Zevenhuizen*
- Seehausen O (2002) Patterns in fish radiation are compatible with Pleistocene desiccation of Lake Victoria and 14600 year history for its cichlid species flock. *Proc R Soc B* 269:491–497
- Seehausen O, Bouton N (1997) Microdistribution and fluctuations in niche overlap in a rocky shore cichlid community in Lake Victoria. *Ecol Freshw Fish* 6:161–173
- Seehausen O, Terai Y, Magalhaes IS et al (2008) Speciation through sensory drive in cichlid fish. *Nature* 455:620–626

- Tamura K, Dudley J, Nei M et al (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- Terai Y, Mayer WE, Klein J et al (2002) The effect of selection on a long wavelength-sensitive (LWS) opsin gene of Lake Victoria cichlid fishes. *PNAS* 99:15501–15506
- Terai Y, Seehausen O, Sasaki T et al (2006) Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biol* 4:2244–2251
- Tobias JA, Aben J, Brumfield RT et al (2010) Song divergence by sensory drive in Amazonian birds. *Evolution* 64:2820–2839
- Turner GF, Seehausen O, Knight ME et al (2001) How many species of cichlid fishes are there in African Lakes. *Mol Ecol* 10:793–806
- Verheyen E, Salzburger W, Snoeks J et al (2003) Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. *Science* 300:325–329
- Wagner CE, Keller I, Witter S et al (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in Lake Victoria cichlid adaptive radiation. *Mol Ecol* 22:787–793