

# Spatio-temporal modes of speciation

Martin Rost and Michael Lässig

Institut für Theoretische Physik, Universität zu Köln,  
50937 Köln, Germany

**Abstract.** The split of a population into two reproductively isolated subpopulations is studied within a model including spatial heterogeneity. We find three dynamical pathways of speciation resulting from a coupling of space, competition and mating behaviour: (i) sympatric at small habitat heterogeneity, (ii) sympatric with subsequent spatial differentiation at intermediate heterogeneity, and (iii) allopatric under strong heterogeneity.

## 1 Introduction

Speciation is the splitting of a species into two new ones. A population of individuals forms a species if they produce viable offspring among them. Essential steps at the beginning of speciation are differentiation in body characteristics and the formation of two reproductively isolated subpopulations. The dynamics of this division, which is the subject of this paper, can be rather fast. Eventually, if interbreeding of the two subpopulations is cut for long enough, the ability to mate successfully will be lost and the division gets irreversible. Speciation has been a long standing issue in evolutionary biology [1,2,3,4,5,6,7,8]. As we will show in this article, it is also a challenging problem of non-equilibrium statistical physics.

An obvious way of species division is *allopatric speciation* [2]. In this case a population gets separated by an external cause, e.g., a previously connected part of the sea may be divided after the formation of a natural barrier, and inhabitant organisms will be isolated from their relatives on the other side. Different characteristic forms may evolve independently and eventually the two groups will form two different species which cannot produce viable common offspring even if the barrier gets removed again.

In *sympatric speciation*, on the other hand, the two dividing groups continuously share a common habitat. Also in sympatry differentiation into subpopulations with different characteristics can be favoured, a process called *disruptive selection*. [he subpopulations fill different *ecological niches*, e.g., they feed from different resources. Niche populations can be stable if their gain in fitness through reduced competition outweighs the loss in fitness through specialisation [4,6,8].

An obstacle to sympatric speciation lies in sexual reproduction [2]. By the rules of genetics, mating between the two opposite subpopulations always produces offspring of intermediate phenotype and prevents their drifting apart. For a long time therefore speciation has been believed to take place predominantly

in allopatry [2]. If speciation is to happen in sympatry, i.e., with both subpopulations sharing a common habitat, some mechanism has to prevent interbreeding.

Theoretical studies, e.g., [3,4], have addressed the possibility of and conditions for sympatric speciation, and tried to elaborate possible mechanisms to prevent interbreeding. One of the proposed mechanisms is assortative mating, where e.g. a female's mate choice is determined by a male's trait, how close or different it is from her own or her preferred one. Mating preference may also depend on a different male trait and both may mutually enhance their evolution towards extreme characteristics [5].

In more recent years, reconstruction of phylogenetic trees from molecular data obtained in new field studies have dramatically changed our understanding of the processes and brought up striking empirical evidence for sympatric speciation. The most prominent example is cichlid fishes in the great African lakes (Victoria, Malawi), where several hundred sister species of monophyletic origin, i.e., all descending from a single colonising species, have been found to coexist [9]. Even more surprisingly this evolution must have taken place in a surprisingly short time: Geological data from the bottom of Lake Victoria indicate that it was completely dry about 12 000 years ago [10]. All species inhabiting the lake today must have colonised it *after* that desiccation and subsequently have speciated. It seems unlikely that with different low water levels the lake would have been divided in many small lakes to give rise to the opportunity of allopatric speciation for so many species in such a short time.

Moreover, these observations also show spatial structure in the fish populations despite the absence of barriers and their sympatric origin. They prefer different parts of the lake due to different food supply and use. Some feed from organisms on the ground, others from floating particles throughout the lake [9]. Encounters between such different populations may well be reduced by this spatial separation. This has an important consequence on the picture of sympatric speciation: As the subpopulations start to differ in phenotypic characteristic they gain the possibility of a secondary spatial differentiation, also called patching.

Spatial separation plays an even larger role in organisms which are not constraint to a relatively small and spatially well mixed habitat such as a lake. It is more pronounced when, e.g., the splitting populations specialise as parasites on different host plants, or when the interaction range of individuals is small compared to the area inhabited by the entire species. This is e.g. the case for *salamandra* taxa in Europe who seem to consist of different clades whose colonisation steps after the last ice age can be tracked down [11].

Speciation is more than a single event, and only the *history* of the entire process yields an adequate understanding of spatially structured populations. Different evolutionary pathways describe primary spatial (and hence allopatric) division compared to secondary patching with originally sympatric differentiation.

Present models for sympatric speciation neglect space [6,7,8], in terms of statistical physics they are "mean field" models. In order to fully understand the history of a speciation process one needs more refined categories than merely

the alternative of sym- and allopatry. In this article we combine the “internal” trait space of the populations with the “external” topographical space, i.e., to go from a mean field model to a spatio-temporal field theory. The resulting modes of speciation show indeed an intricate interplay of internal and external space.

This article is organised as follows: In the next Section we present the general setup of the spatio-temporal model. Characteristic *states* of the system and their biological interpretations are presented in Section 3. Section 4 then shows the possible *evolutionary pathways* and the transitions between states occurring along these pathways. A brief discussion ends the article.

## 2 Evolution of competing phenotypes

**General form.** The model in its most general form deals with *competitive* interaction of different groups of a population extending over a whole spectrum of phenotypes and living in a structured landscape. The dynamical variable is a density  $N(x; \mathbf{r}, t)$ , depending on time  $t$  and the (“external”) spatial coordinate  $\mathbf{r}$ , as well as some suitable variable  $x$  denoting a character in the (“internal”) trait space. In the example of fish this is typically the body size [9], but one can imagine any other variable in which the split into subpopulations and finally speciation becomes first visible. A “minimal” model equation for speciation can be written in the form

$$\partial_t N(x; \mathbf{r}, t) = \nu \nabla_{\mathbf{r}}^2 N(x; \mathbf{r}, t) - \nabla_{\mathbf{r}} \cdot (\lambda \mathbf{v}(x; \mathbf{r}, t) N(x; \mathbf{r}, t)) + \quad (1)$$

$$f(x; \mathbf{r}) R(x; \mathbf{r}, t) - \left[ \alpha(\mathbf{r}) + \int dy \beta(x-y) N(y; \mathbf{r}, t) \right] N(x; \mathbf{r}, t).$$

It has the structure of a nonlinear reaction-diffusion equation and contains the following features:

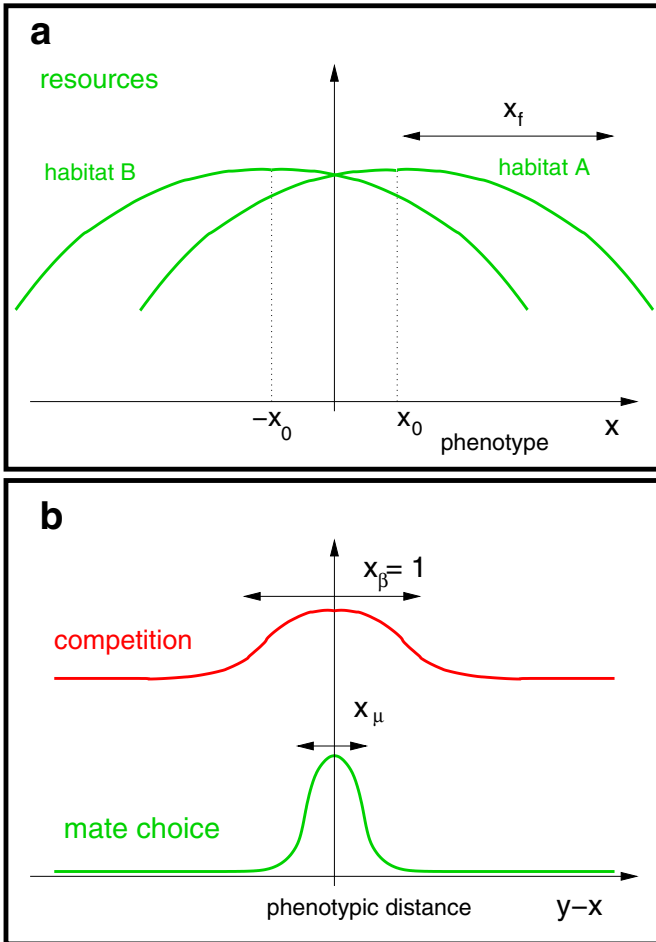
- *Movement in space* is due to diffusion (first term on the right hand side) and to deterministic drift (second term), here expressed by a “velocity” field  $\mathbf{v}$  which could, e.g., be due to a gradient in habitat quality.
- $R$  is the *birth rate* of offspring. In a purely asexual model it would be proportional to  $N$  itself but here it reflects the genetic inheritance patterns, as will be specified below.  $R$  is multiplied by  $f(x; \mathbf{r})$ , which denotes the *habitat quality* at point  $\mathbf{r}$  and plays the role of a fitness.
- The *competition* is quadratic, a simple general form, familiar from Lotka-Volterra equations, proportional to the number of encounters per time unit between two individuals. Some function  $\beta(x-y)$  denotes the phenotype dependence of its strength.  $\beta$  should be maximal at  $x = y$  and decay monotonically as  $|x - y|$  increases.
- An overall *death rate*  $\alpha(\mathbf{r})$  is used to tune global, e.g., climatic, changes in living conditions.

Equation (1) gives a rather general framework for spatio-temporal population dynamics. The results depend, in particular, on the spatial distribution of fitness values  $f(x; \mathbf{r})$ . Here we discuss the simplest topography relevant for speciation.

**Two-habitat topography.** Topographical space is assumed to consist of two homogeneous habitats, A and B. In our particular model choice a newborn individual with phenotype  $x$  has fitnesses

$$f^A(x) = f^B(-x) = \exp(-(x - x_0)^2/x_f^2) \quad (2)$$

in habitats A and B respectively, which is schematically represented in Figure 1 a. Habitat A favours positive values of  $x$ , B negative ones. The resource qualities in both habitats decay on a scale  $x_f$  and have relative difference  $2x_0$ . For simplicity



**Fig. 1.** Schematic representation of habitat quality (panel a) and phenotypic interactions (panel b). Habitat A favours phenotypes with larger values of  $x$  (maximal at  $x_0$ ), habitat B offers maximal resource quality to phenotype  $-x_0$ . The resource curves decay on a lateral scale of  $x_f$ . Competition and mate preference are maximal between equal phenotypes and decay on scales of order  $x_\beta = 1$  (top curve) and  $x_\mu$  respectively (bottom curve).

we assume symmetry between the sign of  $x$  and interchange of A and B. Without loss of generality we take a Gaussian decay for the habitat quality. All properties of environmental quality are then expressed by these two parameters.

It is useful to define an order parameter  $\rho \in [0, 1]$  describing the tendency of the population to settle into the more favourable habitats,

$$\rho = \frac{\int_{x>0} (N^A(x) - N^B(x))}{\int_{x>0} (N^A(x) + N^B(x))}, \quad (3)$$

where  $N^A(x)$  denotes the number density of individuals with trait  $x$  in habitat A, and  $N^B(x)$  the same for B. Without any habitat preference, both are equally populated and  $\rho = 0$ . Populations fully retreated into one habitat, i.e., those with positive  $x$  to A and negative  $x$  to B, give  $\rho = 1$ .

The distribution of subpopulations, and hence the value of  $\rho$ , follow from local fitness differences and the population flux from one habitat to another. These quantities are described by the first two term of Equation (1). The diffusion constant  $\nu$  and the transport coefficient  $\lambda$  are subject to evolutionary changes. As a consequence also  $\rho$  changes under evolution. It turns out to be a simpler approach to take  $\rho$  itself as the primary evolving variable. This will be specified below in detail, once the other model components have been introduced.

**Competition.** Competitive interaction decays with phenotypic distance, see the top curve in Figure 1 b,

$$\beta(x-y) = \exp\left(-\frac{(x-y)^2}{x_\beta^2}\right). \quad (4)$$

Again the Gaussian shape is a particular choice without any qualitative difference to others. Moreover, we use competition to set the scale in phenotype space by choosing  $x_\beta = 1$ . This gives  $x_f$  in Equation (2) a simple interpretation as an estimate for the number of subpopulations that can coexist, or in other words for the number of possible ecological niches.

**Reproduction rate.** An expression for the reproduction rates  $R$  is straightforward to construct. As limiting factor we take the breeding capacity of the females, which should be a reasonable assumption in many cases. The number density of individuals of type  $x$  born per time unit is then given by

$$R(x) = \int_{y,z} C(x; y, z) m(z; y) N(y). \quad (5)$$

One has to sum up all possible couples of parents.  $N(y)$  denotes the number of mothers,  $z$  is the phenotype of the males, and  $m(z; y)$  the probability for a  $y$ -female to mate a  $z$ -male.

**Inheritance and quantitative genetics.**  $C(x; y, z)$  is a genetic tensor giving the probability that a couple of parents  $y$  and  $z$  will have  $x$ -offspring. This needs the normalisation  $\int_x C(x; y, z) = 1$ . It reflects the underlying genetic representation of the trait with respect to which we study the possibility of speciation. Different explicit forms for  $C$  are possible, but as  $x$  is a quantitative trait depending on many loci on the genome, we can make use of the principles of quantitative genetics [7,12]. The distribution  $C(x; y, z)$  of offspring trait values has a mean close to  $(y + z)/2$  and a variance that depends on the number of loci involved.

**Mate choice.** Of course  $m(z; y)$  depends on population sizes, the frequency of males to choose from, as well as the absolute number. In too sparse a population females may lack suitable mating partners or be forced to choose against their preference, which may cause interesting effects [13,14]. In this work we focus on what sometimes is called *saturated mating*, where each female finds a partner, so  $\int_z m(z; y) = 1$  for all  $y$  and consequently  $\int_x R(x) = \int_x N(x)$ . Other than on the number of available males,  $m(z; y)$  depends on the preference of females for certain types of males, which are expressed in terms of preference factors  $\mu(z; y)$ . Similar to competition, in our model female attraction towards a given male depends on their mutual phenotypic distance,

$$\mu(z; y) = \exp\left(-(z - y)^2/x_\mu^2\right), \quad (6)$$

relative to a distance  $x_\mu$  which we define as the *mating range*. An example with relatively narrow mating range is shown in the bottom curve of Figure 1 b. The actual mating probabilities then are given by the numbers of available males weighted by the female preference factors,

$$m(z; y) = \frac{\mu(z; y)N(z)}{\int_{z'} \mu(z'; y)N(z')}. \quad (7)$$

Here it becomes clear how the mating range  $x_\mu$  interpolates between indifferent or “random” mating with  $x_\mu = \infty$  and strong mating preference or assortativity as  $x_\mu \rightarrow 0$ .

**Discrete phenotypes.** As one more simplifying step trait space is discretised into a finite number of “bins”. Three bins, where the population at a point is described by the numbers  $N_1$ ,  $N_2$ , and  $N_3$ , are needed in the simplest case to tell the difference between a population split into two independent subpopulations (when  $N_2 = 0$ ) and a contiguous population profile extending over the whole range (when  $N_2 > 0$ ).

The habitat preference order parameter  $\rho$  defined in Equation (3) gains a simple meaning:  $(1 + \rho)/2$  is the fraction of 1-individuals in habitat A, and of 3 in B. Diffusion and migration between the habitats, still explicitly present in Equation (1) are fully included into the spatial order parameter  $\rho$  in this simple discrete case.

There is a simple way to mimic a continuous trait space even when only three bins explicitly appear in the model. Phenotype 2 is located in the middle at  $x = 0$ , but the positions of 1 and 3,  $\pm x$ , are adapted to an optimal value of phenotypic width. At a given value of  $x$  we ask, whether populations at  $x \pm dx$  can invade and suppress the previous ones at  $x$ . If so,  $x$  is replaced by the new value, until a final stable value is reached.

**Evolutionary adaptations and pathways.** So far we have explained the *population dynamics* of competing phenotypes. For fixed external ecological parameters (resource quality and competition) and for fixed internal parameters (mating and habitat preference) the population dynamical equations lead to some stable fixed point describing a population dynamical equilibrium. This can be seen as a resident population subject to the appearance of mutants with different strategies in mating and habitat preference. Generally mutants will be able to invade and push out the previous residents until an evolutionary stable state is reached [6,7,8].

We assume that the rate at which mutants with new characteristics appear is much smaller than the relaxation rate of population dynamics. The same should be true for the rate of global changes in the environment. In Section 4 we present a slow decrease in the external death rate  $\alpha$  from 1 to 0, in order to model a slow increase in habitat quality, e.g., as a result of climatic changes. The slowest changing variable parametrises time in the model.

For instance,  $\alpha$  may change adiabatically and approach a transition point, where a small change in  $\alpha$  causes a large jump in  $x_\mu$ , an example of which is given in Section 4. During this change  $x_\mu$  is the slowest variable and here a natural way to measure “time” is by the rate of mutations in  $x_\mu$ . These evolutionary pathways define the history of evolutionary adaptations.

For a full description one still needs the initial conditions. It makes sense to assume *no* habitat and mating preference,  $\rho = 0$  and  $x_\mu = \infty$ , before any diversification in phenotypic space appears, and to see if spatial structure and mating preference can evolve.

### 3 States of the system

In this Section we characterise the various states of stationary populations in the model and their meaning with respect to speciation.

If the range of resources is narrow, only the middle population is viable,  $N_2 > 0$ , but not the outer ones and therefore  $N_1 = N_3 = 0$ . Then we also find cases of coexistence where all three blocks are populated. The population profile may be structured and show some tendency of splitting into two independent blocks. However, there is gene flow across the population. Besides direct mating between the two opposite outer populations there will be indirect gene flow: If individuals of phenotype 2 mate with both 1 and 3, none of them is isolated from its counterpart. Differentiations can take place in three ways:

- **Spatial separation.** If the populations are fully retreated to their respective more favourable habitats, if 1 lives only in A and 3 in B (in the model parameters this is the case  $\rho = 1$ ), individuals of opposite phenotype just don't meet each other and therefore don't mate. Gene flow then is suppressed by spatial separation.
- **Complete mating assortativity.** Under fully developed mating preference (or assortative mating),  $0 \simeq x_\mu \ll x$ , there will be no cross mating between 1 and 3 and no offspring of intermediate phenotype from such matings.
- **Trait differentiation.** Absence of intermediate traits separates the population profile. In the discrete case:  $N_2 = 0$  and  $N_1 = N_3 > 0$ . Trait differentiation necessarily needs one of the first two types of separation. The inheritance rules give a parent couple of 1 and 3 offspring of type 2 (which is at the arithmetic mean). Absence of the middle phenotype therefore needs absence of mating between 1 and 3.

This shows that the types of differentiation are not independent. In the evolutionary model dynamics spatial separation or assortativity sets in as the *primary* separation. Both cases induce trait differentiation. Spatial separation may also come in as *secondary* differentiation. This case is of particular interest since trait differentiation is stabilised by a *double boundary* interrupting gene flow, spatially and by mate choice.

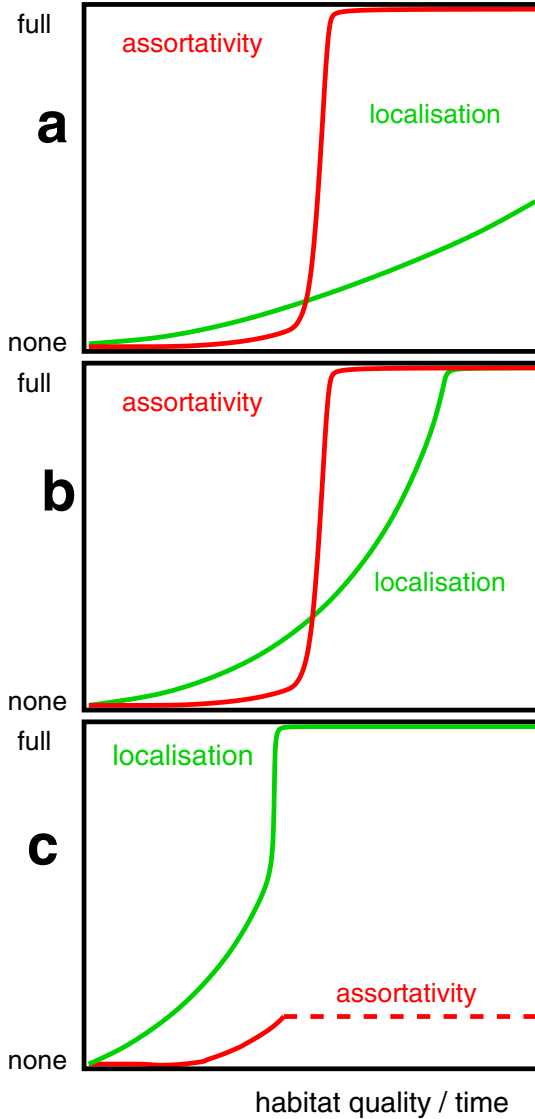
## 4 Evolutionary pathways

Let us now turn to the transitions between these states and present typical sequences of states in an evolutionary context. We focus on a setup inspired by the already mentioned field studies of recolonisation of Europe after the last glacial period [11]. Initially the colonising population finds very poor living conditions, in the model we have the external death rate  $\alpha \simeq 1$ . Under such conditions only the central population at  $x = 0$  is viable, the system is in the state  $N_1 = N_3 = 0$  described first in the previous Section. Now we model slowly warming climatic conditions with increasing habitat quality;  $\alpha$  is gradually lowered and in wider regions of phenotype space populations become viable.

When this region has opened up wide enough, two neighbouring populations will become viable and make better use of the resources than a single one in the middle. This gives selective advantage to any properties in the population enhancing its outer parts. In the previous Section we have seen the different states of the system in which this is achieved. They are reached through evolutionary pathways during which the parameters for mating and habitat preference are adapted. We find three different kinds, which are also illustrated in Figure 2.

- The **sympatric mode** of speciation, shown in Figure 2 a, at small or no habitat heterogeneity. Mating preference slowly develops but remains only partial until at a critical value of habitat quality full assortativity develops with  $x_\mu \rightarrow 0$ . At this point, mating between opposite phenotypes is suppressed and so is its offspring as a source for the population of middle phenotype. On





**Fig. 2.** Three examples of evolutionary pathways, panel **a** and **b** show two varieties of the *sympatric mode* of speciation, panel **c** the *allopatric mode*. **a** has lowest habitat heterogeneity (smallest value of  $x_0$ ) and **c** highest. The slowest time scale with the gradual increase in habitat quality is shown on the horizontal axis. In panels **a** and **b** mating preference switches rapidly from partial to full assortativity when the habitat quality is good enough. At this point the middle phenotypes become suppressed,  $N_2 \rightarrow 0$ . In panel **b** full localisation develops after assortativity. Now the two subpopulations are *doubly isolated*, spatially and by mating preference. In panel **c** localisation develops more rapidly than assortativity and causes  $N_2 \rightarrow 0$ .

its own this phenotype is not viable, the competition load from the two outer ones is too large; so it disappears together with  $x_\mu$ . The entire population is now in a state where 1 and 3 coexist but are reproductively isolated from each other.

- The **sympatric mode with patching**, Figure 2 b, at moderate heterogeneity, where after establishing assortativity there is another transition in the model. If  $x_0$  is large enough, the system will reach a point where  $\rho = 1$  and both 1 and 3 have fully retreated into their respective more favourable habitats. Now there is a double boundary against mating and gene flow between the subpopulations, a sexual and a spatial one.
- The third type of pathway at strong habitat heterogeneity, the **allopatric mode** in Figure 2 c, is different. Here the tendency towards localisation is so strong that the two subpopulations become spatially isolated before assortativity can develop. Patching suppresses cross mating and causes the central phenotype to go extinct. Now mating preference becomes meaningless as different phenotypes do not share any pieces of habitat and  $x_\mu$  remains at the value of partial assortativity reached last.

## 5 Discussion

The main goal of this work is to study the interplay of trait and real space, the internal and external degrees of freedom in a population at the onset of speciation. The model constructed for this purpose shows a complex interaction of its degrees of freedom despite its simplicity. *The pathways of speciation are the combined result of habitat topography, genetics, and ecology.* The traditional alternative of sympatry and allopatry is obsolete.

The evolutionary history is crucial for interpreting observed states of the system. Spatial separation can occur as primary differentiation in the allopatric mode or as a secondary step, after trait differentiation and mating assortativity have already evolved in sympatry.

The effect of spatial heterogeneity on (the possibility of) sympatric speciation is twofold: On one hand it may prevent the evolution of mating preference, as emerging extreme phenotypes retreat into their respective favourite habitat patches before differentiation in sympatry together with mating assortativity has fully developed. On the other hand, if they retreat *after* differentiation, patching forms an additional boundary against gene flow. *Sympatric speciation with subsequent patching is the most efficient way to cut the gene flow between two subpopulations.*

Clearly, the model can be generalised in several ways. For example a more detailed spatial model, closer to Equation (1), can be used to study the profile of the phenotypical population structure at the interface between habitats and the gene flow across the boundary. Moreover, mating can and will be unsaturated in certain cases, in contrast to the definition of Equation (7). If not all females are able to find a desired mating partner, in particular at small population sizes, the transitions between the different states of the system and hence the evolutionary pathways will be modified [13,14].

## Acknowledgements

It is a pleasure to thank Alex Kondrashov, Arne Nolte, Sebastian Steinfartz, and Diethard Tautz for sharing their insight. In particular, Diethard Tautz has pointed out to us the ubiquity of spatial structure in speciating populations.

## References

1. C. Darwin, *The Origin of Species*, John Murray, London (1859).
2. E. Mayr, *Animal species and evolution*, Belknap Press, Cambridge (1963).
3. J. Maynard Smith, *American Naturalist* **100**, 637–650 (1966).
4. M.L. Rosenzweig, *Biological Journal of the Linnean Society (London)* **10**, 275–289 (1978).
5. R. Lande, *Proceedings of the National Academy of Sciences of the USA* **78**, 3721–3725 (1981).
6. S.A.H. Geritz, É. Kisdi, G. Meszéna, and J.A.J. Metz, *Evolutionary Ecology* **12**, 35–57 (1998).
7. A.S. Kondrashov and F.A. Kondrashov, *Nature* **400**, 354–357 (1999).
8. U. Dieckmann and M. Doebeli, *Nature* **400**, 351–354 (1999). Clarendon Press, Oxford (1980).
9. A. Meyer, T.D. Kocher, P. Basasibwaki, and A.C. Wilson, *Nature* **347**, 550–553 (1990).
10. T.C. Johnson, C.A. Scholz, M.R. Talbot, K. Kelts, R.D. Ricketts, G. Ngobi, K. Beuning, I. Ssemmanda, and J.W. McGill, *Science* **273**, 1091–1093 (1996).
11. S. Steinfartz, M. Veith, and D. Tautz, *Molecular Ecology* **9**, 397–410 (2000).
12. M.G. Bulmer, *The mathematical theory of quantitative genetics*, Clarendon Press, Oxford (1980).
13. M. Rost and M. Lässig, in preparation.
14. A. Atik, Diploma Thesis, University of Cologne (2002).