

Univerzita Karlova v Praze, PřF UK  
Katedra filosofie a dějin přírodních věd

# DIPLOMOVÁ PRÁCE

Univerzita Karlova v Praze

Přírodovědecká fakulta

Katedra filosofie a dějin přírodních věd /  
Institute of Science and Technology, Austria

**The Role of Genetic Variance in Speciation /  
Role genetické variance ve speciaci**



**Pavel Payne**

Praha 2011

Vedoucí diplomové práce: Mgr. Jitka Polechová, PhD.

Garant: doc. RNDr. Anton Markoš, CSc.

# Obsah

1. Prohlášení .....	1
2. Poděkování .....	2
3. Abstract.....	3
3.1. Abstrakt (česky).....	3
4. Keywords / Klíčová slova.....	5
5. Introduction .....	6
5.1. Models of sympatric speciation by natural selection.....	7
5.1.1. Levene's model.....	9
5.1.2. Maintenance of polymorphism.....	10
5.1.3. Sympatric speciation and Levene's model.....	11
5.1.4. Polygenic traits and the infinitesimal model .....	12
5.2. The role of epistasis.....	13
5.2.1. Epistasis and selection on quantitative traits.....	14
5.2.2. Quantitative traits in heterogeneous environment .....	16
6. The model .....	18
6.1. Maintenance of polymorphism.....	20
6.1.1. Asymmetry in niche proportions .....	20
6.1.2. Asymmetry in strength of selection.....	24
6.1.3. Asymmetry in niche proportions and strength of selection .....	25
6.1.4. Three-locus model.....	26
6.1.5 Functional fitness .....	28
6.2. Summary of the results.....	34
7. Discussion .....	36
8. Conclusion .....	38
9. References .....	39
10. Source code .....	44

# 1. Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně. Všechny použité prameny jsou uvedeny v závěru, v seznamu literatury.

V Praze, dne 28.8.2011

podpis ... ..

## 2. Poděkování

Na tomto místě bych rád poděkoval všem, kteří mi nejen při psaní této práce poskytli cenné rady a ještě cennější kritiku.

- Na prvním místě bych chtěl uvést svou školitelku, Mgr. Jitku Polechovou, PhD., jejíž kritické připomínky, všestranné znalosti a nezměrná trpělivost pomohly posunout práci kýženým směrem.
- Dále patří můj vřelý dík panu prof. Nicku Bartonovi za velmi cenné rady ať už koncepční nebo týkající se odborných detailů, a za trpělivé korekce anglického jazyka.
- Katedře filosofie a dějin přírodních věd za obecné rozšíření obzorů v oboru Teoretické a evoluční biologie a za podporu a toleranci při mém pobytu na zahraniční stáži.

### **3. Abstract**

Sympatric speciation has received much attention both empirically and theoretically. However, the contribution of sympatric speciation to biodiversity remains unclear. One piece missing from the speciation puzzle is the plausibility of sympatric ecological divergence of species through adaptation in polygenic traits. I consider an environment consisting of two niches, where one value of the trait is advantageous in only one niche, and vice versa. The selection regime is described by a trade-off in viabilities between the niches. These polygenic traits can, and often do, involve epistatic interactions among and between loci, so that the contribution of the alleles to viability deviates from additivity. Epistasis then also affects the curvature of the trade-offs: predominant less-than-additive epistasis turns the curve towards concavity and predominant more-than-additive towards convexity. The curvature of the trade-off plays a crucial role in the evolution of populations. With a convex trade-off, extreme values of the trait are favored and the population tends to diverge, but relatively stringent symmetry in strength of selection within the niches and the niche proportions is necessary to maintain polymorphism. In this study I use two and three-locus haploid versions of Levene's model to investigate various epistasis-selection regimes and explain characteristics of the trade-off curve affecting divergence and maintenance of polymorphism. I compare the results of a multi-locus model with predictions based on the infinitesimal model and identify the limits of both approaches.

### ***3.1. Abstrakt (česky)***

Sympatrická speciace byla významným předmětem zájmu jak z hlediska empirického tak teoretického, nicméně její příspěvek k biodiverzitě stále zůstává nejasný. Jedna z chybějících částí této skládačky je možnost sympatrické ekologické divergence druhů na základě jejich adaptace v polygenních znacích. V této studii uvažujeme prostředí skládající se ze dvou nik, přičemž jedna hodnota polygenního znaku je výhodná v jedné nice a druhá hodnota ve druhé nice. Seleční režim je pak popsán pomocí trade-off v tomto znaku mezi dvěma nikami, a díky tomu, že tyto polygenní znaky mohou, a často též zahrnují epistatické interakce mezi geny, příspěvek jednotlivých alel k danému znaku se odchyluje od aditivity. Tato epistáze pak může ovlivnit zmíněné trade-off, takže převládající přítomnost méně než aditivní (negativní) epistáze způsobuje odchýlení trade-off směrem ke konkávní křivce, zatímco přítomnost více než aditivní (pozitivní) epistáze odchyluje trade-off směrem k vyšší konvexitě. Zakřivení trade-off hraje klíčovou roli v evoluci populací, přičemž pod konvexním trade-off jsou zvýhodněni nositelé extrémních hodnot znaku a populace má tendenci k divergenci. V tomto případě však je vyžadována relativně přísná symetrie v síle selekce v jednotlivých nikách a jejich velikostech, aby se v populaci udržel polymorfismus. V této studii používáme dvou a tří-lokusovou haploidní verzi Leveneho modelu k prozkoumání rozličných režimů epistáze potažmo selekce, a přibližujeme charakteristiky trade-off, které ovlivňují divergenci a udržení polymorfismu. Tyto výsledky následně srovnáváme s predikcemi založenými na infinitesimálním modelu a identifikujeme limity obou přístupů.

#### **4. Keywords / Klíčová slova**

Sympatric speciation, maintenance of polymorphism, polygenic trait, Levene model, epistasis, trade-off, ecological divergence, disruptive selection, infinitesimal model, evolutionary genetics, asymmetry, niche proportions, strength of selection.

Sympatrická speciace, udržení polymorfismu, polygenní znak, Leveneho model, epistáze, trade-off, ekologická divergence, disruptivní selekce, infinitesimální model, evoluční genetika, asymetrie, proporce nik, síla selekce.



## 5. Introduction

Biodiversity may arise through ecologically driven divergence. Populations can react to natural selection through inherited variation in fitness between their individuals. According to Fisher's (1930) Fundamental Theorem, mean fitness increases through time and populations become better and better adapted to their actual environment. Several recent studies showed that this ecological adaptation to novel environments can occur very rapidly (Orr and Smith 1998). Then, if a population is divided by some environmental barrier, which prevents dispersal of individuals, the subpopulations accumulate different sets of alleles because of different selection pressures and random genetic drift. Through this accumulation of different alleles, reproductive isolation may emerge (Dobzhansky 1937, Muller 1940) and complete the process of allopatric speciation.

Can populations diverge even while occupying a single heterogeneous environment, i.e. in sympatry? Darwin (1859) considered sympatric origin of species by ecological divergence and an advantage of specialists over generalists. Since then, sympatric speciation has been of great interest to evolutionary biologists and remains consistently controversial (Jiggins 2006). After the development of population genetics and mathematical modeling, many and varied scenarios of sympatric speciation have been proposed (for reviews, see Kirkpatrick and Ravigné 2002, Gavrillets 2004), but although they show that sympatric speciation can occur, very stringent assumptions are usually required. Another complication in quantifying how frequent sympatric speciation can be in nature is the possibility of secondary allopatry and shift of the species' range. Most of the species considered to arise in sympatry have been found in isolated habitat "islands", such as lakes, host-specific parasites or in host races (Coyne and Orr 2004, ch. 4). However, without such well-isolated habitats, species' ranges can shift relatively quickly after any kind of ecological adaptation (Angert et al 2011). Hence, even though a species may arise in sympatry due to ecological divergence, secondary allopatry is not unlikely.

Although allopatry is considered to be the main source of global biodiversity, its contribution may be over-estimated. Allopatry has long been generally accepted as a null model of speciation (Mayr 1963, Coyne and Orr 2004), but it is hard to falsify it and the failure to reject allopatry does not mean that it is supported (Bolnick and

Fitzpatrick 2007). Despite the fact that allopatric speciation still is considered as the main source of new species, the contribution of sympatric speciation to global biodiversity remains unclear (for review, see Via 2001 or Fitzpatrick et al. 2008).

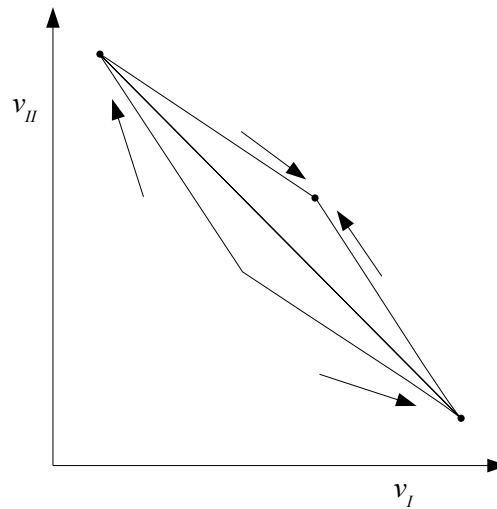
### ***5.1. Models of sympatric speciation by natural selection***

Even though methods and approaches to sympatric speciation by natural selection modeling vary, nearly all of the models share some necessary common properties. The population has to be initially panmictic, consist of individuals that vary in some heritable trait and this variation has to be somehow maintained. Selection has to be disruptive, and so favour specialists over generalists. After the population diverges into two or more distinct types, some reproductive isolation has to be somehow established between the emergent subgroups.

Since I consider ecological disruptive selection, the environment where individuals dwell has to be heterogeneous and carriers of one extreme value of the selected trait are advantaged in one type of the environment and vice versa. Thus, I can say that the environment consists of two niches and depict these varying performances of individuals as a trade-off between viabilities in one and the other niche. Moreover, the appearance of such trade-offs between various traits, trait values or behavioral patterns is quite general and has long been investigated by ecologists (Stearns 1989).

The curvature of such trade-offs (fig. 1) plays a crucial role in the evolution of populations, and so from now on I will use this concept to describe some general properties of the populations. I consider the trade-offs only within one trait. The trait has a range of values, for example an ability to explore one or another source, and individuals, which do well on one source and worse on the other. The relationship, how worse they do on the other source and whether individuals with an intermediate value of the trait are favored over the specialists can be described by the trade-off curve. In general, if the trade-off is linear, there is no difference in overall strength of selection on the two extreme phenotypes and the intermediate phenotype, and so the population stays in equilibrium. If the trade-off curve is concave, natural selection favors intermediate over extreme phenotypes, i.e. selection is stabilizing and being a generalist is an evolutionary stable strategy (ESS). On the other hand, in case of convex trade-off, extreme phenotypes are favored over the intermediate ones, so the selection is disruptive (i.e. being a specialist is an ESS). However, the curvature of the trade-off also affects an ability of the population to maintain polymorphism (through

advantage or disadvantage of the intermediate phenotypes). Therefore it is necessary to deal with these two features together.



**Figure 1.** Trade-off between viabilities in two niches. For a one-locus diploid model, the middle line represents no dominance, the convex curve represents niche dependent underdominance and the concave curve represents niche dependent intermediate dominance of beneficial alleles in each particular niche.

### 5.1.1. Levene's model

Many models of sympatric speciation are based on Levene's (1953) model. This describes the life cycle of a very large population of organisms capable of surviving in a heterogeneous environment. Individuals mate in a common mating pool, the offspring are randomly spread into two niches and generations are non-overlapping. There are two types of alleles and one or the other is beneficial in each of the niches. Selection acts within the niches, the individuals who survive are mating in the common mating pool again and the contribution of each niche is proportional to its size.

In the original Levene model, the ability of individuals to survive in one or the other niche is determined by one biallelic diploid locus with complete dominance, where each allele is beneficial in one of the two niches. Thus, this model maintains polymorphism through an advantage of heterozygotes over homozygotes. The degree of dominance can be described by the trade-off between viabilities in niches. The linear trade-off between viabilities in Figure 1 corresponds to the case with no dominance, i.e. both heterozygotes are generalists and perform equally well in both

niches. The concave trade-off corresponds to intermediate dominance and mean fitness of heterozygotes is higher than mean fitness of homozygotes (i.e. heterozygote advantage). The convex trade-off corresponds to underdominant alleles, so homozygotes are fitter than heterozygotes and selection is disruptive.

The model can also equally well work with haploids, where just two types described by their viabilities in niches are present. These two types are described by points on the trade-off curve and its curvature tells us, when the coexistence of the two types is an evolutionary stable strategy (ESS). If I introduce new types, the ESS is coexistence of two types if the curve is convex.

### **5.1.2. Maintenance of polymorphism**

As mentioned above, maintenance of polymorphism is essential for successful sympatric divergence. To understand how frequent sympatric speciation can be in nature, it is necessary to analyze under which conditions polymorphism is maintained in our models. The conditions for maintenance of polymorphism in Levene's model have been partially analyzed but most of these studies assumed symmetry either in niche proportions or in the selection coefficients in niches. Necessary and sufficient conditions for maintenance of polymorphism in various versions of Levene's model without epistasis was derived by Prout (1968), Gliddon and Strobeck (1975), Strobeck (1979), Nagylaki (2009) and Bürger (2009). The fact that asymmetry in niche proportions significantly influences the ability of the model to maintain genetic polymorphism has been investigated by Maynard Smith (1966), Maynard Smith and Hoekstra (1980) and Hoekstra et al (1985). To quantify this ability they introduced a parameter called "robustness" describing the fraction of parameter space in which genetic polymorphism is maintained. As they focused on the range of niche proportions, this robustness is in fact the same as Gavrilets's (2004, pp. 237) threshold of absolute value of the difference in niche sizes. Maynard Smith and Hoekstra (1980) and Hoekstra et al (1985) have analyzed Levene's and several Levene-like diploid models with one locus and two alleles and showed that the robustness is very sensitive to the niche dependent degree of dominance. They conclude that robustness increases if the fitness of heterozygotes is higher than arithmetic mean of fitnesses of homozygotes in each niche (i.e. heterozygote advantage and concave trade-off in fig. 1). In terms of dominance, robustness is

higher if the beneficial allele in each particular niche is partially or completely dominant. Conversely, underdominant beneficial alleles reduce robustness in comparison with alleles with intermediate dominance (i.e. convex trade-off). They also showed that habitat choice slightly inflates the robustness.

Another widely used simplification in analyses of such models is the assumption of equal strength of selection in both niches for beneficial alleles. Although it eases the analysis by getting rid of one parameter, presumption of this symmetry is unrealistic and it also helps to maintain polymorphism. This was partly analyzed by Maynard Smith (1966) and Maynard Smith and Hoekstra (1980) who also showed that with larger differences in niches sizes, stronger selection in the smaller niche is needed. Substantial difference in strength of selection is possible only if the selection is very strong because stability is second order in selection. Hence, for interpretation of the models of sympatric divergence, this feature has to be taken in account.

### **5.1.3. Sympatric speciation and Levene's model**

To complete the process of sympatric speciation, some mechanism providing nonrandom mating is necessary. Maynard Smith (1966) first studied a population genetic model of sympatric speciation extending Levene's (1953) model with habitat preference and first showed that sympatric speciation can occur under certain, even though restrictive conditions. Thenceforward, many other models and scenarios were introduced (reviewed and extended in Gavrilets 2004, ch.10) and showed that sympatric speciation is theoretically possible under certain conditions.

Although the conditions seem to be stringent and have been criticized as being biologically unrealistic, Bolnick and Fitzpatrick (2007) showed that at least three assumptions are relatively often met in nature. They focused on the general necessary properties of sympatric speciation models, particularly frequency-dependent disruptive selection, assortative mating/ecology pleiotropy and whether assortative mating imposes weak or no costs. However, although they showed that such assumptions are not implausible, many arguable points remain unresolved. This criticism is mainly aimed at highly symmetrical and polymorphic initial conditions, constant ecological parameters, very stringent assortative mating, a stable

environment and resource distribution, and soft selection (Bolnick and Fitzpatrick 2007).

As I have mentioned in the previous section, symmetry in niche sizes or equal strength of selection within niches ensures that polymorphism will be maintained, because the models are typically stable only over a narrow parameter range around the symmetric values. However, the vast majority of models of sympatric speciation are based on such restrictive assumptions, so their credibility is arguable, because even small perturbation can cause fixation of one of the extreme phenotypes.

#### **5.1.4. Polygenic traits and the infinitesimal model**

An arguable assumption frequently made in many sympatric speciation models is their restriction to one or rarely several loci, whereas ecological adaptation of a population often involves a gradual change in a polygenic trait (Rose et al. 2000, Wilson and Turelli 1986). An obvious reason for this common restriction is the computational difficulty of such models, especially if additional evolutionary forces, such as epistasis and nonlinear trade-off between viabilities are included. A multilocus version of Levene's model has been analyzed (Nagylaki 2009, Bürger 2010) but no epistasis was assumed.

Another framework for modeling evolution of polygenic traits was introduced by Fisher (1918) and extended by Bulmer (1980). This approach aimed at modeling a response of populations with a polygenic trait to selection. It does not require detailed knowledge of the underlying genotype frequencies and effects but assumes a linear parent/offspring regression. The model assumes very large (effectively infinite) number of loci, each with infinitesimal effect leading to Gaussian distribution of phenotypes with fixed variance. Although this approach allows us to understand the evolution of such polygenic traits, it also faces certain limitations. The evolutionary response of the population is not driven by change in allele frequencies of underlying loci, but solely by changes in associations among loci, whereas the allele frequencies are implicitly assumed constant.

Barton (2010) analyzed a model where evolution of a population in heterogeneous environment is described by the infinitesimal model. The evolution of the population under various selection schemes was described by trade-off between

viabilities in two niches. He concludes that with a concave trade-off the population evolves a single generalist genotype, whereas with a convex trade-off, disruptive selection leads to two reproductively isolated specialists if variance is high enough. As I have mentioned above, at least for one locus diploid Levene's model, polymorphism is maintained only over a narrow parameter range around the symmetric values. Hence, the assumption of fixed variance might not be appropriate in a large parts of parameter space.

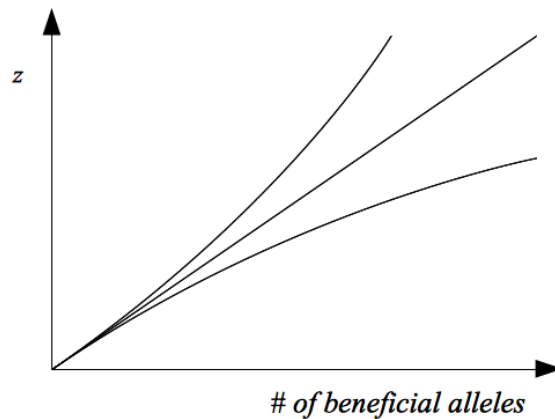


## ***5.2. The role of epistasis***

It has long been clear that epistasis could play a key role in the process of speciation. The fact that some loci can influence the effect of some other loci on the phenotype has been discussed already by founders of evolutionary genetics (Poulton 1903, ch.2; Fisher 1918) and is fundamental to Wright's shifting-balance theory (1931, 1932). Dobzhansky (1937) and Muller (1940) proposed a theory of speciation based on gradual accumulation of different alleles in two separate diploid populations. During allopatry, selection provides that all the alleles in both populations must be compatible with each other, i.e. heterozygotes carrying new mutation cannot have reduced fitness. If the populations come into secondary contact, some combination of new alleles from one or the other population may not be compatible, because these combinations have never been "tested" by selection. Thus, these hybrids have reduced fitness and some postzygotic isolation can emerge. Since then, epistasis was considered an important evolutionary force many times and it appeared in many either verbal (Mayr 1954, Carson 1968, Templeton 1980) or mathematical (e.g. Goodnight 1987, 1988; Cheverud and Routman 1996, Turelli and Orr 2000, Lopez-Fanjul et al. 2002, 2004, Turelli and Barton 2006) theories of evolution of populations.

### **5.2.1. Epistasis and selection on quantitative traits**

Although epistasis is generally accepted as a key evolutionary force, its consequences for the evolution of polygenic traits are still not clear. Generally, one can imagine two different types of epistatic interactions at the developmental, i.e. genotype-phenotype level (biological epistasis defined by Moore and Williams 2005). The contribution of each additional allele (beneficial mutation) to the trait value can be either lower or higher than if the allele was present alone, so the effect can be either less-than-additive or more-than-additive. With increasing number of alleles and only more-than-additive epistasis, the value of the trait increases faster than linearly, and conversely, with less-than-additive epistasis the value increases more slowly (Fig. 2). If I consider only pairwise interactions, the deviation from additivity of  $n$ -th allele is proportional to  $\varepsilon \cdot \binom{n}{2}$ , where  $\varepsilon$  is the epistatic coefficient.



**Figure 2.** Possible effects of increasing number of alleles on a polygenic trait. On the x-axis is increasing number of beneficial alleles and on the y-axis the value of the trait,  $z$ . The upper curve represents more-than-additive effect, the middle one represents pure additivity and the lower curve less-than-additive effect of alleles on a polygenic trait.

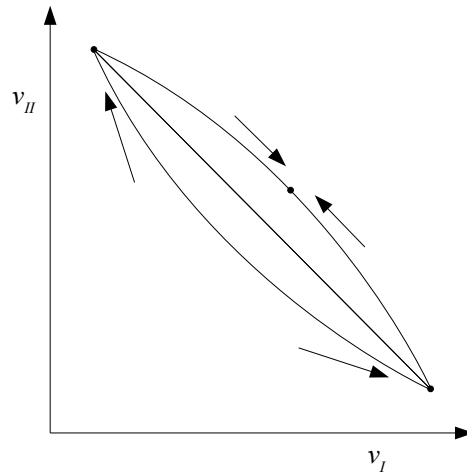
Two recent studies (Khan et al 2011, Chou et al 2011) showed that adaptation to a novel environment through new beneficial mutations in bacterial strains exhibit mostly less-than-additive epistasis in their contribution to fitness. Also Eshed and Zamir (1996) showed that in some additive traits less-than-additive developmental epistasis predominates in tomatoes. The effect of chromosome segments of tomato *Lycopersicon pennelli* introgressed into an otherwise homogeneous genetic background of *Lycopersicon esculentum* is epistatic in 28% of tested interactions ( $P < 0.05$ ). Detected epistasis was predominantly less-than-additive (slopes of regressions in plots of expected vs. observed effects were from 0.71 to 0.79; their figure 3), and was measured for four yield-associated additive traits (plant weight, fruit mass, content of sugar and yield). However, in the evolution of natural populations, the value of a polygenic trait is irrelevant unless natural selection acts on it. And as well as the trait change can deviate from linearity, natural selection can act on the trait nonlinearly. So, the individuals with more beneficial additive alleles and therefore a higher value of the trait might be advantaged either more or less than what would correspond to this trait value if the per-locus strength of selection was constant. From now on, I will call epistasis at this phenotype-environment level epistatic selection. Thus, if I want to understand the evolution of such a polygenic trait, I have to take into account both levels of epistasis and consider the total outcome as their combination. This combination can be understood as the effect of the developmental

epistasis on viabilities of individuals, and in this study, I will consider epistasis as this combination.

### **5.2.2. Quantitative traits in heterogeneous environment**

If I assume a heterogeneous environment, as in Levene's model, one value of the trait can be beneficial in one type of the environment and another value in the other type. Then, I can depict different viabilities of individuals within different niches as a trade-off similarly as for dominance in Fig. 1. The trade-offs in these polygenic traits represent a boundary to the set of all possible phenotypes and various interactions within and among genes, and other epigenetic and environmental influences cause that the most of phenotypes lie under this curve. However, if I neglect all the other influences than epistasis, and assume always only one type of epistatic interactions, I can say that the phenotypes lie on the trade-off and so the trade-off is an interpretation of the underlying genetics.

In the case of no epistasis, i.e. pure additivity of the effects of alleles on viabilities, this trade-off will be linear and all individuals will perform equal mean fitness over all niches (middle line in Fig. 3). The less-than-additive effect of alleles at the genotype-phenotype level, will cause a deviation of the trade-off curve from linearity towards concavity (upper line in Fig. 3). Conversely, more-than-additive epistasis will push the trade-off curve towards convexity (lower line in Fig. 3).



**Figure 3.** Trade-offs between viabilities in two niches. The lower convex curve represents more-than-additive developmental epistasis in a polygenic trait and/or disruptive epistatic selection. The upper concave curve represents less-than-additive developmental epistasis and/or stabilizing epistatic selection. The middle line represents the case without epistasis.

Thence, the convex trade-off will result in disruptive selection favoring carriers of extreme values of the trait and the concave trade-off will result in stabilizing selection, favoring generalists.

At the developmental level, one can imagine almost arbitrary epistatic interactions, which may lead to increase of phenotypic variation in the trait (Weber et al. 1999, Cheverud and Routman 1996, Cheverud and Routman 1995), but not necessarily to a deviation of the trade-off curve from linearity. However, if there is a pattern in how loci epistatically interact, the trade-offs tend to be either convex or concave and thence affect evolution of the population.

In this study, I focus on the two clearcut cases, purely convex versus concave trade-offs. Particularly, I analyze how various degrees of convexity or concavity affect maintenance of polymorphism in a population in two and three-locus haploid version of Levene model with mating within niches. Then I compare the results with predictions for divergence in a quantitative trait approximated by the infinitesimal model (Barton 2010).

## 6. The model

Since we focus on the effect of epistatic interactions between and among loci, we analyze a haploid biallelic version of Levene's model with mating within niches, which is more favourable to speciation. The environment consists of two niches, where alleles of the first type (small letters) are beneficial in one niche and alleles of the second type (capital letters) in the other niche. For further analysis, it is convenient to define the fitness to be between zero and one, so the presence of alleles which are deleterious in the particular genotype in niche  $\gamma$  reduce its fitness by selection coefficient  $s_\gamma$ . Epistasis ( $\epsilon_\gamma$ ) is defined as the deviation from additivity of deleterious alleles in each niche if more than one of those alleles is present. Thus, the relative advantage of having all beneficial alleles (i.e. fitness equal to 1) is either higher or lower than what would correspond to purely additive effects of the alleles. Here we are assuming a special kind of symmetric epistasis, which acts both among alleles beneficial in one niche and among alleles beneficial in the other niche. As we have discussed in the introduction, the scheme of selection can also include deviations the effects of alleles from additivity, so from now on, we will use the concept of trade-offs (Fig. 3) to describe the state of the epistasis-selection scheme. Less-than-additive epistasis is then represented by a positive epistatic coefficient and leads to a concave trade-off, which implies stabilizing selection. More-than-additive epistasis is represented by a negative epistatic coefficient and leads to a convex trade-off and thence to disruptive selection.

Generally, the trade-offs represent an upper bound to the set of all possible phenotypes. Thus, only the most fit genotypes lie on that curve and the other are somewhere in the  $\{v_1, v_2\}$  space below this curve. However, in this study, we use it to describe the actual trade-off between viabilities of individuals defined by exact individual fitnesses (or later functions). Because these fitness definitions do not contain more complex epistatic interactions than the special symmetric ones, all genotypes lie on the trade-off curve.

Genotype frequencies in the mixing pool are given by  $g(X)$ , where  $X$  represents each particular genotype. Then, individuals are randomly distributed into the niches and are under niche dependent natural selection which depends on their fitnesses  $w_\gamma(X)$ , so the genotype frequencies in the niches after selection are  $g'_\gamma(X) = w_\gamma(X) g(X)$ . Mating and recombination of the rate  $r$  occurs within niches after selection and each niche contributes by newborns to the mixing pool proportionally to its size,  $c_\gamma$ . In this two niche

model  $c_I = c$  and  $c_{II} = 1 - c$ .

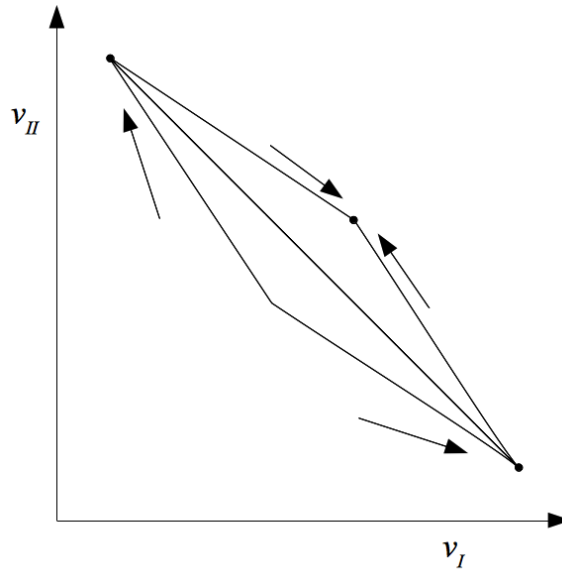
From now on, we will mainly operate with allele frequencies and linkage disequilibria, which are more convenient for the analyses. So,  $p_i$  is the frequency of an allele beneficial in one niche (small letters) in the  $i$ -th locus. Because frequencies are defined to be between 0 and 1, the frequency of allele beneficial in the other niche (capital letters) is  $q_i = 1 - p_i$ .

We begin by analysis of the simplest version of the model, the haploid biallelic two-locus model. Fitnesses of all genotypes are defined in Table 1.

genotype	niche I	niche II
AB	$1 - 2s_I - \epsilon_I$	1
Ab/aB	$1 - s_I$	$1 - s_{II}$
ab	1	$1 - 2s_{II} - \epsilon_{II}$

**Table 1.** Fitnesses of individual genotypes.  $s_I$  is selection coefficient for niche specific deleterious alleles in niche I,  $s_{II}$  is selection coefficient for niche specific deleterious alleles in niche II and  $\epsilon_I$  and  $\epsilon_{II}$  are epistatic coefficients which are defined as the deviation from additivity of the two deleterious alleles present in a genotype together. Positive  $\epsilon_y$  ( $\epsilon_y > 0$ ) represents less-than-additive epistasis (and results in the concave trade-off) and negative ( $\epsilon_y < 0$ ) represents more-than-additive epistasis (which results in the convex trade-off).

The curvature of the trade-offs in this two-locus model is discrete and is depicted in Fig. 4. Because the curvature of the trade-off also affects maintenance of polymorphism in the population, and polymorphism is necessary for any kind of speciation, we analyze in which region of parameter space is this polymorphism maintained.



**Figure 4.** Trade-offs between viabilities in niches. The lower convex trade-off ( $\epsilon_y < 0$ ) results in an advantage of individuals carrying extreme genotypes  $AB$  or  $ab$  relatively to the intermediate genotypes,  $aB$  or  $Ab$ . The upper concave trade-off ( $\epsilon_y > 0$ ) favours intermediate genotypes and despite of the presence of two different environments a generalist predominate.

## ■ 6.1. Maintenance of polymorphism

### ■ 6.1.1. Asymmetry in niche proportions

First we analyzed the range of niche proportions,  $c$ , where genetic polymorphism is maintained in model without epistasis, i.e. with the linear trade-off between viabilities in niches. For simplicity we assumed equal strength of selection and strength of epistasis in both niches ( $s_I = s_{II} = s$ ,  $\epsilon_I = \epsilon_{II} = 0$ ). The model is in equilibrium when all allele frequencies in two following generations are equal, so, when the condition  $p_{i(t+1)} - p_{i(t)} = 0$  is satisfied. These equations have three different roots, where two of them refer to simple monomorphic equilibria ( $\hat{p}_i = 0$  and  $\hat{p}_i = 1$ ) and the third one represents a polymorphic equilibrium. This model is in an equilibrium if the equation

$$\frac{1 - 2c(1-s) - p_1s - p_2s}{2(1-s)} = 0$$

is satisfied.

This polymorphic equilibrium is stable until allele frequencies reach their boundaries at  $\hat{p}_i = 0$  and  $\hat{p}_i = 1$  and it is if

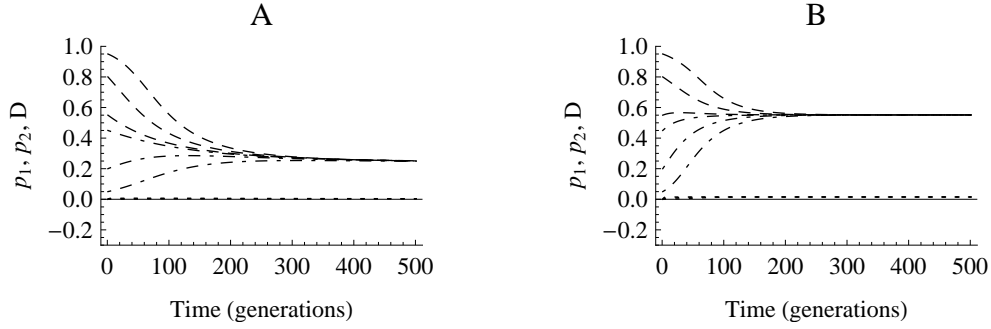
$$\frac{1-2s}{2-2s} < c < \frac{1}{2-2s}.$$

If we want to analyze the model with epistatic interactions involved ( $\epsilon_I = \epsilon_{II} = \epsilon$ ), the analysis of the polymorphic equilibrium becomes unsolvable.

Because we assume recombination between loci ( $0 < r < 0.5$ ), the population stays polymorphic only if the trade-off is convex. With the concave trade-off, extreme genotypes suffer from the epistasis which lowers their fitness and if the allele frequencies are not precisely equal, one of the two intermediate genotypes fixes. With the convex trade-off, the individuals carrying genotypes with alleles of the same type are favoured in one or the other niche.

So, under the convex trade-off, we can assume allele frequencies to be equal ( $p_1 = p_2 = p$ ) because stability of the equilibria does not depend on the initial allele frequencies. The equilibrium frequencies depend only on the proportions of niches and the strength of selection in the niches (Fig. 5). But even in this simplified case, the exact solution for a polymorphic equilibrium is the solution of equation  $p_{(t+1)} - p_{(t)} = 0$  for  $p$  and it is a polynomial of degree five, which has five generally unsolvable roots. Two of them represent monomorphic equilibria ( $\hat{p} = 0$  and  $\hat{p} = 1$ ), so the model has at most three polymorphic equilibria. Numerical simulations showed, that under realistic values of parameters ( $c \in (0,1)$ ,  $s \in (0, 1)$ ,  $\epsilon \in \pm (0, s)$ ) two of them are mostly complex and only one of them goes up in realistic values between  $\hat{p} = 0$  and  $\hat{p} = 1$ .





**Figure 5.** Plots of numerical simulations with different initial allele frequencies. Initial allele frequencies are 0.05, 0.2 and 0.45 for allele A in locus 1 (dot-dashed curves) and 0.95, 0.8 and 0.55 for allele B in locus 2 (dashed curves). The dotted lines represent linkage disequilibria between loci. A) Convex trade-off,  $\epsilon_\gamma = 0.25 s_\gamma$ ,  $s_I = s_{II} = 0.2$ ,  $c = 0.52$ ; B) Convex trade-off,  $\epsilon_\gamma = 0.25 s_\gamma$ ,  $s_I = 0.3$ ,  $s_{II} = 0.2$ ,  $c = 0.5$ .

Therefore, we can analyse an instability of the two monomorphic equilibria ( $\hat{p} = 0$  and  $\hat{p} = 1$ ) and assume that between these two monomorphic equilibria only one realistic polymorphic equilibrium exists. This approach was used by Levene (1953), Prout (1968) and Hoekstra et al (1985). Gliddon and Strobeck (1975) analyzed the haploid multilocus analogue of Levene's (1953) model and prove that conditions for instability of the monomorphic equilibria are also necessary and sufficient conditions for stability of the unique polymorphic equilibrium. This conclusion is valid only for the model with linear trade-off (solution of equation  $p_{i(t+1)} - p_{i(t)} = 0$  for  $p$  is polynomial of the third order), but nevertheless allows us to estimate the polymorphic equilibria even for a model with nonlinear trade-offs.

The analysis of instability of monomorphic equilibria gives us the transition points between monomorphic and polymorphic equilibria (points of neutral stability where leading eigenvalues are equal to zero). Then, the region of parameter space (in this case the range of niche proportions,  $c$ ) where these eigenvalues are higher than zero (i.e. are unstable) represent the parameter space where a polymorphic equilibrium exists. Thus, any perturbation in allele frequencies from the monomorphic equilibrium (and consequently in linkage disequilibrium) leads to a polymorphic equilibrium. With recombination between loci ( $r = 0.5$ ) the model maintains polymorphism if

$$\frac{s(1 - 2s - \epsilon)}{(1 - s)(2s + \epsilon)} < c < \frac{s + \epsilon}{(1 - s)(2s + \epsilon)}.$$

These conditions holds for recombination rates ( $r > \frac{-\epsilon}{1-s-\epsilon}$ ). If recombination rate is

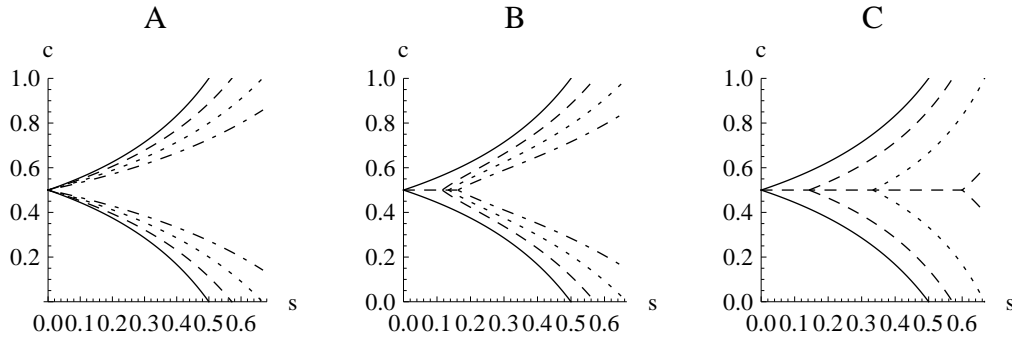
low ( $0 < r < \frac{-\epsilon}{1-s-\epsilon}$ ), the conditions above do not hold anymore and maintenance of polymorphism is determined by

$$\frac{(1-2s-\epsilon)(2s+\epsilon(1-r)+r(1-2s))}{(1-r)(4(1-s)s+2\epsilon(1-2s)-\epsilon^2)} < c < \frac{2s+\epsilon-r}{(1-r)(4(1-s)s+2\epsilon(1-2s)-\epsilon^2)}.$$

This transition is because the leading eigenvalues for the first case with faster recombination are overcome by the eigenvalues including recombination rate parameter. In the other words, these eigenvalues are the leading eigenvalues until the strength of selection overcome the recombination rate.

If selection is weak and more-than-additive epistasis is appreciable ( $\epsilon < 0$ ), the first condition (of each pair of conditions) becomes higher than the second one because the region of niche proportions, where the leading eigenvalues are higher than zero (i.e. monomorphic equilibria are unstable) do not overlap. Thence, both these conditions cannot be satisfied and a polymorphic equilibrium does not exist.

Fig. 6 shows how increasing convexity of the trade-off (i.e. increasing strength of more-than-additive epistasis) and increasing recombination rate narrow the range of niche proportions where polymorphism is maintained. Even very low recombination rate shifts the region of parameter space where polymorphism is maintained towards stronger selection (Fig. 6A and 6B). This shifting occurs only until the threshold for any recombination rate is reached ( $\frac{-\epsilon}{1-s-\epsilon}$  Fig. 6C). Then, any recombination rate leads to the same niche proportion ratio boundaries.



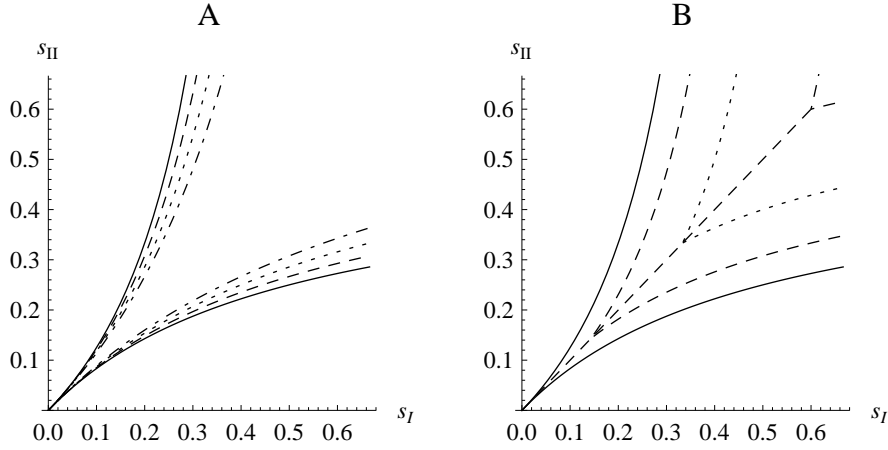
**Figure 6.** Range of niche proportions where polymorphism is maintained. A polymorphic equilibrium is achieved for the parameter combinations between the lines. On the x-axis is the strength of selection and on the y-axis are niche proportions. The outer solid curves represent linear trade-off, the dashed curves are for  $\epsilon = -0.25s$ , the dotted curves are for  $\epsilon = -0.5s$  and the dash-dotted curves for  $\epsilon = -0.75s$ . A) *o* recombination; B) recombination rate,  $r = 0.025$ ; C) recombination rate,  $r = 0.5$ .

### ■ 6.1.2. Asymmetry in strength of selection

Maintenance of polymorphism in the population does not depend only on the range of niche proportions, but is also affected by the difference in the strength of selection in niches. First, we set the niche sizes to be equal ( $c = 1/2$ ) and analyze under which combinations of  $s_I$  and  $s_{II}$  the population is polymorphic. Similarly as in the case with varying proportions of niches, the fraction of parameter space where the population achieves a polymorphic equilibrium depends on recombination rate,  $r$ , and the strength of more-than-additive epistasis, i.e. convexness of the trade-off between viabilities in niches. We analyzed the instability of the monomorphic equilibria and obtain conditions, where a polymorphic equilibrium exists. With recombination, the population achieves a polymorphic equilibrium if

$$\frac{s_I - \epsilon_{II}(1 + s_I)}{1 + 2s_I} < s_{II} < \frac{s_I + \epsilon_I}{1 - 2s_I - \epsilon_I}.$$

Because the system is symmetric, the same conditions with substituted subscripts holds also for  $s_I$ . The parameter space where the population achieves a polymorphic equilibrium is also reduced by increasing more-than-additive epistasis, i.e. increasing convexity of the trade-off. Figure 7 shows some examples of the parameter space for different values of the epistatic coefficients  $\epsilon_\gamma$ .



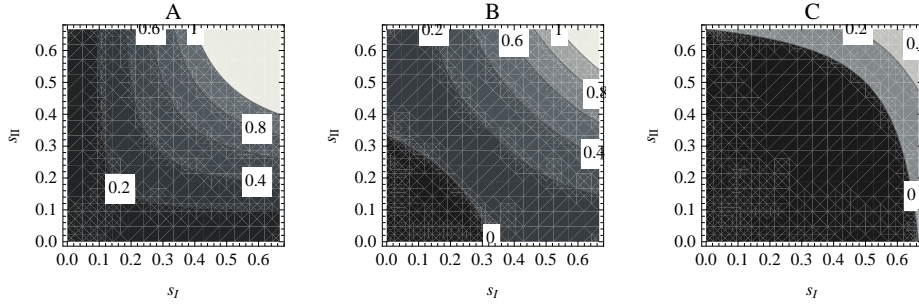
**Figure 7.** Parameter space of  $s_I$  and  $s_{II}$  where polymorphism is maintained in a population. Niche proportions,  $c$ , are held fixed at  $1/2$ . The outer solid curves represent linear trade-off, the dashed curves are for  $\epsilon_\gamma = -0.25 s_\gamma$ , the dotted curves are for  $\epsilon_\gamma = -0.5 s_\gamma$  and the dash-dotted curves for  $\epsilon_\gamma = -0.75 s_\gamma$ . A) No recombination; B) free recombination between loci.

### ■ 6.1.3. Asymmetry in niche proportions and strength of selection

Finally, we analyzed for which part of parameter space of  $c$ ,  $s_I$  and  $s_{II}$  the model is in polymorphic equilibrium. Similarly to the case with equal strength of selection, we analyzed instability of monomorphic equilibria, which gives us the conditions when at least one polymorphic equilibrium exists. With recombination  $r=1/2$ , the model reaches a polymorphic equilibrium if

$$\frac{1 - s_{II} \epsilon_I - 2 s_I s_{II} - s_{II}}{s_I + s_{II} + \epsilon_I - s_{II} \epsilon_I - 2 s_I s_{II}} < c < \frac{s_{II} + \epsilon_{II}}{s_I + s_{II} + \epsilon_{II} - \epsilon_{II} s_I - 2 s_I s_{II}}.$$

The actual parameter space between these conditions for various values of  $\epsilon_\gamma$  is depicted in Fig. 8 as an absolute value of the difference in niche sizes where the population is polymorphic,  $|\Delta c|$  (as in Gavrillets 2004). With increasing convexity of the trade-off, the parameter space with polymorphic equilibria shrinks significantly, and with  $\epsilon_\gamma \rightarrow -s_\gamma$ , selection has to be extremely strong to maintain polymorphism.



**Figure 8.** Contour plots of the parameter space, where polymorphism is maintained in two-locus model. Contour plots depict absolute value of difference in niche proportions,  $|\Delta c|$ , with respect to selection coefficient in niche I and niche II ( $s_I$  and  $s_{II}$  respectively) where polymorphic equilibria exist. A)  $\epsilon_\gamma = 0$ ; D)  $\epsilon_\gamma = -\frac{1}{3} s_\gamma$ ; c)  $\epsilon_\gamma = -\frac{2}{3} s_\gamma$ ;  $\gamma \in \{I, II\}$ .

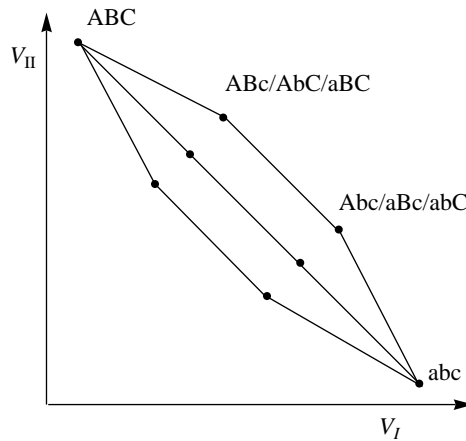
#### ■ 6.1.4. Three-locus model

Since we aim to discover the role of epistatic interaction in maintenance of polymorphism and possibility of divergence in a polygenic trait, we have extended the two-locus model to three loci. The number of combinations of each epistatic interaction is proportional to  $\binom{n}{m}$ , where  $n$  is number of loci and  $m$  is number of interacting loci. Thus, in the three-locus model we have three possible pairwise epistatic interactions ( $\epsilon_\gamma$ ) and one possible three allelic interaction ( $\epsilon_{123,\gamma}$ ) for each niche ( $\gamma$ ). Fitnesses of all genotypes are defined in Table 2.

genotype	niche I	niche II
ABC	$1 - 3 s_I - 3 \epsilon_I - \epsilon_{123,I}$	1
ABc/AbC/aBC	$1 - 2 s_I - \epsilon_I$	$1 - s_{II}$
Abc/aBc/abC	$1 - s_I$	$1 - 2 s_{II} - \epsilon_{II}$
abc	1	$1 - 3 s_{II} - 3 \epsilon_{II} - \epsilon_{123,II}$

**Table 2.** Fitnesses of individual genotypes.  $s_I$  is selection coefficient for deleterious alleles in niche I,  $s_{II}$  is selection coefficient for deleterious alleles in niche II and  $\epsilon_I$  and  $\epsilon_{II}$  are pairwise epistatic coefficients which are defined as the deviation from additivity of any two beneficial alleles present in a genotype together. Epistatic interaction of all three deleterious alleles is determined by the coefficient  $\epsilon_{123,I}$  and  $\epsilon_{123,II}$ , respectively.

As the number of loci increases, also the trade-off curves between viabilities in niches become smoother and are depicted in Figure 9.



**Figure 9.** Trade-offs between viabilities in niches. The lower convex trade-off ( $\epsilon_\gamma < 0$  and/or  $\epsilon_{123,\gamma} < 0$ ) results in an advantage of individuals carrying extreme genotypes *ABC* or *abc* relative to the intermediate genotypes. The upper concave trade-off ( $\epsilon_\gamma > 0$  and/or  $\epsilon_{123,\gamma} > 0$ ) favours intermediate genotypes and despite of the presence of two different environments a generalist predominate.

We have analyzed this model similarly as the two-locus model. The three-locus model with linear trade-off, i.e. without epistatic interactions and equal strength of per-locus selection in niches ( $s_I = s_{II} = s$ ) maintains polymorphism if

$$\frac{1 - 3s}{2 - 3s} < c < \frac{1}{2 - 3s}.$$

Analysis of the model with epistatic interactions faces similar difficulties like the two-locus model, so we have analyzed instability of monomorphic equilibria again and focus only on the case when recombination is fast enough ( $r \rightarrow 1/2$ ). The model with convex trade-off maintains polymorphism if

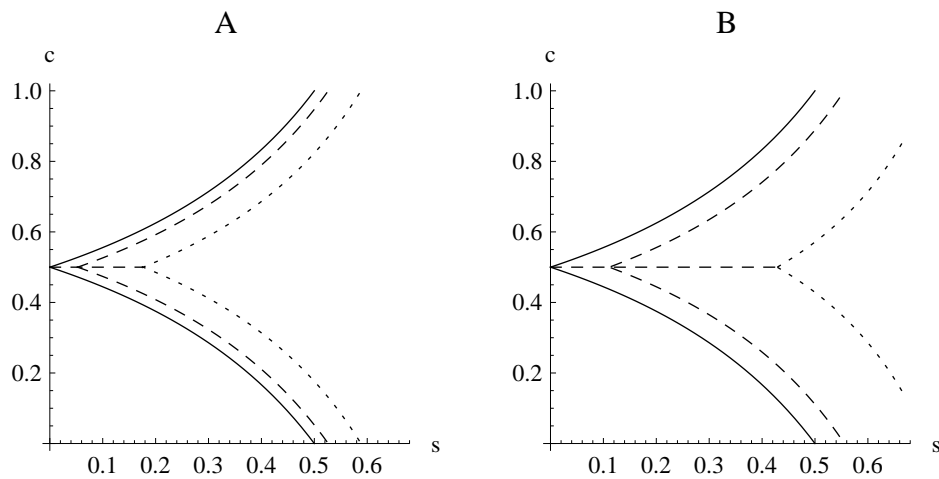
$$\frac{s - 3s^2 - 3\epsilon s}{(s + \epsilon)(2 - 3s)} < c < \frac{s + 2\epsilon}{(s + \epsilon)(2 - 3s)}.$$

Because the fitnesses are defined as a subtraction of particular selection coefficients of the deleterious alleles from 1, the overall strength of selection on the trait in the three-locus model with the same  $s$  is higher than in the two-locus model. Thus, if we want to understand what is the role of epistatic interactions in maintenance of genetic polymorphism, we must normalize the strength of selection on the trait (i.e. multiply the

per-locus strength of selection in the three-locus model by the factor of  $\frac{2}{3}$ , so  $\hat{s} = \frac{2}{3} s$ . Then, the conditions for the case with linear trade-off are equal for both models. For the three-locus model with convex trade-off we obtain conditions

$$\frac{\hat{s} - 2\hat{s}^2 - 3\epsilon\hat{s}}{(1-\hat{s})(2\hat{s}+3\epsilon)} < c < \frac{\hat{s} + 3\epsilon}{(1-\hat{s})(2\hat{s}+3\epsilon)}$$

Comparison with the conditions for the two-locus model shows us that an increasing number of pairwise epistatic interactions narrows the range of niche proportions where polymorphism is maintained (Fig. 10), because the relative advantage of being a specialist is higher and therefore the trade-off is more convex.



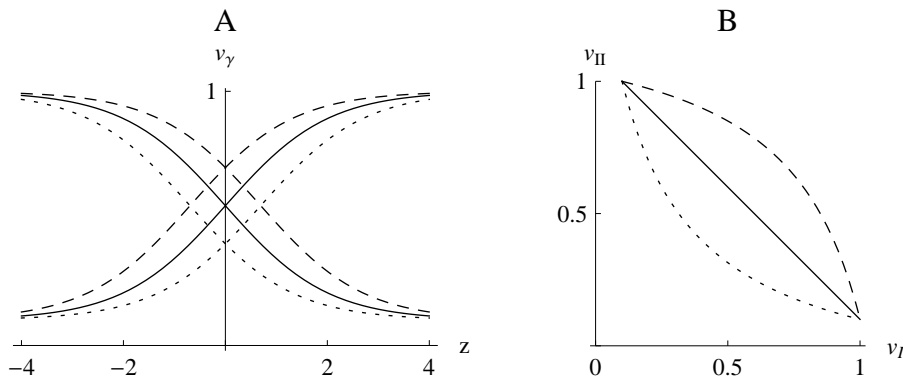
**Figure 10.** Comparison of the ranges of niche proportions where polymorphism is maintained in the two- and three-locus model with normalized strength of selection. Population is polymorphic in the parameter space between corresponding curves. Solid curves represent the case with linear trade-off, the dashed curves represent the two-locus model and the dotted curves the three-locus model. A)  $\epsilon = -0.1s$ ; B)  $\epsilon = -0.2s$ .

### ■ 6.1.5. Functional fitness

To compare these simple models with classic definitions of fitnesses with the infinitesimal version of Levene model, we redefine them as in Barton (2010). In this model, the viability of an individual with particular value of the trait  $z$  in niche I and niche II is defined as follows:

$$v_I = \theta + \frac{1-\theta}{1+\beta e^z} \quad \text{and} \quad v_{II} = \theta + \frac{1-\theta}{1+\beta e^{-z}}.$$

Parameter  $\beta$  determines how high is the relative advantage of being a specialist or a generalist and thence the curvature of the trade-off. If  $\beta = 1$ , the trade-off is linear, with  $\beta < 1$  the trade-off is concave and if  $\beta > 1$  the trade-off is convex (Fig. 11). The actual position of particular genotypes on the viability curves in Fig. 11A are determined by the effect of alleles on the trait,  $\alpha$ . Alleles beneficial in one niche contribute to the trait by  $-\alpha$  and alleles beneficial in the other niche by  $+\alpha$ . Then, the intermediate genotypes which contains equal number of alleles of the first and the second type have the value of the trait equal to 0 and the specialist genotypes have trait values equal to  $\pm n\alpha$ , where  $n$  is the number of loci. Minimal viabilities in the “wrong” niches are determined by  $\theta$ . Strength of selection then varies from 0 to 1, while if  $\theta = 1$ , there is no selection on the trait (viabilities are equal to 1 for any value of the trait in any niche).



**Figure 11.** A) Plot of viabilities in niches,  $v_\gamma$  with respect to the value of the trait,  $z$ . B) Trade-off in those viabilities. The solid curves represent  $\beta = 1$ , the dashed curves  $\beta = 0.5$  and the dotted curves  $\beta = 2$ . Minimal viability in the “wrong” niche is  $\theta = 0.1$ .

The actual viabilities of all genotypes for the two-locus model are defined as follows:

genotype	niche I	niche II
AB	$\theta + \frac{1-\theta}{1+\beta e^{2\alpha}}$	$\theta + \frac{1-\theta}{1+\beta e^{-2\alpha}}$
Ab/aB	$\theta + \frac{1-\theta}{1+\beta e^0}$	$\theta + \frac{1-\theta}{1+\beta e^0}$
ab	$\theta + \frac{1-\theta}{1+\beta e^{-2\alpha}}$	$\theta + \frac{1-\theta}{1+\beta e^{2\alpha}}$

Now, for given genotype frequencies, we can calculate the trait mean and vari-



ance. Trait values of genotypes for the two-locus model are  $z(X) = \{-2\alpha, 0, 0, 2\alpha\}$  and its mean value is

$$\bar{z} = \sum_{X=1}^4 g(X) z(X),$$

where  $g(X)$  is frequency of particular genotype  $X$ . Variance of the trait in population is given by

$$\text{Var}(z) = \sum_{X=1}^4 g(X) [z(X) - \bar{z}]^2 = \sum_{i=1}^2 2\alpha^2 p_i q_i.$$

The actual strength of selection is given both by parameters  $\beta$  and  $\theta$ , and by the variance of the trait in population. The variance of the trait is determined both by allele frequencies and by the allelic effect  $\alpha$ , whereas if all the allele frequencies are at 1/2, the variance is maximal and  $\alpha$  is the only parameter, which determines it. If the variance is high enough, the specialists lie on the viability curves close to their asymptotes, so the strength of selection against maladapted genotypes in each niche is maximal. Thus, we know that viabilities of the adapted genotypes are equal to 1 and viabilities of the maladapted genotypes are equal to  $\theta$ . For the linear trade-off ( $\beta = 1$ ), the actual strength of per locus selection is equal to the viability of the intermediate genotypes (i.e. intersection of the  $v_\gamma$  axis at  $z = 0$ ) and it is

$$s_\gamma = \frac{1 - \theta}{2}.$$

For the concave trade-off with  $\beta = 0.5$ , this intersection is at 2/3 of the relative difference in viabilities ( $1 - \theta$ ). Thus, relative viability of the intermediate genotypes is  $\frac{2+\theta}{3}$  and it is equal to  $1 - s_\gamma$ . Similarly, viabilities of the maladapted genotypes,  $\theta$ , are equal to their fitness according to the first definition,  $1 - 2s_\gamma - \epsilon_\gamma$ . If we solve these two equations for  $s_\gamma$  and  $\epsilon_\gamma$ , we obtain the solution, which is

$$s_\gamma = \frac{1 - \theta}{3} \text{ and } \epsilon_\gamma = \frac{1 - \theta}{3}.$$

For the convex trade-off with  $\beta = 2$ , the intersection is at  $1/3$  of the relative difference in viabilities,  $1 - \theta$ . Thence, the equations are

$$\frac{1 + 2\theta}{3} = 1 - s_\gamma \text{ and } \theta = 1 - 2s_\gamma - \epsilon_\gamma$$

and their solution gives us the corresponding strength of selection and epistasis,

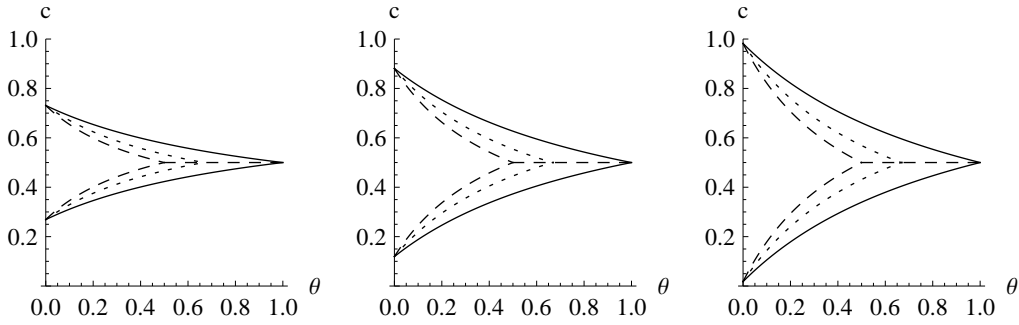
$$s_\gamma = \frac{2 - \theta}{3} \text{ and } \epsilon_\gamma = \frac{\theta - 1}{3} = -\frac{s_\gamma}{2}.$$

Now, we analyze maintenance of polymorphism in this model and from now on focus only on the case with convex trade-offs.

If variance in the trait is low, i.e. the allelic effect  $\alpha$  is small, the viabilities of the specialists lie on the viability curves closer to their intersection with  $v_\gamma$  axis (i.e. the viability of the generalist), so the relative strength of selection against the maladapted genotypes is lower. As we have inferred above, the niche size difference where polymorphism is maintained increases with increasing strength of selection. And in accordance with those results, also with these definitions of viabilities the niche size difference where polymorphism is maintained increases with decreasing  $\theta$  and increasing  $\alpha$  (Fig. 12). The conditions for maintenance of polymorphism are

$$\frac{1 + e^{2\alpha} \beta \theta}{(1 + e^{2\alpha})(1 + \beta \theta)} < c < \frac{e^{2\alpha} + \beta \theta}{(1 + e^{2\alpha})(1 + \beta \theta)}.$$

Similarly as with the first definitions of fitness, when recombination rate is very small, i.e.  $r < 0.1$ , the boundaries of the parameter spaces are slightly inflated.

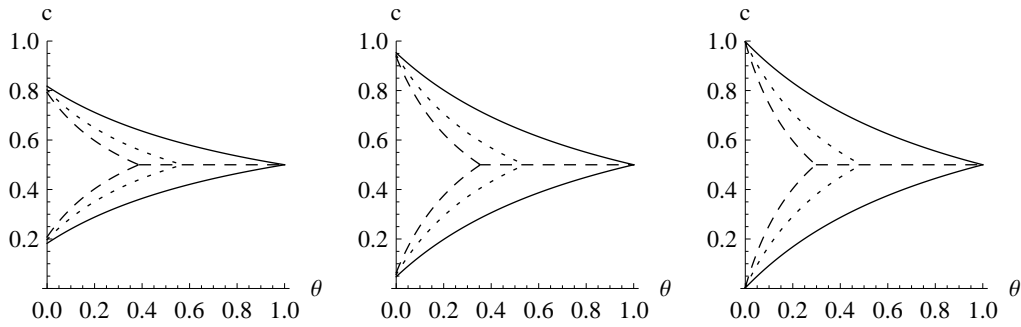


**Figure 12.** Fractions of parameter space where the population reaches a polymorphic equilibrium. The solid curves represent  $\beta = 1$ , dotted curves  $\beta = 1.5$  and the dashed curves  $\beta = 2$ . A)  $\alpha = 0.5$ , B)  $\alpha = 1$  and C)  $\alpha = 2$ .

As we inferred above, maintenance of polymorphism generally depends on the strength of selection and the relative advantage of being a specialist. In the first part, we concluded that with increasing number of loci and convex trade-off, the range of niche proportions where a polymorphic equilibrium can be achieved is lower because if the strength of selection on the trait is the same, more epistatic interactions are involved and the convexity of the trade-off is higher. Similarly, as with the functional definitions of viabilities, increasing number of loci narrows the range of niche proportions where population is polymorphic, if we normalize the maximal variance of the trait. In the three-locus model, the maximal variance is 1.5 times higher, so also the strength of selection is higher and thence also the niche proportions where the population is polymorphic is inflated. However, if we multiply the allelic effect  $\alpha$  in the three-locus model by a factor of  $2/3$ , for  $\theta = 0$  we obtain the same boundaries, where the population is polymorphic. If  $\theta$  increases, the the parameter space gets slightly narrower then in the two locus model.

The conditions for maintenance of polymorphism in the three-locus model are

$$\frac{(1 + e^\alpha \beta)(1 + e^{3\alpha} \beta \theta)}{1 + \beta^2 \theta + e^\alpha \beta (1 + \theta) + e^{3\alpha} \beta (1 + \theta) + e^{4\alpha} (1 + \beta^2 \theta)} < c < \frac{(e^\alpha + \beta)(e^{3\alpha} + \beta \theta)}{1 + \beta^2 \theta + e^\alpha \beta (1 + \theta) + e^{3\alpha} \beta (1 + \theta) + e^{4\alpha} (1 + \beta^2 \theta)}$$



**Figure 13.** Fractions of parameter space where the population reaches a polymorphic equilibrium in the three-locus model. The solid curves represent  $\beta = 1$ , dotted curves  $\beta = 1.5$  and the dashed curves  $\beta = 2$ . A)  $\alpha = 0.5$ , B)  $\alpha = 1$  and C)  $\alpha = 2$ .

Under the convex disruptive trade-off populations can maintain polymorphism, but the proportions of niches are even narrower than if the trade-off is linear, and are considerably shifted towards stronger selection. This parameter space also shrinks if the convexness of the trade-off between viabilities in niches is higher.

Since we assume recombination between loci ( $r = 1/2$ ), a relatively large fraction of the parameter space where the population is polymorphic exhibit only slightly bimodal distribution of genotypes. Frequencies of the intermediate genotypes goes to zero, if strength of selection and relative advantage of the specialists is high (i.e.  $\theta \rightarrow 0$ ,  $\beta > 1$  and  $\alpha$  high enough to reach it's asymptotes). Also, the parameter space with polymorphic equilibria under these values of the parameters is large. However, if some of these conditions are not satisfied and for example minimal viability in the “wrong” niche is appreciable (i.e.  $\theta > 0$ ), the parameter space where the population is polymorphic rapidly narrows (Fig. 13).

Even if the strength of selection is not high enough to ensure that the population will consist only of the extreme genotypes, the distribution of the genotypes is still considerably bimodal. So, some additional mechanism reducing recombination, such as assortative mating or habitat preference, might lead to complete reproductive isolation and thence speciation. Nevertheless, under such conditions where specialists have only small advantage over generalists, is hard to maintain polymorphism in the population and the niche proportions and strength of selection in niches have to be very close to symmetry.

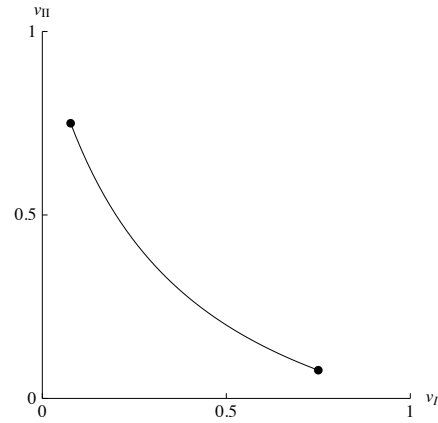
## **6.2. Summary of the results**

I first analyzed two and three-locus Levene model with epistatic interactions involved and assessed the parameter space in which polymorphism is maintained. The key conclusion from this part of the analysis is that the parameter space where polymorphism is maintained broadens with increasing strength of selection against maladapted genotypes within niches, but this region shifts towards stronger selection with increasing strength of more-than-additive epistasis among beneficial alleles. Thus, if selection against maladapted genotypes within niches is not strong enough, the population stays in a polymorphic equilibrium only if the system is completely symmetric. In the interpretation of trade-offs between viabilities in niches, with increasing convexity of the trade-off, polymorphism is maintained in the population only if selection is very strong. Also, an increasing number of loci and thence an increasing number of pairwise epistatic interactions among them, shifts the parameter space where the population is polymorphic towards stronger selection because the convexity of the trade-off increases.

Then, I redefined the definition of viabilities used by Barton (2010) in his infinitesimal model and analyzed the actual strength of selection and epistasis that corresponds to the functional definitions used in his model. Consistent with our previous results, I conclude that with increasing convexity of the trade-off, the parameter space where the population maintains polymorphism is shifting towards lower minimal viability in the “wrong” niche,  $\vartheta$  (Fig. 12). Similarly as above, increasing number of loci shifts the parameter space towards even lower minimal viability in the “wrong” niche (Fig. 13).

The next issue was to analyze the role of variance in the trait. I assign an allelic effect on the trait,  $\alpha$ , to each allele and analyzed its influence on the maintenance of polymorphism in population. Here, I conclude that the crucial point in the role of variance in the trait is whether the specialists lie on the viability functions (Fig. 11A) far enough that their viabilities in the wrong niches are close to zero. If the variance in the trait is low (i.e.  $\alpha$  or the number of loci is low) and the specialists lie on the viability functions closer to the generalists (i.e. they are closer to the trait mean) and thence have an appreciable viability in the wrong niche. Thus, the effect of low variance in the trait is in fact almost the same as the case when the minimal

viability in the wrong niche is appreciable ( $\vartheta > 0$ ). This results in cropped trade-off in the viabilities between niches (Fig. 14).



**Figure 14.** Trade-off between viabilities in niches with lower variance in the trait. If specialists do not have a high enough value of the trait, their viabilities in the wrong niche are higher than zero and the evolutionary outcome is similar to the case, when variance is high but minimal viability in the wrong niche is appreciable.

## 7. Discussion

Maintenance of polymorphism in populations is one of the most persistent questions in evolutionary biology and is related to many evolutionary phenomena. Since many traits involved in ecological adaptation and perhaps ecologically driven speciation are polygenic, I am trying in particular to answer the question: How is polymorphism in these polygenic traits maintained under disruptive natural selection in sexually reproducing populations? This issue is important because disruptive selection, which can drive sympatric speciation and therefore act on one panmictic population, also strongly affects maintenance of polymorphism in this population.

I conclude that the traditional scenario of sympatric speciation when a population experiencing weak or intermediate disruptive selection (lets say up to 10%) evolves some mechanism causing reproductive isolation (such as habitat preference or assortative mating) becomes highly implausible, because the system has to be extremely symmetric in strength of selection and niche proportions to maintain polymorphism (Fig. 12, 13). Polymorphism is maintained if viabilities of specialists in the wrong niches are close to zero (as in Barton 2010, Fig. 4a). As to my knowledge, there is not such an example of a polygenic trait (with recombining loci), where one value of the trait is lethal in one environment and the other value in the other environment. This case seems to be very implausible in nature. Though sympatric species with different hosts could be a good example, the adaptations tend to involve inversions and/or arise in allopatry (Feder et al. 2003a, 2003b).

I conclude that ecological sympatric speciation based on a polygenic trait and the life cycle assumed in this model is only likely if some other mechanisms of balancing selection are present. Because I analyzed only the haploid version of the model, for example heterozygote advantage, which acts as a mechanism maintaining polymorphism might play an important role and has been neglected in this study. Further research is needed to investigate whether polymorphism can also be maintained under weak or intermediate selection and thence lead to sympatric divergence.

Even though I analyzed only two and three-locus model, the comparison with the results of Barton (2010) lets us to conclude that maintenance of polymorphism is very hard unless viabilities of the specialists in the wrong niches are close to zero. In terms of selection coefficients, polymorphism is maintained only if selection against

maladapted genotypes is extremely strong. Whereas increasing the linear component of the strength of selection for specialists facilitates maintenance of polymorphism between niches, with increasing convexity of the trade-off, (i.e. increasing “disruptivness” of selection), maintenance of polymorphism gets harder (Fig. 6C, 12 and 13). Increasing number of loci also seems to reduce the possibility of maintenance of polymorphism, but this might be caused by increased number of epistatic interactions. The number of epistatic interactions rises with the number of loci involved, which may increase the convexity of the trade-off and hence reduce the parameter range where polymorphism is maintained.

Thus, we have to be careful in the interpretation of results obtained using the infinitesimal model in sympatric speciation modeling, because the assumption of fixed allelic variance can be misleading.

I have mentioned in the introduction the study by Bolnick and Fitzpatrick (2007), who revealed some properties of sympatric speciation and showed that these properties (frequency-dependent disruptive selection, assortative mating/ecology pleiotropy and whether assortative mating imposes weak or no costs) can be found in nature. However, they claimed that many arguable points remain unresolved. I have partly analyzed one of the remaining properties, maintenance of polymorphism under non-symmetric conditions of the niche proportions (e.g., amount of a limiting resource) and the strengths of selection within niches. Our results suggest that maintenance of polymorphisms and hence the coexistence of the incipient species may be rather hard to achieve. Increasing convexity of the trade-off of viabilities between niches makes maintenance of polymorphism even harder and requires extremely strong selection in the wrong niches.

To figure out possibilities of divergence in sympatry a further research covering additional mechanisms maintaining polymorphism, such as diploidy and heterozygote advantage, is necessary.



## **8. Conclusion**

The results inferred in this study suggest that sympatric ecological divergence through gradual adaptation in a polygenic trait is unlikely. Increasing strength of disruptive ecological selection (i.e. increasing convexity of the trade-off between viabilities in niches) causes that maintenance of polymorphism in the population requires extremely strong selection in the wrong niches, or precise symmetry in niche sizes and strength of selection in the niches. The traditional scenario of sympatric speciation when a population is under weak or intermediate disruptive selection and then evolves some mechanism of reproductive isolation is unlikely. Under these conditions polymorphism is lost unless stringent symmetry is satisfied, so some other mechanism maintaining polymorphism is needed to allow this speciation scenario. Thus, we have to be careful in the interpretation of results obtained using infinitesimal model, which assumes fixed allelic variance.

## 9. References

- Angert A. L., Crozier L. G., Gilman S. E., Rissler L. J., Tewksbury J. J. and Chunco A. J. 2011: Do species' traits predict recent shifts at expanding range edges? *Ecology Letters*, in press.
- Barton N. H. 2010: What role does natural selection play in speciation? *Phil. Trans. R. Soc. B* 2010 365, 1825-1840.
- Bolnick D. I. and Fitzpatrick B. M. 2007: *Annu. Rev. Ecol. Evol. Syst.* 38:459–87.
- Bulmer M. G. 1980: *The mathematical theory of quantitative genetics*. Clarendon Press, Oxford, U.K.
- Bürger R. 2009: Polymorphism in the two-locus Levene model with nonepistatic directional selection. *Theor. Popul. Biol.* 76, 214–228.
- Bürger R. 2010: Evolution and polymorphism in the multilocus Levene model without epistasis. *Theor. Popul. Biol.* 78, 123-138.
- Carson H. L. 1968: The population flush and its genetic consequences. In: *Population Biology and Evolution* (Lewontin, R. C., ed.), pp. 123–137. Syracuse: Syracuse University Press, NY.
- Cheverud J. M. and Routman E. J. 1995: Epistasis and its contribution to genetic variance components. *Genetics* 139: 1455-1461.
- Cheverud J. M. and Routman E. J. 1996: Epistasis as a source of increased additive genetic variance at population bottlenecks. *Evolution* 50(3): 1042-1051.
- Chou H. H., Chiu H.C., Delaney N. F., Segrè D., Marx C. J. 2011: Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science* 332, 1190-2.
- Coyne J. A. and Orr H. A. 2004: *Speciation*. Sinauer Associates, Inc. Massachusetts, USA; 545 p.

- Darwin C. 1859: *On the origin of species by means of natural selection*. London, UK: John Murray
- Dobzhansky T. 1937: *Genetics and the origin of species*. New York, NY: Columbia University Press, 1982; 364 p.
- Eshed Y. and Zamir D. 1996: Less-than-additive epistatic interactions of quantitative trait loci in tomato. *Genetics* 143, 1807–1817.
- Feder J.L., Berlocher S.H., Roethele J.B., Dambroski H., Smith J.J., Perry W.L., Gavrilovic V., Filchak K.E., Rull J. and Aluja M. (2003a): Allopatric genetic origins for sympatric host plant shifts and race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences U.S.A.* 100: 10314-10319.
- Feder, J.L., Roethele S.B., Filchak K.E., Niedbalski J. and Romero-Severson J. 2003b: Evidence for inversions related to sympatric host are formation in the apple maggotfly. *Genetics* 163: 939-953.
- Fisher R. A. 1918: The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinb.*, 52: 399–433.
- Fisher R. A. 1930: *The Genetical Theory of Natural Selection*. Oxford Univ. Press, 1999; 318 p.
- Fitzpatrick B. M. , Fordyce J. A. , Gavrillets S. 2008: What, if anything, is sympatric speciation?. *J Evol Biol.* 21(6):1452-9.
- Gavrillets S. 2004: *Fitness landscapes and the origin of species*. Princeton, NJ: Princeton University Press.
- Gliddon C. and Strobeck C. 1975: Necessary and sufficient conditions for multiple-niche polymorphism in haploids. *Am. Nat.* 109(966): 233-235.
- Goodnight C. J. 1987: On the effect of founder events on the epistatic genetic variance. *Evolution* 41: 80-91.
- Goodnight C. J. 1988: Epistasis and the effect of founder events on the additive

genetic variance. *Evolution* 42: 441-454.

Jiggins C. D. 2006: Sympatric speciation: why the controversy. *Current Biology* Vol 16 No 9, 333-334.

Khan A. I., Dinh D. M., Schneider D., Lenski R. E., Cooper T. F. 2011: Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332, 1193-6.

Kirkpatrick M. and Ravigné V. 2002: Speciation by natural and sexual selection: Models and experiments. *Am. Nat.* Vol. 159, pp. S22–S35.

Levene H. 1953: Genetic equilibrium when more than one niche is available. *Am. Nat.* 87, 331–333.

López-Fanjul C., Fernández A., Toro M. A. 2002: The effect of epistasis on the excess of the additive and nonadditive variances after population bottlenecks. *Evolution* 56(5):865-76.

López-Fanjul C., Fernández A., Toro M. A. 2004: Epistasis and the temporal change in the additive variance-covariance matrix induced by drift. *Evolution* 58(8):1655-63.

Matos M., Rose M. R., Rocha Pitea M. T., Rego C. and Avelará T. 2000: Adaptation to the laboratory environment in *Drosophila subobscura*. *J. Evol. Biol.* 13: 9-19.

Maynard Smith J. & Hoekstra R. 1980: Polymorphism in a varied environment—how robust are the models? *Genet. Res.* 35. 45–57.

Maynard Smith J. 1966: Sympatric speciation. *Am. Nat.* 100, 637–650.

Mayr E. 1954: Change of genetic environment and evolution. In *Evolution as a Process*, ed. J. S. Huxley, A. C. Hardy, E. B. Ford, pp. 156-80. London: Allen & Unwin.

Mayr E. 1963 *Animal species and evolution*. Cambridge, MA: Harvard University Press.

- Moore J. H. and Williams S. M. 2005: Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. *BioEssays* 27:637–646.
- Muller H. J. 1942: Isolating mechanisms, evolution and temperature. *Biol. Symp.* 6, 71-125.
- Nagylaki T. 2009: Evolution under the multilocus Levene model without epistasis. *Theor. Popul. Biol.* 76, 197–213.
- Poulton E. B. 1903: What is a species? *Proc. Entomol. Soc. Lond.* 1903, xxvii–cxvi.
- Prout T. 1968: Sufficient conditions for multiple niche polymorphism. *Am. Nat.* 102(928): 493-496
- Stearns S. C. 1989: Trade-offs in life-history evolution. *Functional Ecology*, 3(3): 259-268.
- Strobeck C. 1979: Haploid selection with  $n$  alleles in  $m$  niches, *Am. Nat.* 113(3): 439-444.
- Templeton A. R. 1980: The theory of speciation via the founder principle. *Genetics* 94, 1011–1038.
- Turelli M. and Barton N. H. 2006: Will multilocus epistasis and population bottlenecks increase additive genetic variance? *Evolution* 60:1763-1776.
- Turelli M. and Orr H. A. 2000: Dominance, epistasis and the genetics of postzygotic isolation. *Genetics* 154: 1663–1679.
- Via S. 2001: Sympatric speciation in animals: the ugly duckling grows up. *TREE* Vol. 16 No. 7: 281-390.
- Weber K., Eisman R., Morey L., Patty A., Sparks J., Tausek M. and Zeng Z.B. 1999: An analysis of polygenes affecting wing shape on chromosome 3 in *Drosophila Melanogaster*. *Genetics* 153: 773–786.
- Wilson D. S. and Turelli M. 1986: Stable underdominance and the evolutionary

invasion of empty niches. *Am Nat.* 127(6): 835-850.

Wright S. 1931: Evolution in mendelian populations. *Genetics* 16 (2): 97–159.

Wright S. 1932: The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc. Sixth Int. Congress in Genetics, Ithaca, NY, 1932, vol. 1, pp. 356–366.* Menasha, WI: Brooklyn Botanic Garden.

## **10. Source code**

Source code of the analyses of the models can be found on the attached CD.