**Evolutionary consequences of assortativeness in haploid genotypes**

David Marcelo Schneider
Instituto de Física 'Gleb Wataghin',
Universidade Estadual de Campinas, 13083-970, Campinas, SP, Brazil

Ayana de Brito Martins
Instituto de Biociências, Universidade de São Paulo,
05508-090, São Paulo, SP, Brazil

Eduardo do Carmo
Universidade Federal da Integração Latino Americana
85867-970, Foz do Iguaçu, PR, Brazil

Marcus Aloizio Martinez de Aguiar (corresponding author)
Instituto de Física 'Gleb Wataghin',
Universidade Estadual de Campinas, 13083-970, Campinas, SP, Brazil

**Abstract**

We study the evolution of allele frequencies in a large population where random mating is violated in a particular way that is related to recent works on speciation. Specifically, we consider non-random encounters in the haploid phase, which is relevant for organisms that release gametes in the environment so that fertilization is external. The gametes are described by biallelic genes at two loci and pairs of gametes whose alleles differ at both loci are considered incompatible. Evolution under these conditions lead to the complete disappearance of one of the alleles and substantially reduce the diversity of the population. Surprisingly, certain combinations of allele frequencies remain constant during the evolution, revealing the emergence of strong correlation between the two loci promoted by the epistatic mechanism of incompatibility. The allele frequencies at equilibrium depend only on their initial values, and so does the time to equilibration. We discuss the relation between this model of assortative fertilization and selection against double heterozygous individuals and its relevance to speciation in spatially extended populations.

# 1 Introduction

While the origin of species has always been a central subject in evolutionary biology, the large number of recent empirical and theoretical developments has renewed the interest in the area (Coyne and Orr, 2004; Butlin et al., 2012; Nosil, 2012). Individual-based simulations, in particular, have been successful in fostering relevant discussions in speciation (Dieckmann and Doebeli, 1999; Doorn et al., 2009; de Aguiar et al., 2009). Specifically, simulations in which mating is restricted by spatial and genetic distances have been able to describe empirical patterns of species diversity (de Aguiar et al., 2009) and within-species genetical diversity (Martins et al., 2013).

In order to reflect the dynamics of evolving populations, most simulations incorporate several ingredients simultaneously, such as mutation, genetic drift, recombination, assortativeness in mating and individual's movement and spatial positioning. Gavrilets (1999) proposed and analysed a number of simplified mathematical models that are closely related to these simulations, including selection, mutation, drift and population structure. These complex approaches to speciation, although somewhat realistic, do not allow for the detailed understanding of how each of the mechanisms involved contribute to the emergence and maintenance of reproductive isolation.

The main purpose of the present work is to analyse, from a more theoretical point of view, the evolutionary consequences of restricting mating by genetic distance.

One of the simplest ways of introducing assortativeness in mating in a individual-based simulation is to attribute haploid genomes with $B$ biallelic loci to individuals and allow them to mate only if the genomes differ in no more than $G$ loci (Gavrilets et al., 2000; de Aguiar et al., 2009; Martins et al., 2013). This approach considers that mate choice often relies on multiple cues that are determined genetically (Candolin, 2003). In the case of assortative mating, we assume that individuals have a certain tolerance to differences when choosing a mate, however if the other individual is too different, it will no longer be considered a potential mate. Under these assumptions, reproductive isolation was shown to be maintained among demes in the presence of sufficiently low migration rates (Gavrilets et al., 2000) and to emerge spontaneously from a genetically uniform population subjected to mutations in the case of spatially extended populations for which mating is also constrained by the spatial distance. (de Aguiar et al., 2009; Martins et al., 2013).

The introduction of mating incompatibility for individuals differing in too

many loci can be viewed as a breakdown of the panmixia. Instead of being allowed to mate with any other individual in the population, individuals can only mate with others that are sufficiently similar genetically. Panmixia is one of the central hypothesis of the Hardy-Weinberg (HW) equilibrium law (Hardy, 1908; Weinberg, 1908), which establishes that in very large and well mixed populations where mutations are negligible, the allele frequencies remain constant over generations and genotype frequencies remain constant after the first generation. The result, known today as the Hardy-Weinberg (HW) equilibrium law, set a null hypothesis for evolution and established the mathematical grounds to understand the forces that cause it to break down. Changes in allele or genotype frequencies are related to the breaking of one or more HW hypothesis and can be regarded as evolutionary forces acting on the population. Speciation is a particular case of evolution where subpopulations follow different evolutionary paths. This process is associated with the emergence of reproductive barriers that reduce gene flow and allow differentiation by selection and/or drift. These barriers may pre-copulatory, when individuals fail to copulate or post-copulatory, where copulation happens but fails to produce fertil descendants. Post-copulatory barriers may be pre-zigotic, when individuals copulate but fertilization does not occur, or post-zigotic, when the zigote is formed but has reduced viability or fertility (Coyne and Orr, 2004).

In this work we study the breakdown of the HW law by violating random fertilization in the way proposed by Gavrilets et al. (2000) and successfully applied to speciation by de Aguiar et al. (2009) and Martins et al. (2013). Even though individuals were treated as haploid in these simulations, pairing of two genomes occurred during mating, characterizing a diploid phase. Therefore, the non-random encounter of haploid individuals in these models is equivalent to non-random encounter of gametes in a diploid model. We hope our results will help clarify the processes described there and draw attention to the evolutionary consequences of non-random encounters in the haploid phase which differ from non-random encounters in the diploid phase as described in several classic models (see Crow and Kimura (1970) for a review).

Considering a very simple conceptual model where an organism's life cycle is divided in stages (Fig. 1), it is possible to map different reproductive barriers to specific stages of this model in an attempt to unify the study of reproductive isolation to the violation of assumptions of the HW law.

For the HW theorem to hold mating and fertilization must be random

(Figs. 1(A) and 1(C)), meiosis must follow Mendelian segregation (Fig.1(B)) and survival must be independent of genotype (Fig.1(D)). There are innumerous mechanisms by which each stage may depart from the assumptions of the HW's law. Non-random mating that can drive speciation may be caused by physical separation by geographical barriers (Mayr, 2001) or ecological (*e.g.* habitat, temporal), behavioral or mechanical isolation (Coyne and Orr, 2004). These mechanisms affect the encounter rate of diploid individuals (Fig. 1(A)). Hybrid inviability or sterility (Coyne and Orr, 2004), and strong competition for resources (Dieckmann and Doebeli, 1999) (Fig. 1(D)), are also mechanisms that can drive speciation. Segregation distorters (Fig. 1(B)) are known to cause at least partial reproductive isolation, however their role in speciation is less certain (Hurst and Schilthuizen, 1998; Schluter, 2009).

In this paper, we turn our attention to non-random fertilization (*i.e.* non-random encounters in the haploid phase, Fig. 1(C)). The role of gametic incompatibilities in reproductive isolation is particularly important in organisms for which gametes are released in the environment. If both males and females release gametes, so that fertilization is external (*i.e.* broadcast spawning) (Levitan, 1998), gamete recognition may be the only barrier to interspecific fertilization and, therefore, the only mechanism driving speciation (Eady, 2001). In addition, gametic isolation as an additional barrier to interspecific fertilization appears to be ubiquitious in many taxa and is expected to be historically important in speciation (Coyne and Orr, 2004). While our work may have direct importance in the case of external fertilizers, our main goal is to take a minimalist approach to understand the consequences of non-random gamete encounters.

We will work out the theory for infinitely large populations focusing on two biallelic loci ($B = 2$) without mutations. Gametic incompatibilities will be implemented by allowing gametic recognition only if the alleles form each gamete differ at most in one locus ($G = 1$). This is the simplest system for which the genetical mechanism of interest may be implemented. Contrary to the consequences of non-random encounters in the diploid phase (Ewens, 2004), we will show that this process does lead to evolution by changing the allele frequencies and that it is one of the main ingredients in the process of speciation studied in (de Aguiar et al., 2009). Despite the changes in all allele frequencies, we will demonstrate that a certain combination of frequencies from the two loci remain constant during the evolution, revealing a strong correlation between the loci introduced by the genetic restriction during fer-

tilization. Moreover, we will show that the evolution of such a population is equivalent to a population of diploid individuals with selection against double heterozygous genotypes and will connect the result with previous models that study the breakup of the Hardy-Weinberg theorem, selection in two-loci systems and speciation.

# 2 Statement of the problem

We consider a very large population of diploid individuals with two biallelic loci. Each individual produces gametes by combining the chromosomes inherited by its parents. The alleles at the first locus are $A$ or $a$ and the alleles at the second locus are $B$ or $b$.

Non-random encounters in our model occur in the haploid phase and it will be convenient to focus mainly on this phase, as in previous models of speciation (de Aguiar et al., 2009; Martins et al., 2013). Chromosomes $AB$ and $ab$ of a diploid individual, for example, produce gametes $Ab$, $AB$, $ab$ and $aB$ with 25% probability each. The gametes will later pair up with other gametes in the pool to form new diploid individuals and will crossover to produce the gametes of the next generation.

If all gametes are compatible there will be 9 different diploid genotypes whose frequencies can be calculated from the four independent gametic frequencies, as displayed in Table 1.

Table 1. Diploid frequencies for unrestricted fertilization.

| | $AB$ | $aB$ | $Ab$ | $ab$ |
|---|---|---|---|---|
| $AB$ | $f_{AB}^2$ | $f_{AB}f_{aB}$ | $f_{AB}f_{Ab}$ | $f_{AB}f_{ab}$ |
| $aB$ | $f_{aB}f_{AB}$ | $f_{aB}^2$ | $f_{aB}f_{Ab}$ | $f_{aB}f_{ab}$ |
| $Ab$ | $f_{Ab}f_{AB}$ | $f_{Ab}f_{aB}$ | $f_{Ab}^2$ | $f_{Ab}f_{ab}$ |
| $ab$ | $f_{ab}f_{AB}$ | $f_{ab}f_{aB}$ | $f_{ab}f_{Ab}$ | $f_{ab}^2$ |

In Gavrilets et al. (2000) de Aguiar et al. (2009) and Martins et al. (2013) the gametes were treated as *haploid individuals* and recombination as *mating* that produced the individuals (gametes) of the next generation. Moreover, individuals were considered incompatible if their alleles differed in too many loci. For the present case of two loci we forbid gametic recognition if both

alleles are different. This implies that the gametes $AB-ab$ and $Ab-aB$ will not pair up. The gametic frequencies in this case is shown in Table 2.

Table 2. Diploid frequencies for restricted fertilization.

| | $AB$ | $aB$ | $Ab$ | $ab$ |
|---|---|---|---|---|
| $AB$ | $f_{AB}^2$ | $f_{AB}f_{aB}$ | $f_{AB}f_{Ab}$ | 0 |
| $aB$ | $f_{aB}f_{AB}$ | $f_{aB}^2$ | 0 | $f_{aB}f_{ab}$ |
| $Ab$ | $f_{Ab}f_{AB}$ | 0 | $f_{Ab}^2$ | $f_{Ab}f_{ab}$ |
| $ab$ | 0 | $f_{ab}f_{aB}$ | $f_{ab}f_{aB}$ | $f_{ab}^2$ |

Interestingly, this type of mating restriction can be interpreted as selection against double heterozygous diploid individuals, as can be seen from Table 2. Since the sum of all entries must add to one, the haplotype frequencies in Table 2 are different from those in Table 1.

The purpose of this paper is to write down the evolution equations for the gametic frequencies in both situations, with and without mating restriction, and study the dynamics and the equilibrium solutions. We will show that the population follow the HW law in the latter case but display unexpected features in the former. In particular we will show that one of the alleles completely disappears from the population, changing considerably its original diversity. However, we will also show that a certain combination of alleles from each locus remain constant during the evolution, revealing a strong correlation between the two loci introduced by the dynamics.

# 3 Reproductive mechanism

Consider a population of $N$ biallelic gametes with haplotypes $AB$, $Ab$, $aB$, and $ab$ ($A$ and $a$ being the alleles at the locus 1, and $B$ and $b$ the alleles at the locus 2). The number of gametes of each type at time $t$ is given by $N_{AB}$, $N_{Ab}$, $N_{aB}$ and $N_{ab}$ with $\sum_{i,j} N_{ij} = N$, where $i = \{A, a\}$ and $j = \{B, b\}$. Pairs of gametes in this generation correspond to diploid individuals that will contribute gametes to the next generation through independent segregation. To each pair $g_1$ and $g_2$ there corresponds a fertilization probability $r_{g_1:g_2}$ that includes the effects of gametic compatibility and viability of the newly formed gametes.

Assuming no overlap of generations, the contribution of the current gametes to the gametes $AB$ of the next generation can be inferred from the Table 3.

Table 3. Diploid individuals generating $AB$ gametes.

| Diploid genotypes | Number of gamete encounters | Fraction of successful AB gametes |
|---|---|---|
| $AB \times AB$ | $\frac{1}{2}N_{AB} \times (N_{AB} - 1)$ | $r_{AB:AB}$ |
| $AB \times Ab$ | $N_{AB} \times N_{Ab}$ | $1/2 \times r_{AB:Ab}$ |
| $AB \times aB$ | $N_{AB} \times N_{aB}$ | $1/2 \times r_{AB:aB}$ |
| $AB \times ab$ | $N_{AB} \times N_{ab}$ | $1/4 \times r_{AB:ab}$ |
| $Ab \times aB$ | $N_{Ab} \times N_{aB}$ | $1/4 \times r_{Ab:aB}$ |

The number of gametes with haplotypes $AB$ at time $t+1$ thus obeys the equation

$$N_{AB}^{t+1} = \frac{N_{AB}^t(N_{AB}^t - 1)}{2} r_{AB:AB} + \frac{N_{AB}^t N_{Ab}^t}{2} r_{AB:Ab} + \frac{N_{AB}^t N_{aB}^t}{2} r_{AB:aB} + \frac{N_{AB}^t N_{ab}^t}{4} r_{AB:ab} + \frac{N_{Ab}^t N_{aB}^t}{4} r_{Ab:aB}. \quad (1)$$

By constructing equivalent tables, one obtains the equations for the number of gametes in the new generation for the remaining genotypes (see Appendix A). In the following sections we analyze, in the limit of infinitely large populations, the dynamics of the gametic frequencies $p_{ij} = N_{ij}/N$ and the corresponding allele frequencies $\tilde{p}_i = \sum_j p_{ij}$ and $\tilde{p}_j = \sum_i p_{ij}$. The two scenarios described above, with and without restriction in fertilization, will be specified by the values of the probabilities $r_{g_1:g_2}$.

# 4 Random fertilization

If fertilization is random we set $r_{g_1:g_2} = r$ for all pairs of gametes. We calculate $r$ imposing that the total number of gametes remains constant across generations. Substituting this constant fertilization probability in equation (1) and in equations (A.1)-(A.3) for the other types, summing their right hand sides and setting the result to $N$ we obtain

$$N = rN(N-1)/2 \quad (2)$$

so that $r = 2/N$ for very large populations. Accordingly, the dynamical equations for the evolution of the genotypic frequencies $p_{ij} = N_{ij}/N$, in the infinite size limit, can be written as

$$p_{AB}^{t+1} = p_{AB}^{t} - D^t/2 \tag{3}$$
$$p_{Ab}^{t+1} = p_{Ab}^{t} + D^t/2 \tag{4}$$
$$p_{aB}^{t+1} = p_{aB}^{t} + D^t/2 \tag{5}$$
$$p_{ab}^{t+1} = p_{ab}^{t} - D^t/2, \tag{6}$$

where we used the normalization condition $p_{AB} + p_{Ab} + p_{aB} + p_{ab} = 1$ to simplify the equations and defined the linkage (or gametic) disequilibrium (Hamilton, 2009)

$$D^t \equiv p_{AB}^t p_{ab}^t - p_{Ab}^t p_{aB}^t. \tag{7}$$

It can be shown (see SOM, section 1) that

$$D^t = 2^{-t} D^0, \tag{8}$$

where $D^0$ is the initial value of $D$.

From equations (3)-(6) one immediately sees that a sufficient condition for the equilibrium is $D = 0$, or $p_{AB}p_{ab} = p_{Ab}p_{aB}$. Notice also that the allele frequencies $\tilde{p}_A = p_{AB} + p_{Ab}$ and $\tilde{p}_B = p_{AB} + p_{aB}$ (and therefore $\tilde{p}_a$ and $\tilde{p}_b$) remain constant and equal to their initial values.

The dynamics defined by equations (3-6) can be solved analytically (see SOM, section 1). If $D^0 = 0$ the gametic frequencies also remain constant and are determined by the product of the corresponding allele frequencies, $p_{AB} = \tilde{p}_A\tilde{p}_B$, etc. If $D^0 \neq 0$ the population is in linkage disequilibrium and the gametic frequencies evolve toward the equilibrium where $p_{AB} = \tilde{p}_A\tilde{p}_B$, etc, but the frequencies reach these values only asymptotically as $D^t \to 0$, and not in a single step as in the one-locus HW equilibrium. These results are equivalent to previous known results of evolution towards gametic equilibrium.

# 5 Genetically restricted fertilization

To describe the incompatibility between gametes differing by more than one allele, we redefine the compatibility-viability fertilization rate as follows

$$r_{g_1:g_2} = \begin{cases} 0 & g_1:g_2 = AB:ab \text{ or } g_1:g_2 = Ab:aB \\ r' & \text{otherwise.} \end{cases} \tag{9}$$

To calculate the rate $r'$ we impose again that the total population size remains constant through generations. We obtain

$$r' = \frac{2}{N}\frac{1}{1-2\Delta} \tag{10}$$

where

$$\Delta \equiv \frac{N_{AB}N_{ab} + N_{Ab}N_{aB}}{N^2} = p_{AB}p_{ab} + p_{Ab}p_{aB} \tag{11}$$

Notice that $r'$ depends on the number of incompatible pairs in the population and is no longer constant. The larger the number of incompatible pairs the larger the chances that a compatible pair leaves an offspring to the next generation.

In this case the equations for the evolution of the gametic frequencies can be put in the form

$$p_{AB}^{t+1} = \frac{p_{AB}^t(1-p_{ab}^t)}{1-2\Delta^t} \tag{12}$$

$$p_{Ab}^{t+1} = \frac{p_{Ab}^t(1-p_{aB}^t)}{1-2\Delta^t} \tag{13}$$

$$p_{aB}^{t+1} = \frac{p_{aB}^t(1-p_{Ab}^t)}{1-2\Delta^t} \tag{14}$$

$$p_{ab}^{t+1} = \frac{p_{ab}^t(1-p_{AB}^t)}{1-2\Delta^t}. \tag{15}$$

In what follows, we explore the dynamics governed by these equations on the basis of stability analysis of the equilibrium solutions.

## 5.1 Equilibrium solutions and stability analysis

The equilibrium solutions are obtained by setting the frequencies at generation $t+1$ equal to the corresponding frequencies at generation $t$: $p_{AB}^{t+1} = p_{AB}^t$, $p_{Ab}^{t+1} = p_{Ab}^t$, etc. Equations (12)-(15) display four different types of equilibria which are grouped and labelled in Table 3. As we will show next, only types 1 and 2 are stable.

Since $p_{AB}+p_{Ab}+p_{aB}+p_{ab} = 1$, it is possible to give a graphical description of the dynamics by representing only three independent frequencies. We arbitrarily chose as the *phase space* frequencies $p_{AB}$, $p_{Ab}$ and $p_{aB}$, as depicted in Figure 2. The constrains $p_{AB} \geq 0$, $p_{Ab} \geq 0$, $p_{aB} \geq 0$, and $p_{AB}+p_{Ab}+p_{aB} \leq 1$ gives the phase space the geometry of a tetrahedron having right triangular faces.

Table 3. The four types of equilibrium points.

| Type | Label | Solution |
|---|---|---|
| Type 1. Continuous sets. Two compatible haplotypes have zero frequency; one allele is lost in one locus and the other locus remains polymorphic. | $E_a$ | $p_{AB} = p_{Ab} = 0, p_{aB} = \lambda_a, p_{ab} = 1 - \lambda_a; p_a = 1$ |
| | $E_b$ | $p_{AB} = p_{aB} = 0, p_{Ab} = \lambda_b, p_{ab} = 1 - \lambda_b; p_b = 1$ |
| | $E_A$ | $p_{aB} = p_{ab} = 0, p_{AB} = \lambda_A, p_{Ab} = 1 - \lambda_A; p_A = 1$ |
| | $E_B$ | $p_{Ab} = p_{ab} = 0, p_{AB} = \lambda_B, p_{aB} = 1 - \lambda_B; p_B = 1$ |
| | | where $\lambda_{a,b,A,B} \in (0, 1)$ |
| Type 2. Three haplotypes have zero frequency. One allele is lost in each loci. | $E_{ab}$ | $p_{AB} = p_{Ab} = p_{aB} = 0, p_{ab} = 1$ |
| | $E_{Ab}$ | $p_{AB} = p_{aB} = p_{ab} = 0, p_{Ab} = 1$ |
| | $E_{aB}$ | $p_{Ab} = p_{aB} = p_{ab} = 0, p_{AB} = 1$ |
| | $E_{ab}$ | $p_{AB} = p_{Ab} = p_{ab} = 0, p_{aB} = 1$ |
| Type 3. Two incompatible haplotypes have zero frequency | $EU_1$ | $p_{Ab} = p_{aB} = 0, p_{AB} = p_{ab} = 1/2$ |
| | $EU_2$ | $p_{AB} = p_{ab} = 0, p_{Ab} = p_{aB} = 1/2$ |
| Type 4. Equiprobable distribution | ES | $p_{AB} = p_{Ab} = p_{aB} = p_{ab} = 1/4$ |

The equilibrium points enumerated in Table 3 can be easily identified in figure 2. The continuous sets containing solutions Ea (purple), Eb (blue), EA (red) and EB (cyan) correspond to four of the six edges of the tetrahedron. The points Eab, EAb, EaB and EAB are the vertices of the tetrahedron (black circles), points $EU_1$ and $EU_2$ are located at the midpoints of the edges not containing equilibrium points (orange circles) and finally, the center of

the tetrahedron houses the point ES (brown circle).

The stability analysis of the equilibrium solutions of each of the four types is desbribed in the Appendices B and C and in section 2 of the SOM. The results are summarized in Table 4. The stable equilibria consist of types 1 and 2 and form a continuous set containing of the four edges of the tetrahedron displayed in Figure 2, including the vertices. Given an initial condition the population will converge to one of these equilibrium points, and we still need to find which one will be reached. This is the goal of the next two subsections.

Table 4. Stability of the equilibrium points.

| Type | Label in Figure 2 | Stability |
|---|---|---|
| Type 1. ($E_a, E_b, E_A, E_B$) | Colored lines (purple, blue, red, cyan) | Locally stable |
| Type 2. (Eab, EAb, EaB, Eab) | Black circles | Locally stable |
| Type 3. ($EU_1$, $EU_2$) | Orange circles | Unstable |
| Type 4. (ES) | Brown circle | Saddle point with a single stable direction |

## 5.2 Dynamics and Geometrical interpretation

Quantities that do not change in time give powerful insights in the understanding of dynamical problems. In the absence of restrictions in fertilization, the allele frequencies $\tilde{p}_A$ and $\tilde{p}_B$ remain constant and this property characterizes the gametic equilibrium. Surprisingly, the dynamics with restrictions in fertilization described here also has a conserved quantity that allows for the complete solution of the evolution equations.

In Appendix D we use equations (12)-(15) to show that all allele frequencies obey the same evolution equation, which can be written in the form

$$\tilde{p}_u^{t+1} - 1/2 = \frac{\tilde{p}_u^t - 1/2}{1 - 2\Delta^t} \tag{16}$$

for $u = A, B, a, b$. Writing this equation for $u = A$ and $u = B$ and dividing one by the other implies that the quantity

$$T = \frac{\tilde{p}_A - 1/2}{\tilde{p}_B - 1/2} \tag{17}$$

remains constant from the first generation. Therefore, the dynamic is constrained by two linear relations: constancy of $T$ and normalization ($p_{AB} + p_{Ab} + p_{aB} + p_{ab} = 1$). Given the two linear relations above, the trajectory of any initial condition is restricted to a two-dimensional plane in the four dimensional space of genotype frequencies. We call this plane where the dynamics take place the T-plane. The value of $T$ is fixed by the initial conditions and determines if the T-plane will intersect simultaneously the lines of fixed points $E_b$ and $E_B$ (as illustrated in Figure 3) or the lines $E_a$ and $E_A$.

This construction allows us to predict, for an arbitrary initial condition, the asymptotic equilibrium point of the population. First we need to specify in which plane the initial condition is located. Second, as the plane contains two stable fixed points, we also need to establish which of these equilibrium points will be reached. Since the dynamics is restricted to the T-plane and type 1 solutions are locally stable, it is possible to demonstrate (see Appendix E) that haplotype and allele frequencies will evolve toward one of these two points, losing one allele. The line passing through ES and connecting to $EU_1$ and $EU_2$ lies on the T plane and happens to be the subspace where ES is stable. Therefore, because ES is a saddle point, any trajectory that is not exactly on this line will be repelled towards one of the two available stable points contained in the T plane. Accordingly, knowing on which side of this line the initial point is suffices to determine which equilibrium point will be reached, as illustrated in Figure 3.

Let us assume that $-1 < T < 1$ so that the T-plane intersects the axis as in Figure 3. In this case points below the $EU_1$-$EU_2$ line move towards the $p_{Ab}$ axis, to a Eb fixed point, for which $b$ is fixed and $B$ is lost. Points above this line move towards the EB fixed point. Let us call $X$ the point where the T-plane crosses the $p_{Ab}$ axis. Using $p_{Ab} = X$ and $p_{AB} = p_{aB} = 0$ in equation

(17) we find $X = p_{Ab} = (1-T)/2$, which emphasizes that the T-plane crosses the Eb and EB lines only if $-1 < T < 1$. For $T > 1$ or $T < -1$ the T-plane crosses the Ea and EA lines.

We shown in Appendix E that, if $\tilde{p}_B < 1/2$ the point is the lower half of the T-plane and the equilibrium point is of the type Eb with $p_{Ab} = X = (1-T)/2$, $p_{AB} = p_{aB} = 0$, $p_{ab} = (1+T)/2$. Otherwise the point is on the upper half and goes to the EB fixed point given by $p_{Ab} = 0$, $p_{AB} = (1+T)/2$, $p_{aB} = (1-T)/2$, $p_{ab} = 0$. Also, from equation (17) we find that $\tilde{p}_B < \tilde{p}_A$, otherwise $T > 1$. Similarly, $\tilde{p}_B < \tilde{p}_a$, otherwise $T < -1$. The conclusion is that if $\tilde{p}_B$ is the smallest of the allele frequencies, the equilibrium point is $p_{Ab} = (1+T)/2$, $p_{AB} = p_{aB} = 0$, $p_{ab} = (1-T)/2$ which has $\tilde{p}_B = 0$.

The practical result of this analysis is that the smallest among the initial allelic frequencies always goes to zero. This information, together with the conserved quantity $T$ suffices to determine all frequencies. For example, if $p_b$ is the smallest initial frequency, in the equilibrium $p_b = 0$ and, consequently, $p_B = 1$. From equation (17) we find $p_A = (1+T)/2$ and $p_a = 1 - p_A = (1-T)/2$ and all genotype frequencies have been calculated. In addition, the time it takes to reach the equilibrium also depends on the initial condition and is calculated in detail in section 3 of SOM.

# 6 Discussion

Understanding how allele frequencies changes in a population is central to evolution. Depending on the conditions and constrains to which these changes are subjected, very different outcomes can result. Here we considered an isolated and large population where gametic recognition is controlled by alleles in two independent loci. We postulated that individuals whose genotypes differ in both loci (double heterozygotes) are inviable whereas all other individuals have the same fitness. The model was motivated by previous speciation studies where haploid individuals with $B$ biallelic genes differing in more than $G$ alleles could not mate because of accumulated differences. The problem considered here consists in the simplest case where $B = 2$ and $G = 1$ and we have shown that it is equivalent to a two loci diploid model with selection against double heterozygous individuals (*i.e.* underdominance for fitness). While this approach has been used in multiple computational models of speciation (Gavrilets et al., 1998; de Aguiar et al., 2009; Martins et al., 2013), to our knowledge, its connection to Hardy-Weinberg theory and

equivalence with selection against heterozygous individuals has never been explicitly stated. If we define $n = G+1$, restriction in mating between haploid individuals is equivalent to selection against $n$-uple to $B$-uple heterozygote. For example, a case where $B = 5$ and $G = 2$ is equivalent to selection against triple, quadruple and quintuple heterozygotes. Gavrilets (1999) while working on a equivalent model with multiple loci stated that this mechanism is *similar* to underdominance. By taking a step back and analysing a two-loci model, we have been able to conclude that this type of mating restriction in the haploid phase is actually *the same* as underdominance. This result highlights the importance of carefully evaluating assumptions and mechanisms used in more complex models.

Although selection against heterozygotes and non-random fertilization are completely equivalent under the assumptions taken into account in this work, the biological mechanisms involved are not expected to be the same. Gamete recognition in free-spawning species is the result of the interaction between multiple molecules on gamete surface that mediate the different stages of fertilization (Vieira and Miller, 2006; Lessios, 2011) and genes involved in this process are expressed specically in reproductive cells (Palumbi, 2008). Selection agains heterozygotes may arrise from intrisic genetic incompatibilities or other mechanisms that cause reduction in survival or mating probabilities of heterozygous individuals (Coyne and Orr, 2004) and the genes involved may have very different functions (Wu and Ting, 2004; Nosil and Schluter, 2011)

In particular, intrinsic genetic incompatibilities which are expected to be explained by specific epistatic interactions, are likely to be subject to a different mechanism. The evolutionary dynamics of these incompatibilities are usually studied by Dobzhansky-Muller models (Coyne and Orr, 2004). These models propose that incompatibilities arise from interactions between alleles that have evolved in different subpopulations. In that case the process involve an initial population of diploid individuals with genotype $AABB$, which has been divided in two subpopulations. Allele $a$ may evolve in one population and allele $b$ in the other and when individuals of different subpopulations breed, the hybrid offspring may have reduced fitness due to interactions between alleles $a$ and $b$. These assumptions lead to an important difference with respect to the model of selection against double heterozygotes, where all genotypes are equally fit, except for the double heterozygotes which are less fit. In the Dobzhansky-Muller model, besides the double heterozygotes, some individual which are heterozygous for just one locus and one type of

double homozygotes may also be less fit, depending on assumptions of ancestry and dominance. Therefore, the system here described does not present the asymmetry that is characteristic of the Dobzhansky-Muller mechanism. Despite these differences, our approach shares one important feature with the Dobzhansky-Muller model: in subdivided populations both mechanisms explain how diverging populations may arrive at opposite sides of an adaptive valley without ever crossing it. They evolve through ridges of equal fitness values (Gavrilets, 2004).

It has been known for decades that non-random encounters between genotypes during reproduction often lead to change in alleles frequencies in two-allele models. Nonetheless, the cases of inbreeding and pure positive assortative mating are often emphasized. In both cases the allele frequencies remain constant, although the frequencies of the genotypes do change through generations, and the constancy in allele frequencies is attributed to the fact that random encounters are disrupted in a away that is symmetrical to both alleles. (Hamilton, 2009; Crow and Kimura, 1970). In the present case, reproductive encounters between haploid genotypes are also restricted in a away that is symmetrical to both alleles in each loci, however this restriction is equivalent to selection against double heterozygotes, which leads to disruptive selection resulting in changes in allele frequencies. Moreover, one of the four alleles always disappears from the population. Note that this affects the diversity of the population in a drastic way: of the nine possible genotypes, six disappear, not only the double heterozygote.

Surprisingly we find that, despite theses drastic changes, certain combinations of allele frequencies remain constant during the entire process, even though alleles segregate independently. In order to understand the meaning of these constants, we define

$$d_u = \tilde{p}_u - 1/2 \tag{18}$$

for $u = A$, $a$, $B$ or $b$, as the *deviation* of the frequency of allele $u$ from value $1/2$, where $\tilde{p}_A = \tilde{p}_a = \tilde{p}_B = \tilde{p}_b$. In the phase space, the point where all alleles have equal frequencies corresponds to a saddle point unstable equilibrium.

In terms of $d_u$ the first conserved quantity is

$$T = \frac{d_A}{d_B} = \frac{d_a}{d_b}$$

which measures the ratio between the deviations of alleles $A$ and $B$. The

second conserved quantity is the sign of $d_u$, which also remain fixed for all alleles.

Using these two quantities we were able to compute all allele frequencies in equilibrium. The genotypic frequencies are then obtained by the corresponding product of the allele frequencies. Since the values of $T$ and the signs of $d_u$ depend on the initial conditions, so does the final allele and genotypic frequencies. In particular, we have shown that the allele that disappears from the population is the one that starts with the smallest frequency (smallest $d_u$). This model is closely related to the classical of models of selection in two biallelic loci propsed by Kimura (1956), Karlin (1975) and others (Ewens, 2004) in which arbitrary fitness values can be attributed to each of the nine possible diploid genotypes. These models, however, focus on different modes of overdominance in search of stable polymorphisms. By introducing underdominance under the same framework, we have find that the stable solutions are still polymorphic in one loci.

The constant T also reflects a correlation between the allele frequencies at the two loci that arise due to epistasis. This correlation explicitly shows how allele frequencies in one loci change as a function of the frequencies at the other loci and how a polymorphic stable equilibrium may be attained in a model of underdominance. For one-locus, two-alleles systems, the polymorphic equilibrium is always unstable (Hamilton, 2009). This result, which was also found in a previous model that combined overdominance and underdominance, draws attention to the importance of studying multiple-loci systems, since their dynamics may not be well represented by simpler systems (Hastings, 1981).

The consequences of this result to population diversity and speciation can be understood in the context of structured populations. For one-locus systems, selection against the heterozygote in subdivided population may lead to coexistence of both alleles in the whole population, if migration is low enough (Altrock et al., 2011). If more loci are taken into account, allele $a$ may be lost in one group and allele $b$ may be lost in another group. The diversity of each population would be reduced and the genetic distance between individuals of different populations would become larger than the distance between individuals within each population. For the case studied, $B = 2$ and $G = 1$, such populations would not be reproductively isolated: in terms of gametes, the first would have $AB$ and $Ab$ whereas the second would have $AB$ and $aB$ and gamete recognition between populations would still be possible. Joining the two populations would result in further evolution

and one of the alleles would be lost altogether. However, for larger number of participating genes, reproductive isolation could be achieved. For $B = 3$ and $G = 1$, for instance, the equilibrium solutions $ABc$ and $ABC$ would be isolated from the solution $abC$ and $abc$.

The model proposed in de Aguiar et al. (2009) breaks the HW hypothesis because the population is finite, includes mutations and is not panmictic because of spatial and genetic constrains. It shows that speciation can occur as a consequence of two reinforcing trends: isolation by spatial distance and isolation by genetic distance. Isolation by spatial distance is a consequence of limited dispersal across space and was originally studied by Sewall Wright (1943). Mating between individuals located more than a maximum distance $S$ from each other does not occur simply. If $S$ is much smaller than the size $L$ of population range, alleles become spatially autocorrelated and individuals become isolated by spatial distance. In addition, haploid individuals whose genetic distance exceeds a maximum $G$ are not allowed to mate. This is equivalent to assigning a lower fitness to individuals that are homozygous for more than $G$ loci. Initially, differentiation occurs solely by mutation and drift. However, as diversity increases the genetical constrain aids differentiation by disruptive selection. If $G$ is much smaller than the total number of genes $B$ involved in reproduction, mutations may lead to incompatible haploid genotypes. Mutations that are responsible for the differences between individuals that are spatially distant cannot spread through the population because selection against heterozygous individuals acts removing alleles that have lower local frequency. Analytical expressions indicating when speciation is possible under these conditions where recently obtained (de Aguiar and Bar-Yam, 2011; Baptestini et al., 2013), but the contribution of the genetic mating restriction to the process was included as ansatz based on numerical simulations. The main contribution of the present work is to understand in detail the core ingredient of these models, namely, the dynamics of allele frequencies, and to connect it to previous classic models in population genetics.

# References

de Aguiar, M. A. M., M. Baranger, E. M. Baptestini, L. Kaufman, and Y. Bar-Yam, 2009. Global patterns of speciation and diversity. Nature 460:384–387

de Aguiar M.A.M. and Y. Bar-Yam, 2011. The Moran model as a dynamical process on networks and its implications for neutral speciation. Phys. Rev. E 84:031901

Altrock, P. M., A. Traulsen, and F. A. Reed, 2011. Stability properties of underdominance in finite subdivided populations. PLoS Comput Biol 7:e1002260.

Baptestini, E. M., M.A.M. de Aguiar, Y. Bar-Yam, 2013 Conditions for neutral speciation via isolation by distance J. Theor. Biol. (in press).

Butlin, R., A. Debelle, C. Kerth, R. R. Snook, L. W. Beukeboom, R. F. C. Cajas, W. Diao, M. E. Maan, S. Paolucci, F. J. Weissing, L. van de Zande, A. Hoikkala, E. Geuverink, J. Jennings, M. Kankare, K. E. Knott, V. I. Tyukmaeva, C. Zoumadakis, M. G. Ritchie, D. Barker, E. Immonen, M. Kirkpatrick, M. Noor, C. Macias Garcia, T. Schmitt, and M. Schilthuizen, 2012. What do we need to know about speciation? Trends Ecol. Evol. 27:27–39.

Candolin, U., 2003. The use of multiple cues in mate choice. Biological Reviews 78:575595.

Coyne, J. A. and H. A. Orr, 2004. Speciation. 1 ed. Sinauer Associates, Inc.

Crow, J. F. and M. Kimura, 1970. An Introduction to Population Genetics Theory. The Blackburn Press.

Dieckmann, U. and M. Doebeli, 1999. On the origin of species by sympatric speciation. Nature 400:354–357.

Doorn, G. S. v., P. Edelaar, and F. J. Weissing, 2009. On the origin of species by natural and sexual selection. Science 326:1704–1707.

Eady, P. E., 2001. Postcopulatory, prezygotic reproductive isolation. J Zool 253:4752.

Ewens, W. J., 2004. Mathematical Population Genetics: I. Theoretical Introduction. Springer.

Gavrilets, S., 1999. A dynamical theory of speciation on holey adaptive landscapes. The American Naturalist 154:1–22.

———, 2004. Fitness Landscapes and the Origin of Species (MPB-41). Princeton University Press.

Gavrilets, S., H. Li, and M. D. Vose, 1998. Rapid parapatric speciation on holey adaptive landscapes. Proc Biol Sci 265:1483–1489.

———, 2000. Patterns of parapatric speciation 54:11261134.

Hamilton, M., 2009. Population Genetics. 1 ed. Wiley-Blackwell.

Hardy, G. H., 1908. Mendelian proportions in a mixed population. Science 28:49–50.

Hastings, A., 1981. Marginal underdominance at a stable equilibrium. Proc Natl Acad Sci U S A 78:6558–6559.

Hurst, G. D. D. and M. Schilthuizen, 1998. Selfish genetic elements and speciation. Heredity 80:2–8.

Karlin, S., 1975. General two-locus selection models: Some objectives, results and interpretations. Theoretical Population Biology 7:364–398.

Kimura, M., 1956. A model of a genetic system which leads to closer linkage by natural selection. Evolution 10:278–287.

Lessios, H. A., 2011. Speciation genes in free-spawning marine invertebrates. Integr. Comp. Biol. 51:456–465.

Levitan, D., 1998. 6 - sperm limitation, gamete competition, and sexual selection in external fertilizers. Pp. 175–217, *in* T.R. Birkhead and A.P. Mller, eds. Sperm Competition and Sexual Selection. Academic Press, San Diego.

Martins, A. B., M. A. M. d. Aguiar, and Y. Bar-Yam, 2013. Evolution and stability of ring species. PNAS 110:5080-5084

Mayr, E., 2001. What Evolution Is. Basic Books.

Nosil, P., 2012. Ecological speciation. Oxford University Press, Oxford; New York.

Nosil, P. and D. Schluter, 2011. The genes underlying the process of speciation. Trends Ecol. Evol. (Amst.) 26:160–167.

Palumbi, S. R., 2008. Speciation and the evolution of gamete recognition genes: pattern and process. Heredity 102:66–76.

Schluter, D., 2009. Evidence for ecological speciation and its alternative. Science 323:737–741.

Vieira, A. and D. J. Miller, 2006. Gamete interaction: Is it species-specific? Mol Reprod Dev 73:14221429.

Weinberg, W., 1908. ber den nachweis der vererbung beim menschen. Jh Ver Vaterl Naturk Wurttemb 64:368–382.

Wright, S., 1943. Isolation by distance. Genetics 28:114–138.

Wu, C.-I. and C.-T. Ting, 2004. Genes and speciation. Nat Rev Genet 5:114–122.

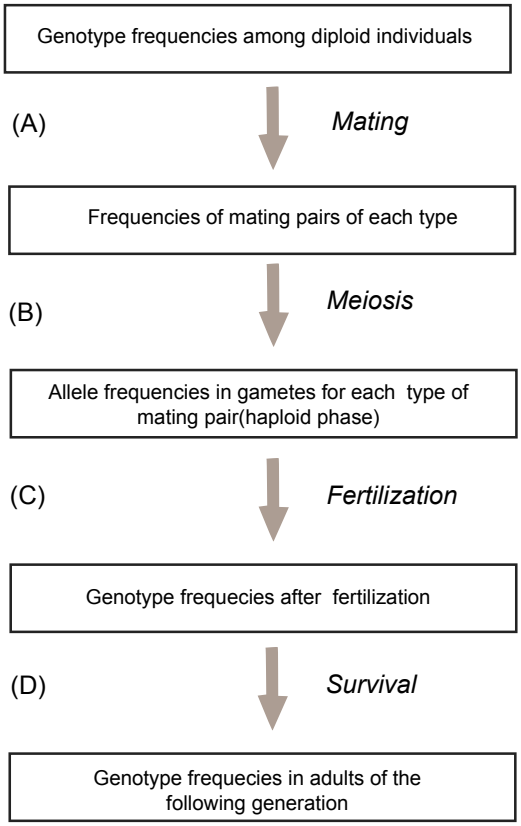


Figure 1: Conceptual model of an life cycle that includes a diploid and an haploid phase.

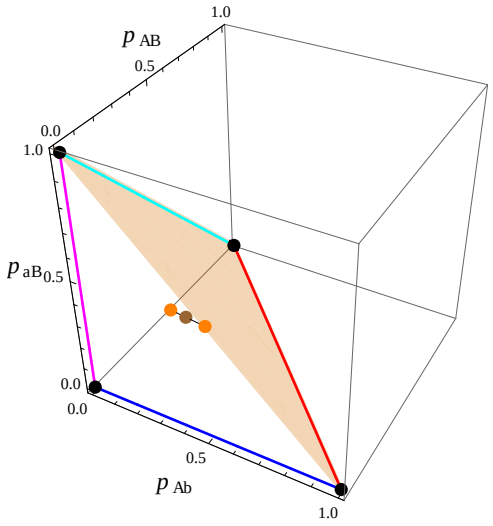


Figure 2: The 4 families of equilibrium solutions: $E_a$ (purple), $E_b$ (blue), $E_A$ (red) and $E_B$ (cyan); $E_{ab}$, $E_{Ab}$, $E_{aB}$ and $E_{AB}$ (black circles); $EU_1$ and $EU_2$ (orange circles) and; ES (brown circle). The light brown face corresponds to $p_{AB} + p_{Ab} + p_{aB} = 1$, or $p_{ab} = 0$, whereas the origin $p_{AB} = p_{Ab} = p_{aB} = 0$ implies is $p_{ab} = 1$.

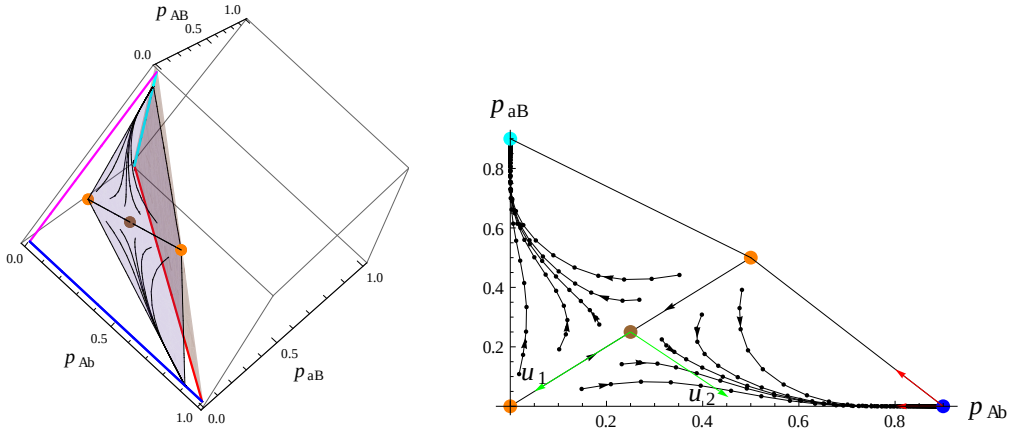


Figure 3: (a) T-plane $T = -0.8$ intersecting the $E_b$ fixed points at $p_{Ab} = 0.9$ and the $E_B$ points at $p_{AB} = 0.1$, $p_{aB} = 0.9$. (b) T-plane projected on the $p_{Ab}$-$p_{AB}$ plane showing the fixed points ES, $EU_1$ and $EU_2$. Vectors $\hat{u}_1$ and $\hat{u}_2$ are shown in green (see Appendix E). Sample trajectories are shown by dots, connected by straight lines to guide the eye.

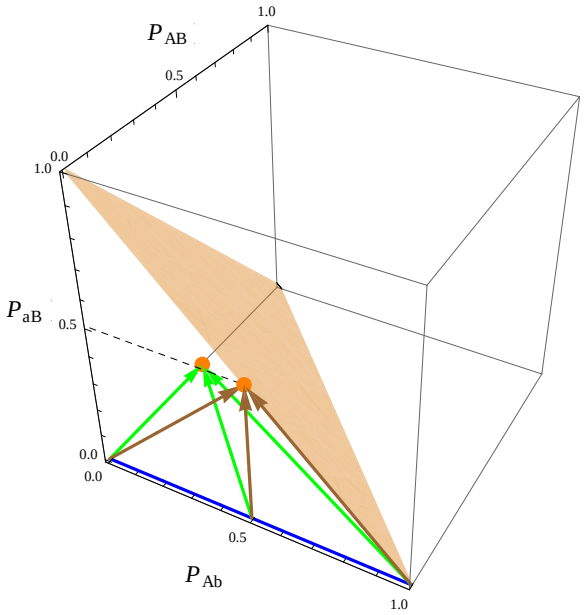


Figure 4: Eigenvectors of equilibrium points $E_b$ for $\lambda_b = 0$, $\lambda_b = 1/2$ and $\lambda_b = 1$. $\mathbf{v}_1$ (green) points to $EU_1$ in the $p_{Ab}-p_{AB}$ plane. $\mathbf{v}_2$ (brown) points to $EU_2$ in the $p_{Ab}-p_{aB}$ plane.

# A Master equations for haplotypes $Ab$, $aB$ and $ab$

$$N_{Ab}^{t+1} = \frac{N_{Ab}^t(N_{Ab}^t-1)}{2} r_{Ab:Ab} + \frac{N_{AB}^t N_{Ab}^t}{2} r_{AB:Ab} + \frac{N_{Ab}^t N_{ab}^t}{2} r_{Ab:ab} + \frac{N_{AB}^t N_{ab}^t}{4} r_{AB:ab} + \frac{N_{Ab}^t N_{aB}^t}{4} r_{Ab:aB} \tag{A.1}$$

$$N_{aB}^{t+1} = \frac{N_{aB}^t(N_{aB}^t-1)}{2} r_{aB:aB} + \frac{N_{AB}^t N_{aB}^t}{2} r_{AB:aB} + \frac{N_{aB}^t N_{ab}^t}{2} r_{aB:ab} + \frac{N_{AB}^t N_{ab}^t}{4} r_{AB:ab} + \frac{N_{Ab}^t N_{aB}^t}{4} r_{Ab:aB} \tag{A.2}$$

$$N_{ab}^{t+1} = \frac{N_{ab}^t(N_{ab}^t-1)}{2} r_{ab:ab} + \frac{N_{Ab}^t N_{ab}^t}{2} r_{Ab:ab} + \frac{N_{aB}^t N_{ab}^t}{2} r_{aB:ab} + \frac{N_{AB}^t N_{ab}^t}{4} r_{AB:ab} + \frac{N_{Ab}^t N_{aB}^t}{4} r_{Ab:aB} \tag{A.3}$$

# B Equilibrium solutions and stability analysis

The stability analysis of the equilibrium solutions consists in linearizing the equations (12)-(15) around the equilibrium point and computing the eigenvalues $\zeta$ of the $3 \times 3$ matrix governing the linear equations. Stability requires $|\zeta| < 1$ for all eigenvalues. In order to make their identification easier, we refer to the equilibirum points by their colors in Figure 2.

**Stability of points $EU_1$ and $EU_2$** (orange dots)

For these points the stability matrix has three equal eigenvalues $\zeta = 2$. As a consequence, both points $EU_1$ and $EU_2$ are unstable, whence the choice of labels EU.

**Stability of points ES** (brown dot)

The eigenvalues of the stability matrix of ES are $\zeta_s = 2/3$ and $\zeta_{u1} = \zeta_{u2} = 4/3$. Accordingly, this equilibrium is a saddle point (ES), being unstable in a two dimensional subspace and stable in a one dimensional subspace. The stable subspace is spanned by the eigenvector associated to $\zeta_s$, $(p_{AB}, p_{Ab}, p_{aB}) = (1, -1, -1)$. Notice that along this direction and equidistant to the point ES, are located the points $EU_1$ and $EU_2$ (figure 2). This fact is very important to the global structure of the phase space.

**Stability of points in types 1 and 2** (colored lines and black dots)

Equilibrium points of types 1 and 2 are certainly peculiar. Not displaying exactly the same properties, they share some common features, so it is instructive to analyze their stability at the same time. We take as an example the set of points Eb (blue line) and its $\lambda_b \to 0$ and $\lambda_b \to 1$ limits, which are the points Eab and EAb (black circles at the end of the blue line), respectively. Section 2 of SOM explains how to transfer the present analysis to the remaining equilibrium points of groups 1 and 2. The stability matrix for any equilibrium point Eb has the following eigenvalues and eigenvectors:

- $\zeta_1 = \lambda_b$: $\mathbf{v}_1 = (1, -2\lambda_b, 0)$.
- $\zeta_2 = 1 - \lambda_b$: $\mathbf{v}_2 = (0, 1 - 2\lambda_b, 1)$.
- $\zeta_3 = 1$: $\mathbf{v}_3 = (0, 1, 0)$.

Because $\zeta_3 = 1$ the $\mathbf{v}_3$ direction is neutral, and initial conditions slightly shifted from the equilibrium points in this direction are not repelled nor attracted. This is consistent with the fact that this direction corresponds to the $p_{Ab}$ axes, where the entire set of equilibrium points $E_b$ is located, as well as points $E_{ab}$ and $E_{Ab}$.

In the directions spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$, the equilibrium points $E_b$ are stable (the points $E_{ab}$ and $E_{Ab}$ are also stable, although the analysis for these point can not be performed in the standard linear way). Moreover, as demonstrated in Appendix C, the line spanned by $\mathbf{v}_1$ contains the equilibrium point $EU_1$ (orange circle), whereas the line spanned by $\mathbf{v}_2$ contains the point $EU_2$ (orange circle), as illustrated in figure B-1.

# C Properties of eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$

In this appendix we demonstrate, for the case of equilibrium points $E_b$ and its limits $\lambda_b \to 0$ (point $E_{ab}$) and $\lambda_b \to 1$ (point $E_{Ab}$), that the line spanned by $\mathbf{v}_1$ contains the point $EU_1$ and the line spanned by $\mathbf{v}_2$ contains the point $EU_2$. The demonstrations can be directly extended to the remaining stable fixed points.

To prove the first statement, consider the straight line $l_1$ defined by

$$(p_{AB}, p_{Ab}, p_{aB}) = \chi\ \mathbf{v}_1 + (0, \lambda_b, 0). \tag{C.1}$$

The value of $\chi$ value for which $p_{Ab} = 0$ is $\chi = 1/2$. Replacing this value in equation (C.1) we find $p_{AB} = 1/2$ which is exactly the coordinates of $EU_1$. Accordingly, for any $0 \leq \lambda_b \leq 1$, the line spanned by the eigenvector $\mathbf{v}_1$ crosses both the specific fixed point in the set $E_b \cup E_{ab} \cup E_{Ab}$ and the fixed point $EU_1$.

By constructing a second straight line $l_2$

$$(p_{AB}, p_{Ab}, p_{aB}) = \chi\ \mathbf{v}_2 + (0, \lambda_b, 0) \tag{C.2}$$

we can prove the second statement. We need now to solve for the $\chi$ value for which $p_{Ab} = 1/2$, which again gives $\chi = 1/2$. The corresponding $p_{aB}$ value is $p_{aB} = 1/2$, which implies that the line spanned by the eigenvector $\mathbf{v}_2$ crosses both the specific fixed point in the set $E_b$ (or $E_{ab}$ and $E_{Ab}$) and the fixed point $EU_2$.

# D Conserved quantities

We start by writing the evolution equations for the allele frequencies using equations (12-15). In this appendix we omit the superscript $t$ on the right hand side of the equations to simplify the notation. For $\tilde{p}_A = p_{AB} + p_{Ab}$ we obtain

$$\tilde{p}_A^{t+1} = p_{AB}^{t+1} + p_{Ab}^{t+1} = \frac{p_{AB} + p_{Ab} - \Delta}{1 - 2\Delta} = \frac{\tilde{p}_A - \Delta}{1 - 2\Delta} \tag{D.1}$$

where

$$\Delta = p_{AB}p_{ab} + p_{Ab}p_{aB}. \tag{D.2}$$

Notice that, as the relationship between the alleles and the genotype frequencies can not be inverted, it is not possible to obtain a closed formula

in terms of the allele frequencies only. Here, the presence of the genotype frequencies is implicit in $\Delta$. The calculation for $\tilde{p}_B$ is similar and we obtain

$$\tilde{p}_B^{t+1} = \frac{\tilde{p}_B - \Delta}{1 - 2\Delta} \tag{D.3}$$

We could proceed in the same way for the remaining two allele frequencies. However, it is instructive to derive them using the normalization conditions. For instance,

$$\tilde{p}_a^{t+1} = 1 - \tilde{p}_A^{t+1} = \frac{1 - 2\Delta}{1 - 2\Delta} - \frac{\tilde{p}_A - \Delta}{1 - 2\Delta} = \frac{1 - \tilde{p}_A - \Delta}{1 - 2\Delta} = \frac{\tilde{p}_a - \Delta}{1 - 2\Delta}. \tag{D.4}$$

and similarly for $\tilde{p}_b$. Interestingly, all allele frequencies obey the same evolution equation, which can thus be condensed in the form

$$\tilde{p}_u^{t+1} = \frac{\tilde{p}_u - \Delta}{1 - 2\Delta}, \tag{D.5}$$

for $u = A, B, a, b$. Subtracting $1/2$ from both sides leads to

$$\tilde{p}_u^{t+1} - 1/2 = \frac{\tilde{p}_u - \Delta}{1 - 2\Delta} - \frac{1 - 2\Delta}{2(1 - 2\Delta)} = \frac{2\tilde{p}_u - 2\Delta - 1 + 2\Delta}{2(1 - 2\Delta)} = \frac{\tilde{p}_u - 1/2}{1 - 2\Delta}. \tag{D.6}$$

Now, consider two arbitrary alleles $u'$ and $u''$. From equation (D.6) it turns out that

$$\frac{\tilde{p}_{u'}^{t+1} - 1/2}{\tilde{p}_{u''}^{t+1} - 1/2} = \frac{\tilde{p}_{u'} - 1/2}{\tilde{p}_{u''} - 1/2}. \tag{D.7}$$

Of course, this relationship holds for $u'$ and $u''$ at the same locus, but in this case it is trivial. However, when the alleles are taken at different loci (i. e., in the combinations $A - B$, $A - b$, $a - B$, or $a - b$), the resulting relationship is very interesting. For instance, for $u' = A$ and $u'' = B$, one gets

$$\frac{\tilde{p}_A^{t+1} - 1/2}{\tilde{p}_B^{t+1} - 1/2} = \frac{\tilde{p}_A - 1/2}{\tilde{p}_B - 1/2}. \tag{D.8}$$

which means that the quantity

$$T = \frac{\tilde{p}_A - 1/2}{\tilde{p}_B - 1/2} \tag{D.9}$$

remains constant from the first generation on. Moreover, as relationship (D.7) for the other 3 possible combinations of the 2 alleles can be derived directly from equation (D.8), it turns out that there is only one independent choice to define the constant $T$. We will keep the definition given in equation (D.9).

# E Dynamics and Geometrical interpretation

The dynamic is constrained by two two linear relations: normalization

$$p_{AB} + p_{Ab} + p_{aB} + p_{ab} = 1$$

and constancy of $T$

$$T(p_{AB} + p_{aB} - 1/2) - p_{AB} - p_{Ab} + 1/2 = 0.$$

The trajectory of any initial condition is, therefore, restricted to a two-dimensional plane in the four dimensional space of genotype frequencies. We have already used the normalization condition to draw the three-dimensional figures 2 and B-1. It is easy to check that the plane defined by equation (17) always contains the fixed points ES, $EU_1$ and $EU_2$, since they trivially satisfy the linear relation above for any value of $T$.

We call this plane where the dynamics take place the T-plane and note that it is completely defined by the initial condition. Since it always contains the straight line connecting $EU_1$ and $EU_2$ passing through ES, changing the value of $T$ only rotates the plane around this line. Depending on the value of $T$ the plane will intersect simultaneously the lines of fixed points Eb and EB (as illustrated in figure 3) or the lines Ea and EA. In the limiting cases $T = 1$ and $T = -1$ it intersects the points EAb-EaB and Eab-EAB respectively.

The T-plane can be fully characterized by containing the point ES and being perpendicular to the vector

$$\hat{t} = \frac{\mathbf{z}_a \times \mathbf{z}_b}{|\mathbf{z}_a \times \mathbf{z}_b|}. \tag{E.1}$$

where $\mathbf{z}_a$ is the vector from the initial condition to the point $EU_1$ and $\mathbf{z}_b$ the vector from the initial condition to the point $EU_2$. After some simplifications we find

$$\hat{t} = A(1, 1 - T, -T)$$

with $A^{-1} = \sqrt{2 - 2T + 2T^2}$.

The line passing through ES and connecting to $EU_1$ and $EU_2$ form the stable manifold of ES. Since no trajectory can cross the stable manifold, knowing on which side of this line the initial point is suffices to determine which equilibrium point will be reached.

This construction allows us to predict, for an arbitrary initial condition, the asymptotic equilibrium point of the population. First we need to specify in which plane the initial condition is located. Second, as the plane contains exactly two stable fixed points, we also need to establish on which side relative to the stable manifold of the point ES the initial condition lies.

Let us assume that $-1 < T < 1$ so that the T-plane intersects the axis as in figure 3. In this case points below the $EU_1$-$EU_2$ line move towards the $p_{Ab}$ axis, to a Eb fixed point, and points above this line move towards the EB fixed point. Let us call $X$ the point where the T-plane crosses the $p_{Ab}$ axis. Using $p_{Ab} = X$ and $p_{AB} = p_{aB} = 0$ in equation (17) we find $X = (1 - T)/2$, which emphasizes that the T-plane crosses the Eb and EB lines only if $-1 < T < 1$. For $T > 1$ or $T < -1$ the T-plane crosses the Ea and EA lines.

We define the unit vectors $\hat{u}_1$ from ES to $EU_1$ and $\hat{u}_2$ as perpendicular to $\hat{u}_1$ (and in the T-plane). These two vectors span the T-plane and are given by

$$\hat{u}_1 = \frac{1}{\sqrt{3}}(-1, 1, -1)$$

and

$$\hat{u}_2 = B(1 - 2T, -(1 + T), -2 + T)$$

where $B^{-1} = \sqrt{6(1 - T + T^2)}$ is the normalization.

Any point $\vec{p}$ on the T-plane can therefore be written as

$$\vec{p} = \vec{ES} + \alpha\hat{u}_1 + \beta\hat{u}_2. \tag{E.2}$$

If $\beta > 0$ the point is the lower half of the T-plane and the equilibrium point is of the type Eb with $p_{Ab} = X = (1-T)/2$, $p_{AB} = p_{aB} = 0$, $p_{ab} = (1+T)/2$. Otherwise the point is on the upper half and goes to the EB fixed point given by $p_{Ab} = 0$, $p_{AB} = (1 + T)/2$, $p_{aB} = (1 - T)/2$, $p_{ab} = 0$.

Taking the scalar product of equation (E.2) with $\hat{u}_2$ we find

$$\beta = \sqrt{\frac{1 - T + T^2}{6}}(1 - 2\tilde{p}_B) \tag{E.3}$$

where $(1 - T + T^2) > 0$ for $-1 < T < +1$.

This rather complicated calculation turns out to have a very simple interpretation: $\beta > 0$ only if $\tilde{p}_B < 1/2$. Also, from equation (17) we find that $\tilde{p}_B < \tilde{p}_A$, otherwise $T > 1$. Similarly, $\tilde{p}_B < \tilde{p}_a$, otherwise $T < -1$. The conclusion is that if $\tilde{p}_B$ is the smallest of the allele frequencies, the equilibrium point is $p_{Ab} = (1 + T)/2$, $p_{AB} = p_{aB} = 0$, $p_{ab} = (1 - T)/2$ which has $\tilde{p}_B = 0$.