Evolution and genetic architecture of disassortative mating at a locus under heterozygote advantage

Ludovic Maisonneuve[1,*], Mathieu Joron[2], Mathieu Chouteau[3] and Violaine Llaurens[1]

1. Institut de Systematique, Evolution, Biodiversité (ISYEB), Museum National d'Histoire Naturelle, CNRS, Sorbonne-Université, EPHE, Université des Antilles, 45 rue Buffon, 75005 Paris, France;

2. Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 CNRS-Université de Montpellier, École Pratique des Hautes Études, Université Paul Valéry, 34293 Montpellier 5, France;

3. Laboratoire Ecologie, Evolution, Interactions Des Systèmes Amazoniens (LEEISA), USR 3456, Université De Guyane, IFREMER, CNRS Guyane, 275 route de Montabo, 97334 Cayenne, French Guiana;

∗ Corresponding author: Ludovic Maisonneuve; e-mail: ludovic.maisonneuve@mnhn.fr.

Manuscript elements: Table 1, figure 1, figure 2, figure 3, figure 4, figure 5, figure S2, figure S3, figure S4,figure S5 , figure S6, figure S7, figure S8, figure S9, figure S10, figure S11, figure S12.

Keywords: Heterogamy, Supergene, Frequency dependent selection, genetic load, mate preference, *Heliconius numata*.

Manuscript type: Article.

Abstract

The evolution of mate preferences may depend on natural selection acting on the mating cues and on the underlying genetic architecture. While the evolution of assortative mating acting on locally adapted traits has been well-characterized, the evolution of disassortative mating is poorly characterized. Here we aim at understanding the evolution of disassortative mating for traits under strong balancing selection, by focusing on polymorphic mimicry as an illustrative example. Positive frequency-dependent selection exerted by predators generates local selection on mimetic colour pattern. In this well-characterized adaptive landscape, polymorphic mimicry is rare but had been reported in a butterfly species where chromosomal inversions control mimetic colour pattern variations. Because inversions are often associated with recessive deleterious mutations, we hypothesize they may induce a heterozygote advantage at the colour pattern locus, putatively favoring the evolution of disassortative mating. To explore the conditions underlying the emergence of disassortative mating, we modeled both a colour pattern locus and a mate preference locus. We confirm that a heterozygote advantage favors the evolution of disassortative mating and show that disassortative mating is more likely to emerge if at least one adaptive allele is free from any genetic load. Comparisons of hypothetical genetic architectures underlying mate choice behaviors show that rejection alleles linked to the colour pattern locus can be under positive selection and enable the emergence of disassortative mating behavior. Our results therefore provide relevant predictions on both the selection regimes and the genetic architecture favouring the emergence of disassortative mating, which could be compared to empirical data that are starting to emerge on mate preferences in wild populations.

Introduction

Mate preferences often play an important role in shaping traits diversity in natural populations, but the mechanisms responsible for their emergence often remain to be characterized. While the evolution of assortative mating on locally adapted trait is relatively well understood (Otto et al., 2008; de Cara et al., 2008; Thibert-Plante and Gavrilets, 2013), the selective forces involved in the evolution of disassortative mating are still largely unknown. Disassortative mating, *i.e.* the preferential crosses between individuals displaying different phenotypes, is a rare form of mate preference (Jiang et al., 2013). In populations where individuals tend to mate with phenotypically distinct partners, individuals with a rare phenotype have a larger number of available mates, resulting in a higher reproductive success. By generating a negative frequency-dependent selection on mating cues, disassortative mating is thus often regarded as a process generating and/or maintaining polymorphism within populations of various species. Obligate disassortative mating for sexes or mating types leads to the persistence of intermediate frequencies of sexes or mating types (Wright, 1939), and promotes polymorphism (e.g. the extreme case of some Basidiomycete fungy where thousands of mating types are maintained (Casselton, 2002)). Some cases of Disassortative mating can be based on different traits: in Amphidromus inversus snails, greater reproductive success is observed in pairs exhibing different shell chirality (Schilthuizen et al., 2007). Disassortative mating based on odors has also been reported in mice (Penn and Potts, 1999) and humans (Wedekind et al., 1995): odor profiles are indeed tightly linked to genotypes at the MHC loci controlling for variations in the immune response, known to be under strong balancing selection (Piertney and Oliver, 2006). The balancing selection in MHC partly stems from heterozygous advantage, whereby heterozygous genotypes might confer an ability to recognize a larger range of pathogens. Such a heterozygote advantage may thus promotes the evolution of disassortative mating (Tregenza and Wedell, 2000). Extreme examples of heterozygotes advantage have been observed for loci at which homozygotes have reduced survival. In the seaweed fly *Coelopa frigida* the heterozygotes ($\alpha\beta$) at the locus Adh have a higher fitness than

homozygotes ($\alpha\alpha$ or $\beta\beta$) (Butlin et al., 1984; Mer) and females prefer males with a genotype that differ from their own (Day and Butlin, 1987). In the white-throated sparrow *Zonotrichia albicollis*, a strong disassortative mating has been reported with respect to the colour of the head stripe, associated with chromosomal dimorphism (Throneycroft, 1975). This plumage polymorphism is controlled by a single locus (Tuttle et al., 2016), where a lack of homokaryotype individuals is documented (Horton et al., 2013).

While the fitness advantage of disassortative mating targeting loci with overdominance seems straightforward, the genetic basis of disassortative mating preferences remains largely unknown. One exception is the self-incompatibility system in *Brassicaceae* for which the S-locus controls for a specific rejection of incompatible pollens (Hiscock and McInnis, 2003). S-haplotypes contains tightly linked, co-evolved SCR and SRK alleles, encoding for a protein of the pollen coat and a receptor kinase located in the pistil membrane respectively, preventing fertilization from self-incompatible pollen due to specific receptor-ligand interactions. Self-rejection has also been proposed as an explanation for the disassortative mating behavior associated with odor in humans. Body odors are strongly influenced by genotypes at the immune genes HLA and rejection of potential partners has been shown to be related to the level of HLA similarity, rather than to a particular HLA genotype (Wedekind and Füri, 1997). In the white-throated sparrow, disassortative mating results from specific preferences for colour plumage that differ between males and females; *tan*-striped males are preferred by all females while *white*-striped females are preferred by all males (Houtman and Falls, 1994). Different mechanisms leading to mate preferences and associated genetic architecture can be hypothesized, that may involve the phenotype of the chooser. Based on the categories described by Kopp et al. (2018), we assume that disassortative mating behavior can emerge from two main mechanisms. (1) *Self-referencing*, *i.e.* when individual used its own signal to choose its mate, may generate a disassortative mating behavior that depends on the phenotypes of both the choosing and the chosen partners. (2) Preferences for or rejection of a given phenotype in the available partners (*recognition/trait* hypothesis), independently from the phenotype of the choosing partner, may also enable the emergence of

disassortative mate preferences. These two mechanisms could involve a two locus architecture where one locus controls the mating cue and the other one the preference toward the different cues (Kopp et al., 2018). The level of linkage disequilibrium between the two loci could have a strong impact on the evolution of disassortative mating. In models investigating the evolution of assortative mating on locally-adapted traits, theoretical simulations have demonstrated that assortative mating is favored when the preference and the cue locus are linked (Kopp et al., 2018).

Here we explore the evolutionary forces leading to the emergence of disassortative mating behavior. We focus on the specific case of the butterfly species *Heliconius numata*, where high polymorphism in wing pattern is maintained within population (Joron et al., 1999) and strong disassortative mating has been documented between wing pattern forms (Chouteau et al., 2017). *H. numata* butterflies are chemically-defended (Arias et al., 2016), and their wing patterns act as warning signals against predators. At a local scale, natural selection leads to the fixation of a single warning signal shared among sympatric defended species (Müllerian mimicry) (Merrill et al., 2015). However, local polymorphism of mimetic colour patterns can still emerge within species in some balancing conditions between migration and local selection for specific mimetic patterns (Joron and Iwasa, 2005). The local polymorphism of several mimetic patterns observed within populations of *H. numata* (Joron et al., 1999) would then require a high migration rate compensating for the strong local selection. However, disassortative mating based on wing pattern has been reported in *H. numata*, with females rejecting males displaying the same colour pattern (Chouteau et al., 2017). Such disassortative mating behavior could then enhance the local polymorphism in colour pattern within this species. This mating behavior could, in turn, promotes migration because immigrant individuals exhibiting a locally rare phenotype would benefit from an increased reproductive success. Nevertheless, the evolution of such disassortative mating pattern is unclear, notably because the mate preference implied should be strongly counter-selected by predators preferentially attacking locally rare, non-mimetic warning patterns (Chouteau et al., 2016). Building on this well-documented case study, we use a theoretical ap-

proach to provide general predictions on the evolution of disassortative mating in polymorphic traits, and on expected genetic architecture underlying this behavior.

Variation in wing colour patterns of *H. numata* is controlled by a single genomic region, called the supergene P (Joron et al., 2006), and displaying chromosomal inversions (Joron et al., 2011). These inversions have recently been shown to be associated with a significant genetic load, resulting in a strong heterozygote advantage (Jay et al., 2019). We thus investigate whether genetic load associated with locally adaptive alleles may favor the evolution of mate preference and promote local polymorphism. We then explored two putative genetic architectures of mate preferences based on (1) *self referencing* and (2) based on *recognition/trait* rule, and test for their respective impact on the evolution of disassortative mating behavior. Under both hypotheses, we assumed that the mating cue and the mating preference are controlled by two distinct loci, and investigate the effect of linkage between loci on the evolution of disassortative mating behavior.

Methods

Based on earlier model of Müllerian mimicry (Joron and Iwasa, 2005) extended to diploid populations (Llaurens et al., 2013), we describe a two-population model with a locus P controlling mimetic colour pattern. Polymorphism in mimetic colour pattern is maintained within each population, by a balance between opposite local selection on colour pattern in the two populations and migration between populations. We assume different levels of genetic load associated with the colour pattern alleles at locus P. We then explicitly model the genetic architecture controlling mate preference toward colour pattern by a locus M assuming either (1) a preference toward similar or dissimilar phenotypes, which thus also depends on the phenotype of the choosing individual, following the *self-referencing* hypothesis or (2) a preference for a given colour pattern displayed by mating partner, independent from the colour pattern of the choosing individual, following the *recognition/trait* hypothesis. A recombination rate ρ between the colour pattern locus P and the preference locus M is assumed.

In the model, each individual is thus described by its genotype i, corresponding to a given colour pattern phenotype $[i]$. Genotypes are composed of four alleles, as follows:

$$i = (p_1, p_2, m_1, m_2), \tag{1}$$

where p_1 and p_2 are two alleles at the locus P and m_1 and m_2 the two alleles at the locus M. We then track down the evolution of genotype frequencies at both the locus P controlling variations in wing colour pattern and the locus M controlling mate preference. The evolution of the number of genotype i within each populations pop through time is influenced by reproduction, survival, predation and migration between patches, following the general equation :

$$\frac{d}{dt}N_{i,pop} = dP_{i,pop} + dR_{i,pop} + dM_{i,pop} + dS_{i,pop}, \tag{2}$$

where $P_{i,pop}$, $R_{i,pop}$, $M_{i,pop}$, and $S_{i,pop}$ described the contribution of these four phenomena to the change of density of genotype i within each populations pop. The computation of these four contributions is detailed below.

Mimetic colour patterns

At the colour pattern locus P, three alleles are assumed to segregate, namely alleles a, b and c, encoding for phenotypes A, B and C respectively. We assume strict dominance among the three alleles with $a > b > c$ in agreement with the strict dominance observed among supergene P alleles within natural populations of *H. numata* (Le Poul et al., 2014). The three colour pattern phenotypes are assumed to be perceived as strictly different by both mating partners and predators. The resemblance $Res[i][j]$ between pairs of individuals exhibiting phenotypes $[i]$ and $[j]$ respectively is thus set to 1 for identical phenotypes and to 0 for dissimilar ones. The resemblance matrix among the three phenotypes is :

$$Res = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Opposite local adaptations in colour pattern

Local selection promotes convergent evolution of wing colour pattern among defended species (i. e. Müllerian mimicry, (Müller, 1879)), forming so-called mimicry rings composed of individuals from different species displaying the same warning signal. At a larger scale, a spatial mosaic of warning patterns can be observed, through an equilibrium between colonization and selection acting locally (Sherratt, 2006).

Here we assume two populations of an unpalatable species involved in Müllerian mimicry with other chemically-defended species. The environment differs in communities of local species involved in mimicry (*i.e.* mimicry rings): we consider two patches occupied by different mimetic communities: population 1 is located in a patch where the local community (*i.e.* other chemically-defended species, not including *H. numata*) mostly displays phenotype A, and population 2 in a patch where the mimetic community mostly displays phenotype B, therefore generating opposite local adaptation in the two patches. Note that the allele c, and corresponding phenotype C is non-mimetic in both patches and will then be disadvantaged in both patches). Local adaptation is modelled via a parameter σ (Strength of local adaptation), which modulates the predation risk associated with the different phenotypes in the two patches (Joron and Iwasa, 2005). When $\sigma = 0$, all phenotypes A, B and C are equally adapted to the environment in both patches (the patch-dependent component of predation is equal for all phenotypes). When $\sigma > 0$, individuals carrying phenotype A and B are favored in the first and the second patch respectively, because their level of predation is decreased by $1 - \sigma$. Hence, the higher σ, the greater the difference between the two communities leading to spatial heterogeneity favouring phenotype A in patch 1 and phenotype B in patch 2. This strength of local adaptation σ plays a central role in the

predation exerted on the different phenotypes in the two patches (see Predation section below).

Positive frequency-dependent predation

Every individual of the focal (polymorphic) species is exposed to a predation risk modulated by its resemblance to the local mimetic community of butterflies. We assume a symmetrical condition where the mortality coefficient was $d(1-\sigma)$ for phenotypes matching the local mimicry ring (*i.e.* diminishing predation exerted on genotypes displaying phenotype A in population 1 and genotypes displaying B in population 2) and $d(1+\sigma)$ otherwise (*i.e.* increasing predation exerted on genotypes displaying phenotype B or C in population 1 and on genotypes displaying phenotype A or C in population 2), where d represents the baseline predation risk and σ the strength of local adaptation due to mimicry communities in patch 1 and 2.

Predation exerted on a given phenotype depends on its match to the local mimetic environment, but also on its own abundance in the patch. Predators learn to associate warning patterns to chemical defense. This learning behavior generates a positive frequency-dependent selection (pFDS) on butterfly wing patterns (Chouteau et al., 2016), because displaying a widely shared colour pattern decreases the risk of encountering a naive predator (Sherratt, 2006). Number-dependent predator avoidance in the focal species is assumed to depend on its unpalatability coefficient (λ) and on the density of each phenotype, so that the protection gained by phenotypic resemblance is greater for higher values of the unpalatability coefficient λ. This results in the following change in the number of each genotype i in population pop due to predation :

$$dP_{i,pop} \quad = \quad -\frac{d}{1+\lambda(\sum_j Res_{[i],[j]} N_{j,pop})}[(1\ +\ \sigma)(1\ -\ Res_{[i],[pop]})\ +\ (1\ -\ \sigma)Res_{[i],[pop]}]N_{i,pop}. \quad (3)$$

with $N_{i,pop}$ representing the total number of individuals with genotype i in population pop,$[pop]$ the phenotype of individual favors by the local slection of the patch pop ($[pop]$ is equal to A in the first patch and to B in the second) then $Res_{[i],[pop]}$ the resemblance of the phenotype expressed by genotype i to the local mimetic community. The predation rate is indeed lower for individuals

displaying the phenotype mimetic to the local community (*i.e.* the phenotype A in population 1 and B in population 2). Individuals displaying phenotype C being non-mimetic in both populations, suffer from a high predation risk in both populations. The denominator models the positive number dependent selection, the protection against predators being stronger for higher values of unpalatability.

Migration

We assume a constant migration rate mig corresponding to the proportion of the population migrating to the other patch. The change in the number of individuals with genotype i in population pop due to migration between populations pop and pop' is given by:

$$dM_{i,pop} = mig(N_{i,pop'} - N_{i,pop}). \tag{4}$$

where mig is the migration coefficient $mig \in [0, 1]$).

Mate preferences

The mate preference is considered as strict, implying that choosy individuals never mate with individuals displaying their non-preferred phenotype. Two hypothetical mate preference mechanisms are investigated. Under the *self-referencing* hypothesis (hyp 1), two alleles are assumed at loci M, coding for (i) random mating (r) and (ii) preferential mating behavior (either assortative sim or disassortative dis) respectively (see fig. S2 for more details). We assume that the *self-referencing* preference alleles sim and dis are dominant over the random mating allele r (see fig. S1 for more details). The dominance relationships between the sim and dis alleles are not specified however, because we investigate independently the evolution of assortative and disassortative mating from a population ancestrally mating at random. Note that under the *self-referencing* hypothesis (hyp. 1), mate choice thus crucially depends not only on the colour pattern of choosen partner, but also on the phenotype of the individual expressing the preference.

An alternative mechanism of mate preference is also investigated, assuming a specific recognition of colour patterns acting as mating cue (*recognition/trait*, hyp.2). Under hyp.2, four alleles segregate at locus M: allele M_r, coding for an absence of colour pattern recognition (leading to random mating behavior), and M_a, M_b and M_c coding for specific recognition of colour pattern phenotypes A, B and C. The 'no preference' allele M_r is recessive over all the preference alleles M_a, M_b and M_c, and preference alleles are co-dominant, so that heterozygotes at the locus M can recognize two different alleles. Then, the recognition enabled by preference alleles M_a, M_b and M_c triggers either attraction (hyp.2.a) or repulsion (hyp.2.b) toward the recognized colour pattern, leading to assortative or disassortative mating behavior depending on the genotype i of the chooser and the phenotype of the choosen individual (see figure S3 and S4 for more details).

To characterize female mating preferences generated by the different genotypes at locus M and the link with their own colour pattern phenotype, we distinguish two main behaviors emerging under hyp.2 (fig. S3 and S4 for attraction (hyp.2.a) and rejection (hyp.2.b) hypotheses respectively):

- Self-acceptance : females mate with males displaying their own colour pattern phenotype.
- Self-avoidance : females do not mate with males displaying their own colour pattern phenotype.

Note that under the rejection hypothesis (hyp.2.b), haplotypes $a - M_a$, $b - M_b$ and $c - M_c$ will trigger rejection of their own colour pattern alleles. In contrast, under the attraction hypothesis (hyp.2.a), haplotype $a - M_b$ and $a - M_c$ will trigger self-avoidance but also avoidance of another allele (here allele c and b respectively). The evolution of disassortative mating under the attraction hypothesis (hyp.2.a) may thus be impaired by the more limited number of available partners for individuals carrying these haplotypes.

In order to compare the mating behaviors observed under attraction (hyp.2.a) and rejection (hyp.2.b) hypotheses, we thus compute the population statistics, $P_{s\text{-}acc}$ (see equation (5)) and $P_{s\text{-}av}$ (see equation (6)) as the proportion of individuals exhibiting respectively a self-acceptance or a

self-avoidance behavior toward their own phenotype throughout both populations. These two inferred behaviors can be directly compared with mate preferences empirically estimated. For example, in experiments where females can choose partners among males displaying colour patterns (Chouteau et al., 2017), the proportion of females mating with males displaying their own phenotype colour pattern can be easily scored and compared to the proportion of self-accepting individuals computed in our model.

$$P_{s\text{-}acc} = \sum_i f_i Pref_{i[i]}, \tag{5}$$

$$P_{s\text{-}av} = \sum_i f_i (1 - Pref_{i[i]}). \tag{6}$$

Reproduction

We assume that the two populations have identical carrying capacity K and growth rate r. We name $N_{tot,pop}$ the total density of individuals in population pop. We assume separated sexes and obligate sexual reproduction, and therefore compute explicitly the Mendelian segregation of alleles during reproduction, assuming a recombination rate ρ between the colour pattern locus P and the preference locus M. We assume that the frequency of male and female of a given phenotype is the same. The frequency of the offspring of genotype i in population pop ($f'_{i,pop}$) then also depends on the frequencies of each genotype in the population and the mate preferences computed in equation (9). We assume a single choosy sex: only females can express preferences toward male phenotypes, while males have no preference and can mate with any accepting females. The genotype of the choosy partners i is thus determining the probability of crosses with partners displaying phenotype $[j]$.

The preference matrix $Pref$ is defined as $Pref_{i,[j]} = 1$ when females with genotype i accept males with genotype j as mating partners and $Pref_{i,[j]} = 0$ otherwise. We define the fertility $F_{i,pop}$

of individuals carrying genotype i in the population pop as:

$$F_{i,pop} = Pref_{i,A}P_{A,pop} + Pref_{i,B}P_{B,pop} + Pref_{i,C}P_{C,pop}. \tag{7}$$

We note $\overline{F}_{pop}$ the average fertility in the population defined as

$$\overline{F}_{pop} = \sum_i f_{i,pop}F_{i,pop}, \tag{8}$$

where P_I, pop refer to as the proportion of the phenotype I in the population pop.

Because choosy individuals might have a reduced reproductive success due to limited mate availability (Kirkpatrick and Nuismer, 2004), we also assume a cost associated with choosiness, referred to as $cost$. When this cost is low ($cost = 0$), females have access to a large number of potential mates, so that their fertility is not limited when they become choosy ("Animal" model). When this cost is high ($cost = 1$), females have access to a limited number of potential mates, so that their fertility tends to decrease when they become choosy ("Plant" model). Intermediate values of $cost$ implies that females can partially reallocate the fitness loss due to the encountering of non-preferred males towards reproduction with other males. This cost of choosiness is known to limit the evolution of assortative mating (Otto et al., 2008) and may thus also limit the emergence of disassortative mating.

$$f'_{i,pop} = \sum_{j,k} coef(i,j,k,\rho) \underbrace{f_{j,pop}}_{\substack{\textit{frequency of female} \\ \textit{with genotype } j}} \underbrace{\frac{1 - cost + cost F_{j,pop}}{1 - cost + cost \overline{F}_{pop}}}_{\substack{\textit{relative probability that female} \\ \textit{of genotype } j \textit{ has mated}}} \underbrace{Pref_{j[k]} \frac{f_{k,pop}}{F_{j,pop}}}_{\substack{\textit{probability male is type } k \\ \textit{given female has mated}}}, \tag{9}$$

where $coef$ controls the mendelian segregation of alleles during reproduction between an individual of genotype j and an individual of genotype k, therefore depending on the recombination rate ρ between the colour pattern locus P and the preference locus M (see supplementary S1 for detailed expression of $coef(i,j,k,\rho)$).

Overall, the change in the number of genotype i in population pop generated by sexual reproduction is given by:

$$dR_{i,pop} = r(1 - \frac{N_{tot,pop}}{K})N_{tot,pop}f'_{i,pop}. \tag{10}$$

Survival

We assume a mortality rate δ. The recessive genetic loads δ_1, δ_2, δ_3 associated with the respective alleles a, b and c limit the survival probabilities of homozygous genotypes at locus P. Dominant alleles are usually derived alleles associated with inversions (see Llaurens et al. (2017) for a review) whereas recessive alleles are generally carried by the ancestral gene order. We thus expect that the genetic load associated with the most dominant allele a and the intermediately dominant allele b have similar strength because of the deleterious mutations captured by the inversions, i.e. $\delta_1 = \delta_2$. These genetic loads associated with dominant alleles could then be higher than the genetic load associated with the recessive allele c, namely δ_3.

$$\delta_i = \begin{cases} \delta_1 & \text{if } (p_1, p_2) = (a, a) \\ \delta_2 & \text{if } (p_1, p_2) = (b, b) \\ \delta_3 & \text{if } (p_1, p_2) = (c, c) \\ 0 & \text{else} \end{cases}. \tag{11}$$

$$dS_{i,pop} = -(1 - (1 - \delta)(1 - \delta_i))N_{i,pop}. \tag{12}$$

Tracking the evolution of the two populations using numerical analyses

Since the model describes the change in genetic densities at a population level, we assume that predation P, migration M, reproduction R and survival S occur at the same time (see first Equation). In a given population, some individuals can die because of predator attack, while other may reproduce or migrate.

Abbreviation	Name	Parameter range
N_i^0	Initial size of the population i	100
d	Predation strength	[0,1]
σ	Strength of local adaptation	0.5
λ	Unpalatability coefficient	0.0002
mig	Migration rate	[0,1]
ρ	Recombination rate	[0, 0.5]
r	Growth rate	2
K	Carrying capacity	2000
δ	Baseline death rate	[0, 1]
δ_i	Genetic load linked to allele i	[0, 1]
$cost$	cost of choosiness	[0, 1]

Table 1: **Description of parameters used in the model and range explored in simulations.**

The complexity of this two-locus diploid model prevents comprehensive exploration with analytical methods. The model is thus studied using deterministic simulations in order to provide general prediction while ignoring the effects of stochastic processes such as drift. Our predictions are therefore relevant for species with large effective population sizes, such as *H. numata*. All parameters and range values used in the different simulations are summarized in Table 1 above. The values of the predation strength d, the strength of local adaptation σ, the unpalatability λ and migration rate mig are chosen as conditions maintaining balanced polymorphism at the colour pattern locus P, drawn from (Joron and Iwasa, 2005). Simulations are performed using Python v.3. and use discrete time step as an approximation (Euler method). We check that reducing the value of the time step provide similar dynamics (see Supp. S5), insuring that our discrete-time simulations provide relevant outcomes.

Results

Effect of mate choice on polymorphism

As already highlighted in the literature (Llaurens et al., 2013), polymorphism can be maintained through an equilibrium between spatially heterogeneous selection and migration if random mating is assumed. In the absence of migration, alleles *a* and *b* become fixed in population 1 and 2 respectively, owing to their mimetic advantage within their respective communities. Polymorphism with persistence of alleles *a* and *b* within each patch can only be maintained with migration, but in all cases the non mimetic allele *c* is lost in both populations (fig.1 (a)).

To test for an effect of mate choice on the selection/migration equilibrium, described above, we carry out simulations introducing *self-referencing* preference alleles (*r*, *dis* or *sim*) at the mate choice locus (Hyp.1), assumed to be fully linked to the colour pattern locus (ρ= 0). We then compute the evolution of frequencies at the colour pattern locus after 10,000 time units for different migration rates *mig*. Assuming assortative mating via *self-referencing* (hyp. 1) leads to the fixation of the dominant allele *a* in both patches for all migration rates explored, because phenotype A is the most frequently expressed due to dominance of allele *a* and therefore benefits from a frequency-dependent advantage (fig.1 (b)). By contrast, disassortative mating maintains higher degree of polymorphism, with the two mimetic alleles *a* and *b*, and the non-mimetic allele *c* persisting within populations, for all migration rates. The non-mimetic phenotype C is rarely expressed because the recessive allele *c* only persists at low frequency, but it is associated with a high reproductive success because of disassortative mating. Indeed, the strict disassortative preference assumed here strongly increases the reproductive success of individuals displaying a rare phenotype such as C. This effect would be weakened with less stringent mate preferences. Nevertheless, the negative FDS on colour pattern generated by disassortative mating counteracts the positive FDS due to predator behavior acting on the same trait. Disassortative mate preferences can thus promote the polymorphism of alleles within and between patches.

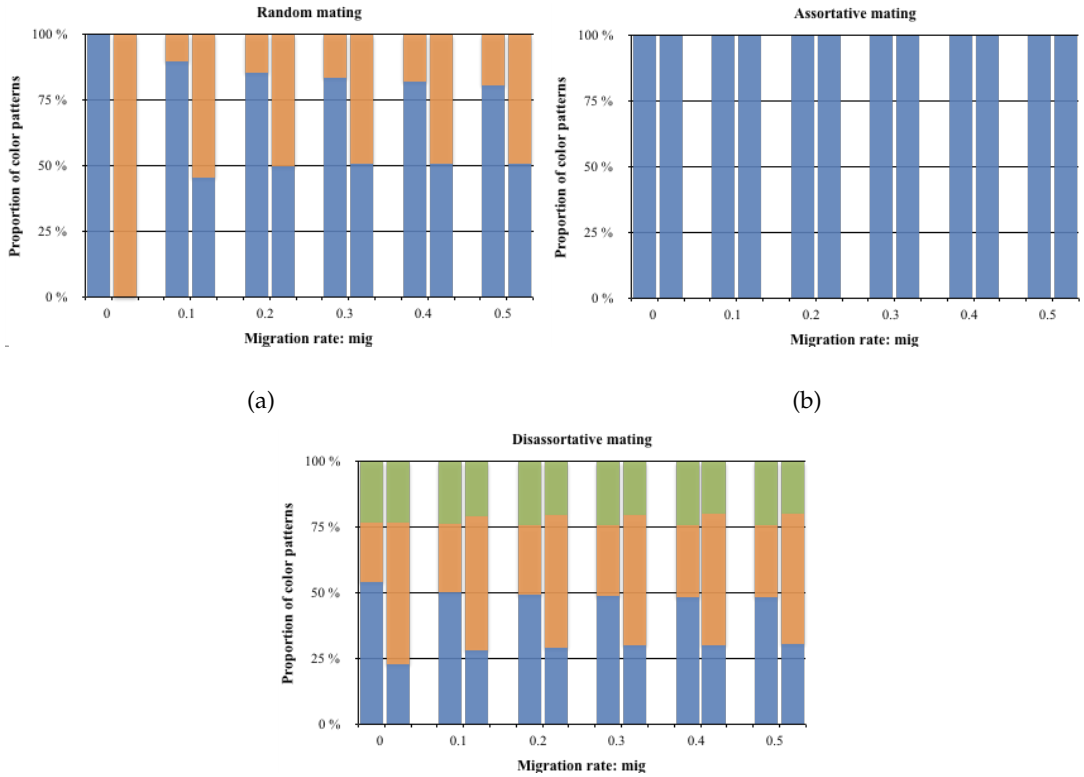


Figure 1: **Impact of mate preferences on colour pattern diversity within both populations**. The equilibrium frequencies of colour pattern phenotypes in population 1 and 2 for different migration rates *mig* are computed assuming different mating behaviors, *i.e.* assortative (a), random (b) or disassortative (c). The heights of the coloured stacked bars indicates the frequencies of colour pattern phenotypes A, B and C (as blue, orange and green areas respectively) in population 1 and 2 (on the left and right side respectively, within each level of migration). The three alleles at the locus P controlling colour pattern variations are introduced in proportion $\frac{1}{3}$ in each population. The locus M controls for the *self-referencing* based mate preferences (hyp. 1). Simulations are run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0$, $\rho = 0$, $cost = 0.1$, $\delta_1 = \delta_2 = \delta_3 = 0$ and $\delta = 0$.

Linked genetic load favors the persistence of a non-mimetic allele

In the following simulations, the migration parameter *mig* is set to 0.1, to allow the persistence of polymorphism of alleles a and b at the colour pattern locus P when assuming random mating. We then investigatethe influence of a genetic load associated with the different colour pattern alleles on polymorphism at the colour pattern locus. This allows inferences regarding the effect

of heterozygote advantage generated by genetic load on polymorphism, independently from the evolution of mating preferences. We observe that phenotypes A and B are maintained, but not phenotype C, when a genetic load is associated with the non mimetic allele c only ($\delta_1 = \delta_2 = 0$ and $\delta_3 > 0$), or when this load is stronger than the one associated with alleles a and b (Supp. table S6). However, the non-mimetic allele c is maintained with the other alleles a and b within both populations, when (i) all three alleles carry a genetic load of similar strength, *i.e.* $\delta_1 = \delta_2 = \delta_3 > 0$ or (ii) when allele c is the only one not carrying a genetic load ($\delta_1 = \delta_2 > 0$ and $\delta_3 = 0$) (Supp. table S6). The heterozygote advantage generated by the genetic load associated with the dominant mimetic alleles at the locus P thus favors the persistence of balanced polymorphism and more specifically promotes the maintenance of the non mimetic allele c within both populations.

Evolution of disassortative mating

Because we expect the heterozygote advantage at colour pattern locus P to enhance the evolution of disassortative mating preferences at the locus M, we first investigate the influence of a genetic load on the evolution of disassortative behavior by testing the invasion of *self-referencing* mutant triggering self-avoidance *dis* (hyp. 1) in a population initially performing random mating with genotype frequencies at equilibrium. We compute the frequency of mutants 100 time units after their introduction, assuming full linkage between loci P and M. Figure 2 shows that the genetic load associated with alleles a and b ($\delta_1 = \delta_2$), has a strong positive impact on the emergence of disassortative mating. The genetic load associated with the recessive allele c (δ_3) has a slighter positive effect on the evolution of disassortative mating. At a larger evolutionary scale, this leads to the fixation of the disassortative mating allele dis (see equilibrium after 10,000 time units in supp. figure S7) when the genetic load associated with the dominant alleles a and b is positive. Simulations assuming different costs associated with choosiness (*cost*) show a similar effect of associated genetic loads, although increasing this cost slows down the invasion of the choosy disassortative mating mutant *dis* (see Sup. fig. S8). Overall, this confirms that genetic load

linked to the colour pattern locus P favors the evolution of disassortative mating behavior in both populations.

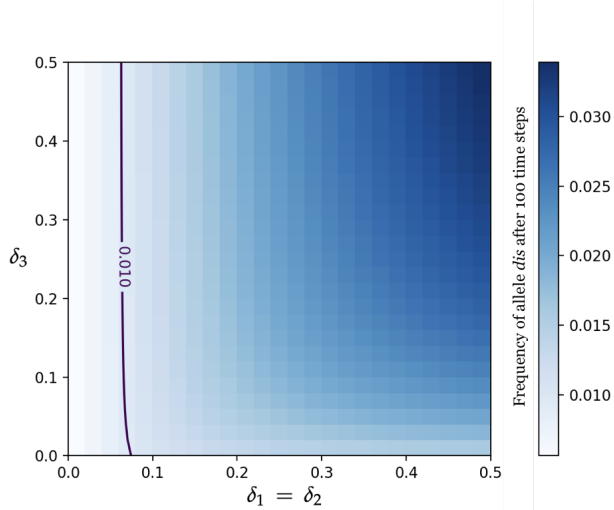
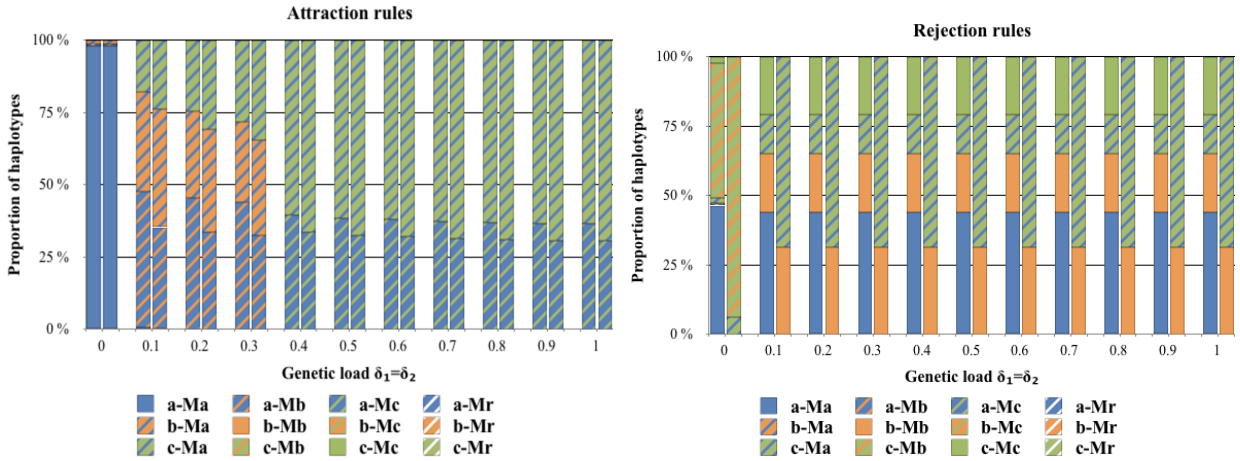


Figure 2: **Impact of linked genetic load on the evolution of disassortative mating, assuming *self-referencing* (Hyp.1).** The invasion of a mutant *dis* with disassortative mating preferences depends on the strength of the genetic load associated with dominant alleles a and b assumed equal ($\delta_1 = \delta_2$) (x axis) and with the recessive allele c (δ_3) (y axis). The shade of blue indicates the frequency of the mutant with disassortative mating preferences *dis*, inducing rejection of partners displaying the same phenotype (hyp. 1), after 100 time units. The purple line indicates the initial frequency of the mutant, set to 0.01, therefore highlighting the conditions above which an invasion by the mutant is observed. The three alleles at the locus controlling colour pattern variations are introduced in even proportion (*i.e.* $\frac{1}{3}$) in each population, and the initial frequency of the mutant are 0.01, shown by the vertical purple line marking the limit of invasion by the mutant. Simulations are run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $\rho = 0$, $mig = 0.1$, $\delta = 0.1$ and $cost = 0.1$.

How does the genetic architecture of mating preference influence the evolution of disassortative mating behavior ?

In order to study the evolution of mating behavior under the assumption of different genetic architecture of mate preferences, we investigate the invasion of *recognition/trait* alleles M_r, M_a, M_b and M_c controlling random mating and recognition of phenotype A, B and C respectively

(Hyp. 2). We assume loci P and M as fully linked ($\rho = 0$), and compare simulations where mate preference alleles trigger either attraction (Hyp. 2a) or rejection (Hyp. 2b) of the recognized colour pattern phenotype (fig.3(a) and fig.3(b) respectively).



(a) (b)

Figure 3: **Influence of a genetic load on haplotype diversity, assuming (a) attraction rule (hyp. 2a) or (b) rejection rule (hyp. 2b) at the preference locus (*recognition/trait*)**. The proportion of haplotypes obtained 2,000 time units after the introduction of preference alleles in both populations are shown for different values of genetic load associated with alleles a and b ($\delta_1 = \delta_2$). For each value of genetic load ($\delta_1 = \delta_2$) the first and second bars represented the haplotypes proportion in the patch 1 and 2 respectively. The locus M controls for a specific recognition of colour pattern alleles inducing either (a) attraction (hyp.2a) or (b) rejection (hyp.2b). The three alleles at the locus P controlling colour pattern variations are initially introduced in even proportion $\frac{1}{3}$ in each population. After 10,000 time units under random mating the four alleles at locus M: M_r, M_a, M_b and M_c are introduced respectively in proportion 0,99, $\frac{0.01}{3}$, $\frac{0.01}{3}$, $\frac{0.01}{3}$. Simulations are run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $\rho = 0$, $mig = 0.1$, $\delta_3 = 0$, $\delta = 0.1$ and $cost = 0.1$.

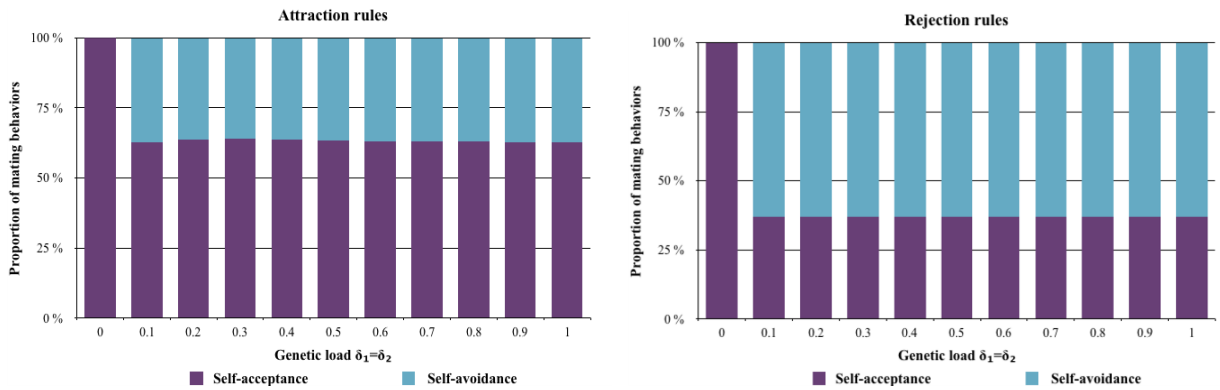
When preference alleles cause female attraction to males exhibiting a given phenotype (Hyp.2a), we observe high frequencies of haplotypes $a - M_b$ and $b - M_a$ in both populations at equilibrium, as soon as the genetic load associated with the dominant alleles δ_1 and δ_2 is greater than 0 (fig.3(a)). These two haplotypes benefit from both positive selection associated with mimicry and limited expression of the genetic load due to the preferential formation of heterozygotes. Haplotype $c - M_a$ is maintained because of the benefit associated with the choice of the most frequent

mimetic phenotype A, and the limited expression of the non-mimetic phenotype C due to the recessiveness of c. Nevertheless, haplotype $b - M_a$ is lost when the genetic load increases. Indeed, when assuming a high genetic load, most individual with phenotype A are heterozygotes (ab or ac). Because the color pattern allele a is dominant over the allele b, the haplotype $b - M_a$ may then induces preference producing a significant proportion of maladaptive bb offspring. As a consequence, the mimetic phenotype B is not maintained in populations where the genetic load is high, and the dominant phenotype A becomes predominantly expressed in both populations.

By contrast, when mate preference is based on alleles causing rejection behavior (Hyp.2b) and when a genetic load is associated with the mimetic alleles a and b at locus P, these alleles become associated with the corresponding rejection alleles at locus M (*i.e.* $a - M_a$ and $b - M_b$ have an intermediate frequency in both populations) (fig.3(b)). Non mimetic allele c becomes either associated with a self-avoiding allele $c - M_c$, or an allele rejecting the dominant allele a ($c - M_a$). The number alleles a in the second patch maintained by migration from the first patch is very small. Then $c - M_a$ is advantaged because it reduces the cost of seeking a rare partner or to produced offspring with rare phenotype in the patch more vulnerable to predation.

The three alleles (a, b and c) persist within patches for all positive values of genetic load (fig.3(b)). This contrasts with the previous case where preference alleles lead to attraction (hyp. 2a) and for which the mimetic allele b is lost when the genetic load is high (fig. 3(a)). Although equilibrium haplotype frequencies are similar for all positive values of genetic load under the assumption of preference allele coding for rejection (Hyp.2b), the strength of genetic load still impacts the temporal dynamics of haplotype frequencies, by allowing the equillibrium to be reached earlier when the generic load is higher (see sup. fig S9). This difference in the timing of invasion of the rejection haplotypes reflects higher selection coefficient associated with these haplotypes in simulations where genetic load is stronger.

We then investigate how haplotype frequencies translate into individual behavior. When we consider preference alleles leading to attraction (hyp.2a), the majority of individuals display assortative preferences at equilibrium ($P_{s\text{-}acc} \approx 63\%$), even with a high genetic load (figure 4(a)).



(a) (b)

Figure 4: **Influence of a genetic load on the distribution of mating behavior observed at the population level, assuming attraction (a) or rejection (b) alleles at the preference locus (*recognition/trait*)**. The proportion of individuals displaying self-acceptance $P_{s\text{-}acc}$ (in purple) and self-avoidance $P_{s\text{-}av}$ (in blue) obtained 2,000 time units after the introduction of preference alleles in both populations are shown for different values of the level of genetic load of δ_1 and δ_2. The locus M controls for a specific recognition of colour pattern alleles inducing either (a) attraction (hyp.2a) or (b) rejection (hyp.2b). The three alleles at the locus P controlling colour pattern variations are initially introduced in even proportion $\frac{1}{3}$ in each population. After 10,000 time units under random mating the four alleles at locus M M_r, M_a, M_b and M_c are introduced respectively in proportion 0,99, $\frac{0.01}{3}$, $\frac{0.01}{3}$, $\frac{0.01}{3}$. Simulations are run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $\rho = 0$, $mig = 0.1$, $\delta_3 = 0$, $\delta = 0.1$ and $cost = 0.1$.

This is surprising given that most haplotypes are of a "disassortative" type, *i.e.* linking a colour pattern allele with an attraction allele to a different colour pattern. Nevertheless, colour pattern alleles b and c are both linked to M_a coding for attraction to A. As a consequence, most individuals formed are heterozygous at both the colour pattern locus (with one allele a and another allele) and at the preference locus (with one allele coding for attraction to phenotype a and another allele). These double heterozygotes thus benefit from mimicry and avoid the expression of deleterious mutations, but can still mate with individuals sharing the same phenotype. By contrast, when we consider preference alleles leading to rejection (hyp.2b), most individuals display a disassortative mating behavior ($P_{s\text{-}av} \approx 63\%$) (figure 4(b)). This highlights that the genetic architecture of mate preference plays a key role in the evolution of the mating behavior of diploid individuals and that the evolution of disassortative haplotypes inducing disassortative prefer-

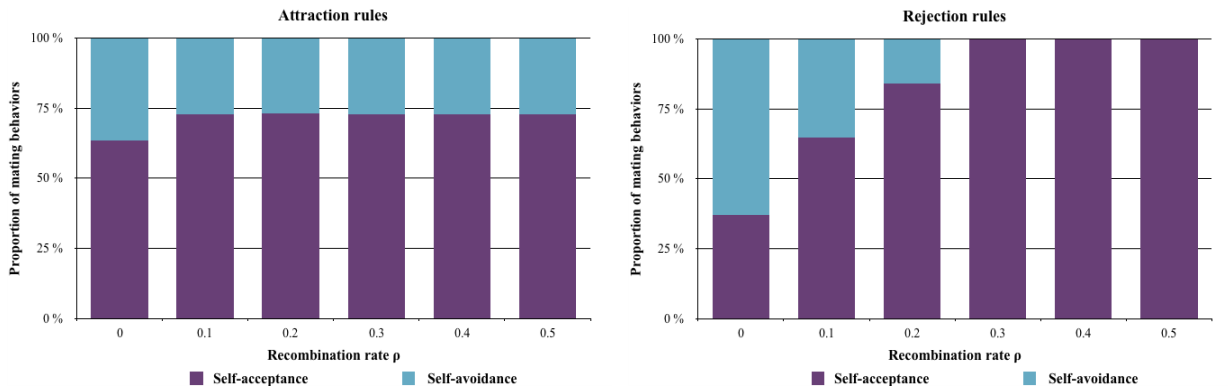
ences do not necessarily cause disassortative mating at the population level. At equilibrium, the proportion of self-avoidance behavior in the population does not depend of the strength of the genetic load (figure 4(b)). However, the strength of the genetic load does impact the speed of evolution of disassortative mating (Supp. fig. S9). This also suggests stronger positive selection on disassortative mating when the genetic load associated with dominant wing colour pattern alleles is higher.

The proportions of self-avoidance behavior observed when assuming the *attraction* rule (Hyp.2.a) ($P_{s\text{-}av} \approx 37\%$) are lower than the proportion of self-avoidance when assuming a *self-referencing* architecture ($P_{s\text{-}av} \approx 50\%$, see Supp. S10). Under *attraction* rule, an individual of a given phenotype mates with individual with at least one preference allele controlling for the attraction for this given phenotype. These crosses generate offspring which contains a color allele together with its corresponding preference allele, triggering a self-accepting behavior, because *recognition/trait* alleles are assumed codominant. In contrast, under the *rejection* rule (Hyp.2.b) disassortative mating is more likely to emerge. Indeed a "disassortative haplotype", *i.e.* a haplotype composed by a phenotype allele and its corresponding preference allele ($a - M_a$ for example) generally immediately translates into a self-avoiding behavior, whatever the genotypic combinations within individuals. The proportions of self-avoidance behavior observed when assuming the *rejection* rule (Hyp.2.b) ($P_{s\text{-}av} \approx 63\%$) are then even higher than the proportion of self-avoidance when assuming a *self-referencing* architecture. Because we assume codominance at the mate choice locus M, heterozygotes at the locus M may reject two phenotypes, and this behavior can further limit the expression of genetic load associated with the locus P in the offspring (*e.g.* a female with the genotype $a - M_a/b - M_b$ will have neither deleterious homozygotes aa nor bb in her offspring). This advantage of rejecting two phenotypes instead of one when assuming the *self-referencing* architecture (Hyp.1) may explained why the evolution of disassortative mating is facilitated when assuming an architecture based *recognition/trait* inducing *rejection* (Hyp.2.b).

Impact of linkage between loci P and M on the evolution of disassortative mating

In previous sections, we observed that the genetic load associated with the two most dominant alleles at the colour pattern locus P impacts the evolution of mate choice, and that disassortative mating is favored by an architecture based *recognition/trait* inducing *rejection* (Hyp.2.b), when the colour pattern locus P and the preference locus M are fully linked. We then test for an effect of recombination between alleles at the two loci on the evolution of mate choice by performing simulations with different values of the recombination rate ρ. Assuming *self-referencing* (hyp.1), increasing recombination rate further promotes the invasion of the disassortative allele *dis* (see Sup. fig. S11). Mate preference depends on the phenotype displayed by the individual, so that the allele *dis* always translates into a disassortative behavior, irrespective of the linkage disequilibrium between the preference locus and colour pattern locus. Increased recombination thus only results in a more rapid fixation of the disassortative mating allele *dis*, which benefits the associated genetic load. The *self-referencing* architecture (hyp. 1) is thus very similar to a single locus architecture, where a single pleiotropic gene controls both the mating cue and the rejection of this cue.

By contrast, when assuming preference for a given colour pattern allele (hyp.2), mating behavior depends on the genotype at the preference locus M, independently from the phenotype of the choosing individuals. We therefore expect a stronger effect of recombination rate on mate choice evolution. Figure 5 indeed confirms this prediction the breaking of the association between preference and wing pattern alleles enabled by recombination between locus P and M decreases the proportion of individuals performing self-avoidance at equilibrium $P_{s\text{-}av}$. Under attraction rule (hyp. 2a), the most striking effect is observed when comparing simulations assuming $\rho = 0$ *vs* $\rho > 0$ but self avoidance behavior is still observed in the population ($P_{s\text{-}av} \approx 27\%$) when recombination is free ($\rho = 0.5$). The evolution of disassortative mating behaviors is more impaired when assuming that preference alleles generate rejection (hyp. 2b): self-avoidance behavior completely disappears when preference and colour pattern loci are unlinked (*i.e.* when $\rho = 0.5$).



(a) (b)

Figure 5: **Influence of recombination between colour pattern and preference alleles on the distribution of mating behaviors at the population level, assuming (a) attraction or (b) rejection alleles at the preference locus (*recognition/trait*).** The proportion of individuals displaying self-acceptance ($P_{s\text{-}acc}$) (in purple) and self-avoidance ($P_{s\text{-}av}$) (in blue) obtained 2,000 time units after the introduction of preference alleles in both populations are shown for different values of recombination rate ρ between the preference locus M and the colour pattern locus P. The locus M controls for a specific recognition of colour pattern alleles inducing either (a) attraction (hyp.2a) or (b) rejection (hyp.2b). The three alleles at the locus P controlling colour pattern variations are initially introduced in even proportion $\frac{1}{3}$ in each population. After 10,000 time units under random mating the four alleles at locus M M_r, M_a, M_b and M_c are introduced respectively in proportion 0,99, $\frac{0.01}{3}$, $\frac{0.01}{3}$, $\frac{0.01}{3}$. Simulations are run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $\rho = 0$, $mig = 0.1$, $\delta_3 = 0$, $\delta = 0.1$ and $cost = 0.1$.

Under the attraction rule (hyp.2.a), two third of the haplotypes containing a given color pattern allele lead to self-avoidance (for instance $a - M_b$ and $a - M_c$ for colour pattern allele a). By contrast, under the rejection rule (hyp.2.b), only one out the three possible haplotypes leads to self-avoidance (for instance $a - M_a$ for colour pattern allele a). Moreover, the allele encoding for the rejection of a given colour pattern (*e.g.* M_a) is rarely linked with the rejected colour pattern allele (*e.g.* a) because mate choice limits crosses between an individual carrying a rejection allele on one hand and an individual displaying the rejected allele on the other hand. This limited linkage between rejecting and rejected alleles further impedes the formation of disassortative haplotypes (*e.g.* $a - M_a$) by recombination. Overall, genetic architectures enabling recombination between colour pattern and preference loci thus limit the evolution of haplotypes linking colour

pattern alleles with the corresponding mate choice allele leading to disassortative mating, when assuming a preference locus acting on the specific recognition of mating cues (Hyp.2).

In contrast, when assuming a *self-referencing* genetic architecture (Hyp.1), recombination increases the proportion of disassortative mating at equilibrium ($P_{s\text{-}av} \approx 99\%$) (Supp. figure S11). Selection generated by the genetic load associated to the colour pattern alleles a and b promotes their linkage with the disassortative *self-referencing* allele dis, while the genetic-load free allele c tends to be linked to the random mating allele r (Supp. figure S12). Because the allele dis reaches a high frequency in the population, as soon as recombination occurs, it generates a large number of recombinant haplotypes $c - dis$. Under the *self-referencing* hypothesis (hyp.1), recombination thus significantly increases the proportion of disassortative mating, reaching higher values than the one observed under the *recognition/trait* architecture (hyp.2). Recombination between the colour pattern locus P and the mate choice locus M therefore makes disassortative mating more likely to evolve assuming a *self-referencing* architecture.

Discussion

Genetic architecture of disassortative mating: theoretical predictions

Our model shows that without recombination between colour pattern (locus P) and preference alleles (locus M) disassortative mating is more likely to emerge when the genetic architecture is based on rejection for trait (Hyp. 2.b). In contrast, when recombination between the two loci does occur, the *self-referencing* architecture (Hyp.1) may facilitate the evolution of disassortative mating. The genetic basis of disassortative mating is largely unknown in natural populations. Assortative mating is better documented, for instance in *Heliconius* butterflies where it is generally associated with attraction towards a specific cue. The locus controlling preference for yellow *vs.* white in *H. cydno* maps close to the gene *aristaless*, whose variations in expression controls for the white/yellow switch in this species (Kronforst et al., 2006; Westerman et al., 2018). In *H. melpomene*, a major QTL associated with preference towards red was identified in crosses be-

tween individuals displaying a red pattern and individuals with a white pattern (Merrill et al., 2019). This QTL is also located close to the gene *optix* involved in the variation of red patterning in *H. melpomene*. Assortative mating in *Heliconius* thus seems to rely on alleles encoding preference for specific cues, linked to with loci involved in the variation of these cues. Contrastingly, our model suggests that the genetic architecture of disassortative mating might differ from those documented in species showing assortative mating behavior.

Some *recognition/trait* genotypes generate similar mate preferences to some *self-referencing* genotypes: for example, under the *rejection* rule, the genotype $a - M_a / a - M_a$ leads to the same mate preference as the genotype $a - dis / a - dis$ under the *self-referencing* genetic architecture. Introducing recombination in the *recognition/trait* architecture then enables the decoupling of the mating cue and of its corresponding preference alleles, thereby disrupting the self rejection behavior. Under the *recognition/trait* architecture, some haplotypes thus generates partial disassortative mating : for instance the haplotype $a - Mb$ under the *rejection* rule allows mating with individuals displaying the non-self phenotype c or the self phenotype a. This self-acceptation behavior may increase the reproductive success associated with these haplotypes. The persistence of these *rejection* rule haplotypes allowing both assortative and disassortative mating to prevent the fixation of strict self-rejection behavior in the population. Furthermore, under the *recognition/trait* architecture, our model distinguishes whether the specific recognition of the cue leads to *rejection* or *attraction*, and highlights that these two hypotheses lead to the evolution of different mate preferences: disassortative mating is more likely to emerge assuming the *rejection rule*. This rule indeed generates a greater number of self-rejecting haplotypes than the *attraction* rule, although recombination limits this effect.

The effect of dominance at the colour pattern locus has a variable impact on the evolution of disassortative mating depending on the genetic architecture of preference. Under both *recognition/trait* rules, the mate choice is based on the phenotype of the chosen individual, so that the dominance relationships at the colour pattern locus influences the evolution of disassortative mating. Nevertheless, under *self-referencing* the mate preference also depends on the phenotype

of the choosing individual, so that dominance at the colour pattern locus of both the choosing and chosen individuals determines the choice. Under *recognition/trait*, however, the mate preference does not depend on the phenotype of the choosing individual, but on the dominance relationships at the mate preference locus, allowing for different types of preference to emerge, including individuals reproducing with different phenotypes only, or individuals mating with either similar or different phenotypes.

Altogether, our theoretical model shows that the genetic basis of mate preferences has a strong impact on the evolution of disassortative mating at loci under heterozygote advantage. This emphasizes the need to characterize the genetic basis of mate preference empirically and the linkage disequilibrium with the locus controlling variation in the mating cues.

Evolution of disassortative mating results from interactions between dominance and deleterious mutations

Here, we confirm that the evolution of disassortative mating is promoted by the heterozygote advantage associated with alleles determining the mating cue. As mentioned below, the phenotype of the chosen individuals depends on the dominance relationships at the colour pattern locus. Our model highlights that a genetic load associated with the dominant alleles contributes more to disassortative mating than a genetic load associated with the most recessive haplotype. This theoretical prediction is in accordance with the few documented cases of polymorphism promoted by disassortative mating. In the polymorphic butterfly *Heliconius numata* for instance, the top dominant haplotype *bicolouratus* is associated with a strong genetic load (Jay et al., 2019). Similarly, in the white throated sparrow, the dominant *white* allele is also associated with a significant genetic load (Tuttle et al., 2016). Again, in the self-incompatibility locus of the *Brassicaceae*, dominant haplotypes carry a higher genetic load than recessive haplotypes (Llaurens et al., 2009). Disassortative mating is beneficial because it increases the number of heterozygous offspring with higher fitness. Once disassortative mating is established within a population, recessive deleteri-

ous mutations associated with the dominant haplotype become sheltered because the formation of homozygotes carrying two dominant alleles is strongly reduced, thereby limiting the opportunities for purging via recombination (Llaurens et al., 2009). Similarly, the model of Karlin and Feldman (1968) suggests that disassortative mating slows down the purge of deleterious alleles. Falk and Li (1969) proved that disassortative mate choice promotes polymorphism, and therefore limits the loss of alleles under negative selection. Disassortative mating might thus shelter deleterious mutations linked to dominant alleles, and thus reinforces heterozygote advantage. The sheltering of deleterious mutations is favoured by the interaction between two aspects of the genetic architecture: dominance at the mating cue locus and limited recombination. This is likely to happen in polymorphic traits involving chromosomal rearrangements, where recombination is limited. Many rearranged haplotypes are indeed associated with serious fitness reduction at homozygotes state (Faria et al., 2019), such as in the derived haplotypes of the supergene controlling controlling plumage and mate preferences in the white-throated sparrow (Thomas et al., 2008). The deleterious elements in the inverted segment can be due to an initial capture by the inversions (Kirkpatrick, 2010), but they could also accumulate through time, resulting in different series of deleterious mutations associated to inverted and non-inverted haplotypes (Berdan et al., 2019).

Here, we assume that mate choice relied purely on a single cue. Nevertheless, mate choice could be based on other cues, controlled by linked loci and enabling discrimination between homozygotes and heterozygotes, thereby further increasing the proportion of heterozygous offsprings with high fitness. We also modelled strict preferences regarding colour patterns, but choosiness might be less stringent in the wild, and may limit the evolution of disassortative mating. Depending on the cues and dominance relationships among haplotypes, different mate choice behaviors may also evolve, which might modulate the evolution of polymorphism within populations. Our model thus stresses the need to document dominance relationships among haplotypes segregating at polymorphic loci, as well as mate choice behavior and cues, to understand the evolutionary forces involved in the emergence of disassortative mating.

Conclusions

Inspired by a well-documented case of disassortative mating based on cues subject to natural selection, our model shows that balancing selection promoting local polymorphism and heterozygote advantage is likely to favor the evolution of disassortative mating preferences. We highlights that disassortative mating is more likely to emerge under rejection toward specific cues. Such rejection loci promote disassortative mating when they are in tight linkage with the locus controlling mating cue variations.

Acknowledgments

The authors would like to thank Charline Smadi and Emmanuelle Porcher for feedback on the modeling approach developed here. We also thank Thomas Aubier and Richard Merrill and the whole Heliconius group for stimulating discussion on mate choice evolution in our favorite butterflies. We are also grateful to Roger Butlin and Tom van Dooren for their thoughtful review of a previous version of this manuscript. We also thank Sylvain Gerber for his help in improving the English of this article. This work was supported by the Emergence program from Paris City Council to VL and the ANR grant SUPERGENE to MJ.

References

M. Arias, A. Meichanetzoglou, M. Elias, N. Rosser, D. L. De-Silva, B. Nay, and V. Llaurens. Variation in cyanogenic compounds concentration within a heliconius butterfly community: does mimicry explain everything? BMC evolutionary biology, 16(1):272, 2016.

E. L. Berdan, A. Blanckaert, R. K. Butlin, and C. Bank. Muller's ratchet and the

long-term fate of chromosomal inversions. bioRxiv, 2019. 10.1101/606012. URL https://www.biorxiv.org/content/early/2019/04/13/606012.

R. K. Butlin, P. M. Collins, and T. H. Day. The effect of larval density on an inversion polymorphism in the seaweed fly coelopa frigida. Heredity, 1984.

L. Casselton. Mate recognition in fungi. Heredity, 88(2):142, 2002.

M. Chouteau, M. Arias, and M. Joron. Warning signals are under positive frequency-dependent selection in nature. Proceedings of the national Academy of Sciences, 113(8):2164–2169, 2016.

M. Chouteau, V. Llaurens, F. Piron-Prunier, and M. Joron. Polymorphism at a mimicry supergene maintained by opposing frequency-dependent selection pressures. Proceedings of the National Academy of Sciences, 114(31):8325–8329, 2017.

T. H. Day and Butlin. Non-random mating in natural populations of the seaweed fly, coelopa frigida. Heredity, 1987.

M. de Cara, N. Barton, and M. Kirkpatrick. A model for the evolution of assortative mating. The American Naturalist, 171(5):580–596, 2008. 10.1086/587062. URL https://doi.org/10.1086/587062. PMID: 18419568.

C. T. Falk and C. C. Li. Negative assortative mating: Exact solution to a simple model. Genetics, 1969.

R. Faria, K. Johannesson, R. K. Butlin, and A. M. Westram. Evolving inversions. Trends in ecology & evolution, 2019.

S. J. Hiscock and S. M. McInnis. Pollen recognition and rejection during the sporophytic self-incompatibility response: Brassica and beyond. Trends in plant science, 8(12):606–613, 2003.

B. M. Horton, Y. Hu, C. L. Martin, B. P. Bunke, B. S. Matthews, I. T. Moore, J. W. Thomas, and D. L. Maney. Behavioral characterization of a white-throated sparrow homozygous for the zal2 m chromosomal rearrangement. Behavior genetics, 43(1):60–70, 2013.

A. M. Houtman and J. B. Falls. Negative assortative mating in the white-throated sparrow, zonotrichia albicollis: the role of mate choice and intra-sexual competition. Animal Behaviour, 48(2):377–383, 1994.

P. Jay, M. Chouteau, A. Whibley, H. Bastide, V. Llaurens, H. Parrinello, and M. Joron. Mutation accumulation in chromosomal inversions maintains wing pattern polymorphism in a butterfly. bioRxiv, 2019. 10.1101/736504. URL https://www.biorxiv.org/content/early/2019/08/15/736504.

Y. Jiang, D. I. Bolnick, and M. Kirkpatrick. Assortative mating in animals. The American Naturalist, 181(6):E125–E138, 2013. 10.1086/670160. URL https://doi.org/10.1086/670160. PMID: 23669548.

M. Joron and Y. Iwasa. The evolution of a müllerian mimic in a spatially distributed community. Journal of Theoretical Biology, 237(1):87–103, 2005.

M. Joron, I. R. Wynne, G. Lamas, and J. Mallet. Variable selection and the coexistence of multiple mimetic forms of the butterfly heliconius numata. Evolutionary Ecology, 13(7-8):721–754, 1999.

M. Joron, R. Papa, M. Beltrán, N. Chamberlain, J. Mavárez, S. Baxter, M. Abanto, E. Bermingham, S. J. Humphray, J. Rogers, et al. A conserved supergene locus controls colour pattern diversity in heliconius butterflies. PLoS biology, 4(10):e303, 2006.

M. Joron, L. Frezal, R. T. Jones, N. L. Chamberlain, S. F. Lee, C. R. Haag, A. Whibley, M. Becuwe, S. W. Baxter, L. Ferguson, et al. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature, 477(7363):203, 2011.

S. Karlin and M. W. Feldman. Further analysis of negative assortative mating. Genetics, 1968.

M. Kirkpatrick. How and why chromosome inversions evolve. PLoS biology, 8(9):e1000501, 2010.

M. Kirkpatrick and S. L. Nuismer. Sexual selection can constrain sympatric speciation.

Proceedings of the Royal Society of London. Series B: Biological Sciences, 271(1540):687–693, 2004.

M. Kopp, M. R. Servedio, T. C. Mendelson, R. J. Safran, R. L. Rodríguez, M. E. Hauber, E. C. Scordato, L. B. Symes, C. N. Balakrishnan, D. M. Zonana, et al. Mechanisms of assortative mating in speciation with gene flow: connecting theory and empirical research. The American Naturalist, 191(1):1–20, 2018.

M. R. Kronforst, L. G. Young, D. D. Kapan, C. McNeely, R. J. O'Neill, and L. E. Gilbert. Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. Proceedings of the National Academy of Sciences, 103(17):6575–6580, 2006.

Y. Le Poul, A. Whibley, M. Chouteau, F. Prunier, V. Llaurens, and M. Joron. Evolution of dominance mechanisms at a butterfly mimicry supergene. Nature Communications, 5:5644, 2014.

V. Llaurens, L. Gonthier, and S. Billiard. The sheltered genetic load linked to the s locus in plants: new insights from theoretical and empirical approaches in sporophytic self-incompatibility. Genetics, 183(3):1105–1118, 2009.

V. Llaurens, S. Billiard, and M. Joron. The effect of dominance on polymorphism in müllerian mimicry. Journal of theoretical biology, 337:101–110, 2013.

V. Llaurens, A. Whibley, and M. Joron. Genetic architecture and balancing selection: the life and death of differentiated variants. Molecular ecology, 26(9):2430–2448, 2017.

R. M. Merrill, K. K. Dasmahapatra, J. Davey, D. Dell'Aglio, J. Hanly, B. Huber, C. D. Jiggins, M. Joron, K. Kozak, V. Llaurens, et al. The diversification of heliconius butterflies: what have we learned in 150 years? Journal of Evolutionary Biology, 28(8):1417–1438, 2015.

R. M. Merrill, P. Rastas, S. H. Martin, M. C. Melo, S. Barker, J. Davey, W. O. McMillan, and C. D. Jiggins. Genetic dissection of assortative mating behavior. PLoS biology, 17(2):e2005902, 2019.

F. Müller. Ituna and thyridia; a remarkable case of mimicry in butterflies (transl. by ralph meldola from the original german article in kosmos, may 1879, 100). Transactions of the Entomological Society of London, 1879.

S. P. Otto, M. R. Servedio, and S. L. Nuismer. Frequency-dependent selection and the evolution of assortative mating. Genetics, 179(4):2091–2112, 2008.

D. J. Penn and W. K. Potts. The evolution of mating preferences and major histocompatibility complex genes. The American Naturalist, 153(2):145–164, 1999.

S. Piertney and M. Oliver. The evolutionary ecology of the major histocompatibility complex. Heredity, 96(1):7, 2006.

M. Schilthuizen, P. G. Craze, A. S. Cabanban, A. Davison, J. Stone, E. Gittenberger, and B. Scott. Sexual selection maintains whole-body chiral dimorphism in snails. Journal of evolutionary biology, 20(5):1941–1949, 2007.

T. N. Sherratt. Spatial mosaic formation through frequency-dependent selection in müllerian mimicry complexes. Journal of Theoretical Biology, 2006.

X. Thibert-Plante and S. Gavrilets. Evolution of mate choice and the so-called magic traits in ecological speciation. Ecology letters, 16, 06 2013. 10.1111/ele.12131.

J. W. Thomas, M. Cáceres, J. J. Lowman, C. B. Morehouse, M. E. Short, E. L. Baldwin, D. L. Maney, and C. L. Martin. The chromosomal polymorphism linked to variation in social behavior in the white-throated sparrow (zonotrichia albicollis) is a complex rearrangement and suppressor of recombination. Genetics, 179(3):1455–1468, 2008. ISSN 0016-6731. 10.1534/genetics.108.088229. URL `https://www.genetics.org/content/179/3/1455`.

H. B. Throneycroft. A cytogenetic study of the white-throated sparrow, zonotrichia albicollis (gmelin). Evolution, 29(4):611–621, 1975. ISSN 00143820, 15585646. URL `http://www.jstor.org/stable/2407072`.

T. Tregenza and N. Wedell. Genetic compatibility, mate choice and patterns of parentage: invited review. Molecular Ecology, 9(8):1013–1027, 2000.

E. M. Tuttle, A. O. Bergland, M. L. Korody, M. S. Brewer, D. J. Newhouse, P. Minx, M. Stager, A. Betuel, Z. A. Cheviron, W. C. Warren, et al. Divergence and functional degradation of a sex chromosome-like supergene. Current Biology, 26(3):344–350, 2016.

C. Wedekind and S. Füri. Body odour preferences in men and women: do they aim for specific mhc combinations or simply heterozygosity? Proceedings of the Royal Society of London. Series B: Biological Sciences, 264(1387):1471–1479, 1997.

C. Wedekind, T. Seebeck, F. Bettens, and A. J. Paepke. Mhc-dependent mate preferences in humans. Proceedings of the Royal Society of London. Series B: Biological Sciences, 260(1359): 245–249, 1995.

E. L. Westerman, N. W. VanKuren, D. Massardo, A. Tenger-Trolander, W. Zhang, R. I. Hill, M. Perry, E. Bayala, K. Barr, N. Chamberlain, et al. Aristaless controls butterfly wing color variation used in mimicry and mate choice. Current Biology, 28(21):3469–3474, 2018.

S. Wright. The distribution of self-sterility alleles in populations. Genetics, 24(4):538, 1939.

S1: mendelian coefficient

Here we define a function returning the Mendelian coefficient in a diploid model with two loci and a rate of recombination ρ. Let c, m and f be three genotypes. $coef(c, m, f, \rho)$ corresponds to the average proportion of offspring of genotype c resulting from a mating between a mother of genotype m and a father of genotype f given a recombination rate ρ.

We first define a function $coef_{haplotype}(h_c, p, \rho)$ which returns the proportion of haplotype h_c produced by a parent of genotype p given a recombination rate ρ.

We have $h_c = (P_c, M_c)$ where P_c is an allele at locus P and M_c is an allele at locus M and $p = (P^m, P^p, M^m, M^p)$ with P^m and P^p being the alleles at locus P and M^m and M^p being the alleles at locus M. P^m and M^m lies on the maternal chromosomes and P^p and M^p lies on the paternal chromosomes.

$$
\begin{aligned}
coef_{haplotype}(h_c, p, \rho) = & \mathbb{1}_{\{P^m = P^p\} \cap \{P_c = P^m\}} (\mathbb{1}_{\{M^m = M^p\} \cap \{M_c = M^m\}} + \frac{1}{2} \mathbb{1}_{\{M^m \neq M^p\} \cap \{\{M_c = M^m\} \cup \{M_c = M^p\}\}}) \\
& + \mathbb{1}_{\{P^m \neq P^p\} \cap \{\{P_c = P^m\} \cup \{P_c = P^p\}\}} (\frac{1}{2} \mathbb{1}_{\{M^m = M^p\} \cap \{M_c = M^m\}} \\
& + \mathbb{1}_{\{M^m \neq M^p\} \cap \{\{M_c = M^m\} \cup \{M_c = M^p\}\}} \\
& (\mathbb{1}_{\{\{P_c = P^m\} \cap \{M_c = M^m\}\} \cup \{\{P_c = P^p\} \cap \{M_c = M^p\}} \frac{1 - \rho}{2} \\
& + \mathbb{1}_{\{\{P_c = P^m\} \cap \{M_c = M^p\}\} \cup \{\{P_c = P^p\} \cap \{M_c = M^m\}} \frac{\rho}{2})).
\end{aligned}
$$

Let $c = (h^m, h^p)$ with h^m being the maternal haplotype and h^p being the paternal haplotype.

Then we define $coef$ as:

$$
coef(c, m, f, \rho) = coef_{haplotype}(h^m, m, \rho) coef_{haplotype}(h^p, p, \rho). \tag{13}
$$

Supplementary Figures

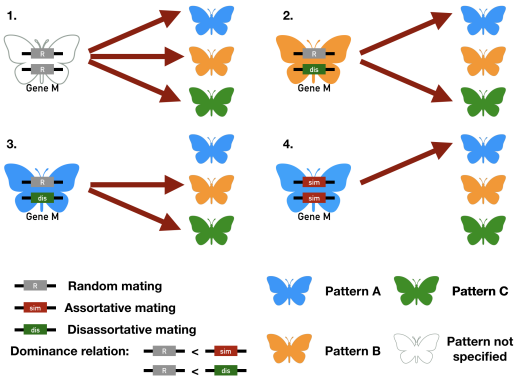


Figure S2: **Mate preferences expressed by the different genotypes at locus *M*, assuming *self-referencing* (Hyp.1)**. 1. Butterflies carrying two *r* alleles mate at random, independently from either their own colour pattern or the colour pattern displayed by mating partners. 2-3. Butterflies carrying a *dis* allele display disassortative mating behavior, and mate preferentially with individuals whose colour pattern differ from their own. 4. Butterflies carrying a *sim* allele display an assortative mating behavior and thus preferentially mate with individuals displaying the same colour pattern. Cases 1 and 4 therefore lead to *self-acceptance*, while cases 2 and 3 lead to *self-avoidance*.

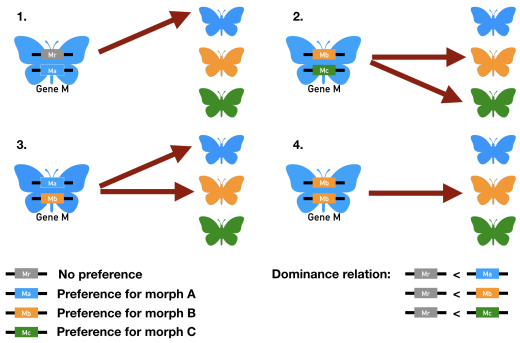


Figure S3: **Mate preferences expressed by the different genotypes at locus *M* assuming preference allele encoding for attraction of specific colour patterns (*recognition/trait*) (hyp.2.a)**. 1. A butterfly displaying phenotype A (in blue) carried one allele coding for specific attraction toward partner displaying phenotype A (in blue) and the allele coding for random mating at the locus M controlling the mate choice. This butterfly will mate preferentially with individuals displaying phenotype A, resulting in assortative mating. 2. A butterfly displaying phenotype A (in blue) carries one allele coding for specific attraction toward partner displaying phenotype B (in orange) and one allele coding for specific attraction toward partner displaying phenotype C (in green). This individual will preferentially mate with individuals displaying phenotype B and C, resulting in disassortative mating 3. A butterfly displaying phenotype A (in blue) carries one allele coding for specific attraction toward partner displaying phenotype A (in blue) and one allele coding for specific attraction toward partner displaying phenotype B (in orange). This individual will preferentially mate with individuals displaying phenotype A and B 4. A butterfly displaying phenotype A (in blue) carries two alleles coding for specific attraction toward partner displaying phenotype B (in orange). This individual will preferentially mate with individuals displaying phenotype B, resulting in disassortative mating. Cases 1 and 3 therefore lead to *self-acceptance*, while cases 2 and 4 lead to *self-avoidance*.

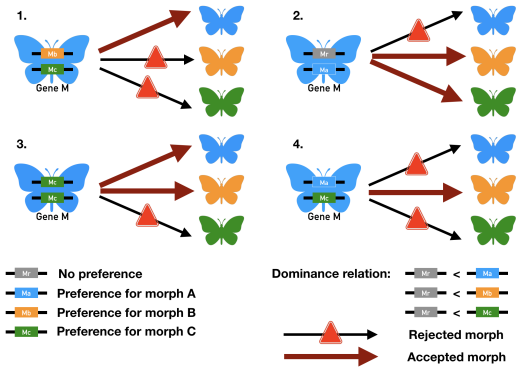


Figure S4: **Mate preferences expressed by the different genotypes at locus *M* preference allele encoding for rejection of specific colour patterns (*recognition/trait*) (hyp.2.a)**. 1. A butterfly displaying phenotype A (in blue) carried one allele coding for specific rejection toward partner displaying phenotype B (in orange) and one allele one allele coding for specific rejection toward partner displaying phenotype C (in orange). This butterfly will mate preferentially with individuals displaying phenotype A, resulting in assortative mating. 2. A butterfly displaying phenotype A (in blue) carried one allele coding for specific rejection toward partner displaying phenotype A (in orange) and one allele coding for random mating (in grey). This butterfly will mate preferentially with individuals displaying phenotypes B and C, resulting in disassortative mating. 3. A butterfly displaying phenotype A (in blue) carried two alleles coding for specific rejection toward partners displaying phenotype C (in green). This butterfly will mate preferentially with individuals displaying phenotypes A and B. 4. A butterfly displaying phenotype A (in blue) carried one allele coding for specific rejection toward partner displaying phenotype A (in blue) and one allele coding for specific rejection toward partner displaying phenotype C (in green). This butterfly will mate preferentially with individuals displaying phenotype B resulting in disassortative mating. Cases 1 and 3 therefore lead to *self-acceptance*, while cases 2 and 4 lead to *self-avoidance*.

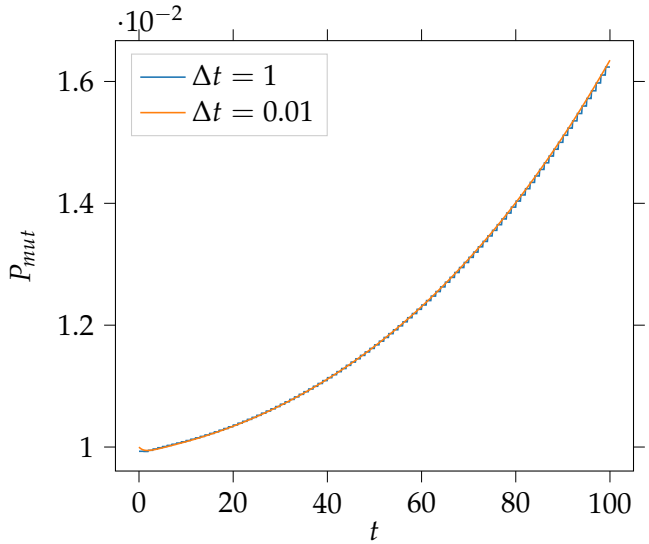


Figure S5: **Evolution of the proportion of mutant** *dis* **in the population after its introduction, using simulations with two different time units ($\Delta t = 1$ or $\Delta t = 0.01$) following Euler method, under the *self-referencing* hypothesis (Hyp.1). Both simulations give a similar dynamics of the proportion of the mutant P_mut, confirming that using discrete time simulations provide relevant estimations of the evolution of disassortative mating. Simulations are run during 100 time steps and assuming $\delta_1 = \delta_2 = 0.5$), $\delta_3 = 0$, $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $\rho = 0$, $mig = 0.1$, $\delta = 0.1$ and $cost = 0.1$.**

$\delta_1 = \delta_2$	δ_3	Population 1			Population 2		
		Proportion of morph A	Proportion of morph B	Proportion of morph C	Proportion of morph A	Proportion of morph B	Proportion of morph C
0,00	0,00	90 %	10 %	0 %	46 %	54 %	0 %
0,00	0,25	90 %	10 %	0 %	46 %	54 %	0 %
0,00	0,50	90 %	10 %	0 %	46 %	54 %	0 %
0,00	1,00	90 %	10 %	0 %	46 %	54 %	0 %
0,25	0,00	63 %	8 %	28 %	22 %	53 %	25 %
0,25	0,25	79 %	18 %	3 %	35 %	59 %	6 %
0,25	0,50	80 %	18 %	2 %	38 %	58 %	4 %
0,25	1,00	82 %	18 %	1 %	41 %	57 %	2 %
0,50	0,00	56 %	7 %	37 %	19 %	51 %	31 %
0,50	0,25	76 %	19 %	5 %	32 %	59 %	9 %
0,50	0,50	78 %	19 %	3 %	36 %	59 %	5 %
0,50	1,00	80 %	19 %	1 %	39 %	58 %	3 %
1,00	0,00	51 %	5 %	43 %	17 %	48 %	35 %
1,00	0,25	74 %	19 %	7 %	31 %	58 %	11 %
1,00	0,50	77 %	19 %	4 %	35 %	59 %	6 %
1,00	1,00	79 %	19 %	2 %	38 %	58 %	3 %

Figure S6: **Impact of linked genetic load on colour pattern polymorphism, assuming random mating.** The proportion of phenotypes A, B and C in the population 1 and 2 after 1000 time units depend on the different values of genetic load associated with the recessive allele c (δ_1), intermediate-dominant allele b (δ_2) and dominant allele c (δ_3). Simulation were run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $\rho = 0$, $mig = 0.1$, $\delta = 0.1$ and $cost = 0.1$.

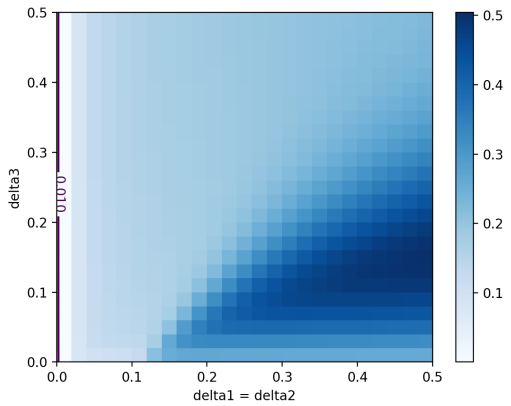
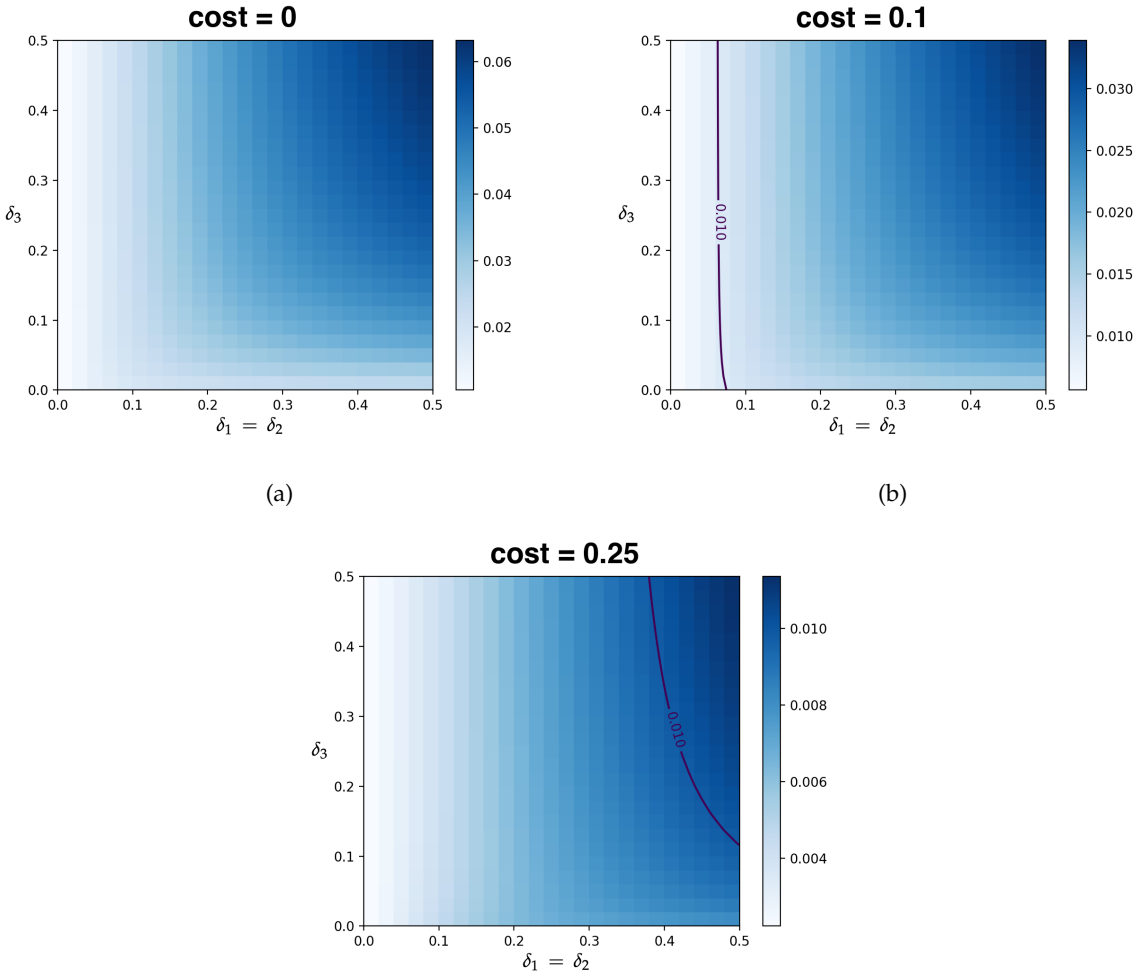


Figure S7: **Effect of the cost of choosiness** *cost* **on the invasion of the disassortative mutant** *dis***, under the** ***self-referencing*** **hypothesis (Hyp.1)**. Simulations are run assuming either low cost of choosiness $cost = 0.1$. The invasion of the disassortative mutant *dis* always depends on the strength of genetic load associated with the dominant alleles a and b ($\delta_1 = \delta_2$) on the x-axis and to the recessive allele c, δ_3, on the y-axis. Level of blue indicates the frequency of the disassortative mutant *dis*, inducing self-avoidance based on phenotype (hyp. 1), after 10,000 time units.The three alleles at the locus P controlling colour pattern variations were introduced in proportion $\frac{1}{3}$ in each population, and the initial frequency of the mutant was 0.01, shown by the vertical purple line, marking the limit of invasion by the mutant. Simulation were run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0$, $mig = 0.1$ and $\rho = 0$.



(a) (b) (c)

Figure S8: **Effect of the cost of choosiness** *cost* **on the invasion of the disassortative mutant** *dis*, **under the** ***self-referencing*** **hypothesis (Hyp.1)**. Simulations were run assuming either *(a)* no cost of choosiness $cost = 0$, *(b)* low cost of choosiness $cost = 0.1$ or *(c)* elevated cost of choosiness $cost = 0.25$. The invasion of the disassortative mutant *dis* always depends on the strength of genetic load associated with the dominant alleles a and b ($\delta_1 = \delta_2$) on the x-axis and to the recessive allele c, δ_3, on the y-axis. Level of blue indicates the frequency of the disassortative mutant *dis*, inducing self-avoidance based on phenotype (hyp. 1), after 100 time units.The three alleles at the locus P controlling colour pattern variations were introduced in proportion $\frac{1}{3}$ in each population, and the initial frequency of the mutant was 0.01, shown by the vertical purple line, marking the limit of invasion by the mutant. Simulation were run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0$, $mig = 0.1$ and $\rho = 0$.

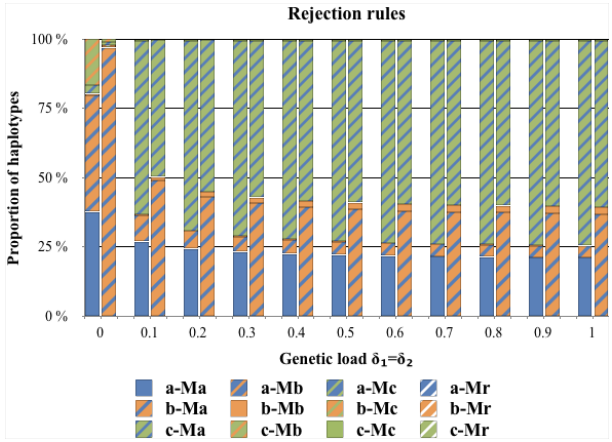


Figure S9: **Impact of the genetic load on haplotype diversity, assuming rejection alleles at the preference locus (Hyp. 2b), during the emergence of preference alleles**. The proportion of haplotypes obtained 200 time units after the introduction of preference alleles in both populations are shown for different values of genetic load associated with alleles a and b ($\delta_1 = \delta_2$). For each value of genetic load ($\delta_1 = \delta_2$) the first and second bars represented respectively the haplotypes proportion in the patch 1 and 2. The locus M controls for a specific recognition of colour pattern alleles inducing either (a) attraction (hyp.2a) or (b) rejection (hyp.2b). The three alleles at the locus P controlling colour pattern variations are initially introduced in even proportion $\frac{1}{3}$ in each population. After 10,000 time units under random mating the four alleles at locus M M_r, M_a, M_b and M_c are introduced respectively in proportion 0,99, $\frac{0.01}{3}$, $\frac{0.01}{3}$, $\frac{0.01}{3}$. Simulations are run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $\rho = 0$, $mig = 0.1$, $\delta_3 = 0$, $\delta = 0.1$ and $cost = 0.1$.

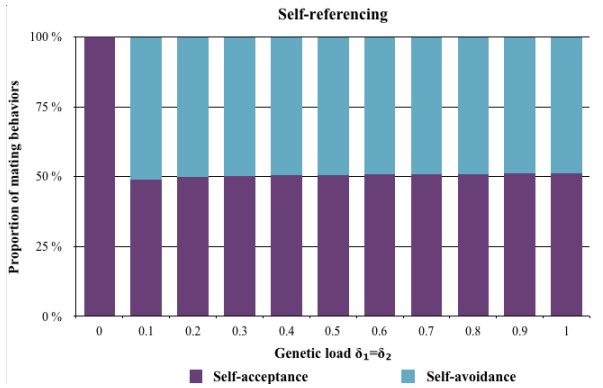


Figure S10: **Influence of the genetic load on the distribution of mating behaviors at the population level, assuming *self-referencing* rule (Hyp.1).** The proportion of individuals displaying self-acceptance $P_{s\text{-}acc}$ (in purple) and self-avoidance $P_{s\text{-}av}$ (in blue) obtained 10,000 time units after the introduction of the mutant allele *dis* in both populations are shown for different values of the level of genetic load of δ_1 and δ_2. The three alleles at the locus P controlling colour pattern variations were introduced in proportion $\frac{1}{3}$ in each population. Simulations are run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $\rho = 0$, $mig = 0.1$, $\delta_3 = 0$, $\delta = 0.1$ and $cost = 0.1$.

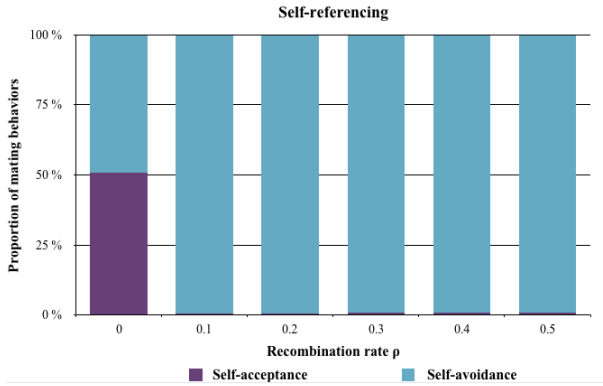


Figure S11: **Influence of recombination between colour pattern and preference alleles on the distribution of mating behaviors at the population level, assuming *self-referencing* rule (Hyp.1).** The proportion of individuals displaying self-acceptance $P_{s\text{-}acc}$ (in purple) and self-avoidance $P_{s\text{-}av}$ (in blue) obtained 10,000 time units after the introduction of the mutant allele *dis* in both populations are shown for different values of the recombination between colour pattern and preference alleles ρ. The three alleles at the locus P controlling colour pattern variations were introduced in proportion $\frac{1}{3}$ in each population. Simulations were run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $mig = 0.1$, $\delta_1 = 0.5$, $\delta_2 = 0.5$ $\delta_3 = 0$, $\delta = 0.1$ and $cost = 0.1$.

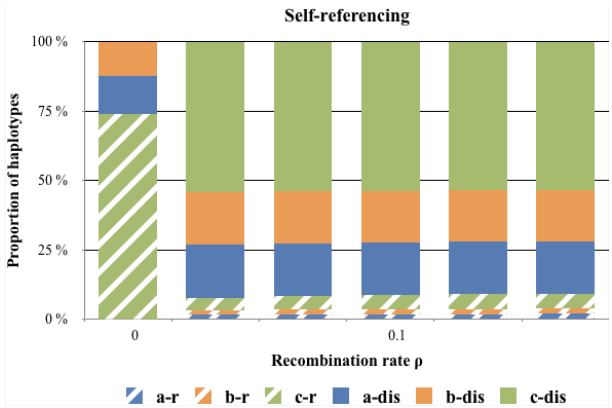


Figure S12: **Impact of recombination between colour pattern (locus *P*) and preference alleles (locus *M*) on haplotype diversity, assuming *self-referencing* preference alleles (Hyp.1).** The proportion of haplotypes for different values of recombination rate ρ obtained 10,000 time units after the introduction of the mutant *dis* are shown. The three alleles at the locus P controlling colour pattern variations were introduced in proportion $\frac{1}{3}$ in each population and the genetic architecture to describe the locus M corresponded to self-referencing (hyp.1). Simulations were run assuming $r = 2$, $K = 2000$, $N^0_{tot,1} = N^0_{tot,2} = 100$, $\lambda = 0.0002$, $\sigma = 0.5$, $d = 0.1$, $mig = 0.1$, $\delta_1 = 0.5$, $\delta_2 = 0.5$ $\delta_3 = 0$, $\delta = 0.1$ and $cost = 0.1$.