

# Conflict between heterozygote advantage and hybrid incompatibility in haplodiploids (and sex chromosomes)

Ana-Hermina Ghenu<sup>\*,1</sup>, Alexandre Blanckaert<sup>\*,1</sup>, Roger K. Butlin<sup>2,3</sup>, Jonna Kulmuni<sup>2,4</sup>, and Claudia Bank<sup>#,1,5</sup>

January 9, 2018

## Affiliations:

<sup>1</sup>Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, Oeiras, 2780-156, Portugal

<sup>2</sup>Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Alfred Benny Building, Sheffield S10 2TN, UK

<sup>3</sup>Department of Marine Sciences, University of Gothenburg, Guldhedsgatan 5a, 41320 Gothenburg, Sweden

<sup>4</sup>Centre of Excellence in Biological Interactions, Department of Biosciences, 00014 University of Helsinki, Finland

<sup>5</sup>Kavli Institute for Theoretical Physics, University of California at Santa Barbara, Santa Barbara, CA 93106-4030

## Keywords:

Hybridization, heterosis, epistasis, haplodiploidy, fitness landscape, speciation

## Correspondence:

Claudia Bank, Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, Oeiras, 2780-156, Portugal  
evodynamics@gmail.com

\*These authors contributed equally.

## Running title:

Heterosis *versus* hybrid breakdown

## Abstract

In many diploid species the sex chromosomes play a special role in mediating reproductive isolation. In haplodiploids, where females are diploid and males haploid, the whole genome behaves similarly to the X/Z chromosomes of diploids. Therefore, haplodiploid systems can serve as a model for the role of sex chromosomes in speciation and hybridization. A previously described population of Finnish *Formica* wood ants displays genome-wide signs of ploidy and sexually antagonistic selection resulting from hybridization. Here, hybrid females have increased survivorship but hybrid males are inviable. To understand how the unusual hybrid population may be maintained, we developed a mathematical model with hybrid incompatibility, female heterozygote advantage, recombination, and assortative mating.

The rugged fitness landscape resulting from the co-occurrence of heterozygote advantage and hybrid incompatibility results in a sexual conflict in haplodiploids, which is caused by the ploidy difference. Thus, whereas heterozygote advantage always promotes long-term polymorphism in diploids, we find various outcomes in haplodiploids in which the population stabilizes either in favor of males, females, or via maximizing the number of introgressed individuals. We discuss these outcomes with respect to the potential long-term fate of the Finnish wood ant population, and provide approximations for the extension of the model to multiple incompatibilities. Moreover, we highlight the general implications of our results for speciation and hybridization in haplodiploids versus diploids, and how the described fitness relationships could contribute to the outstanding role of sex chromosomes as hotspots of sexual antagonism and genes involved in speciation.

## Introduction

Haplodiploids are an emerging system for speciation genetics (Koevoets and Beukeboom, 2009; Kulmuni and Pamilo, 2014; Lohse and Ross, 2015; Knecht et al., 2017). Although  $\approx 20\%$  of animal species are haplodiploid (comprising most *Hymenopterans*, some arthropods, thrips and *Hemipterans*, and several clades of beetles and mites; Crozier and Pamilo, 1996; Evans et al., 2004; de la Filia et al., 2015), little evolutionary theory has been developed specifically for speciation in haplodiploids (Koevoets and Beukeboom, 2009). Under haplodiploidy with arrhenotoky (hereafter simply haplodiploidy; Suomalainen et al., 1987), males develop from the mother's unfertilized eggs and are haploid, whereas eggs fertilized by fathers result in diploid females. Since this mode of inheritance is, from a theoretical viewpoint, similar to that of the X/Z chromosome, most work on speciation of haplodiploids draws on the rich literature of sex chromosome evolution (Jablonka and Lamb, 1991; Presgraves, 2008; Johnson and Lachance, 2012; Lohse and Ross, 2015). An important similarity between haplodiploids and X/Z chromosomes is that recessive mutations in the haploid sex are exposed to selection, but they are masked in diploids. This is expected to lead to faster evolution in the sex chromosomes (Charlesworth et al., 1987) that may partly underlie the large-X effect (Presgraves, 2008). The large-X effect refers to the observation that the sex chromosomes seem to play a special role in speciation by acting as the strongest barrier for gene flow between hybridizing lineages across different species (Höllinger and Hermisson, 2017). Similarly, haplodiploid species have been suggested to acquire reproductive isolation earlier and speciate faster than diploid species (Lohse and Ross, 2015; Lima, 2014). Although the factors influencing haplodiploid and X/Z chromosome evolution are not expected to be exactly the same (e.g. movement of sexually antagonistic genes to the sex chromosomes, dosage compensation between the sex chromosomes and autosomes, and turnover of sex chromosomes cannot occur in haplodiploids; Abbott et al., 2017), by studying haplodiploid models we can both improve our understanding of how speciation happens in the large subgroup of the animal kingdom that is haplodiploid, and gain new insights into the role of X/Z chromosomes in speciation for diploid species.

Recent studies have shown that hybridization and resulting gene flow between diverging populations may be important players in the speciation process since signs of hybridization and introgression are being observed ubiquitously in natural populations (Mallet, 2005; Dieckmann and Doebeli, 1999; Schluter, 2009; Schluter and Conte, 2009; Seehausen et al., 2014). When a hybrid population is formed, various selective forces may act simultaneously to either increase or decrease hybrid fitness, which dictate the fate of the population. One commonly documented finding is hybrid incompatibility (Presgraves, 2008; Fraïsse et al., 2014; Chen et al., 2016), where combinations of alleles at different loci interact to confer poor

fitness when combined in a hybrid individual (Bateson, 1909; Dobzhansky, 1936; Muller, 1942; Orr, 1995). In a hybrid population, the existence of hybrid incompatibility reduces the mean population fitness. This deficit can be resolved either through reinforcement (evolution of increased premating isolation to avoid production of unfit hybrids; Servedio and Noor, 2003) or by purging (demographic swamping leading to extinction of one of the local populations/species or reinstatement of the ancestral allele combinations; Wolf et al., 2001). On the other hand, hybridization can transfer adaptive genetic variation from one lineage to another (Heliconius Genome Consortium, 2012; Song et al., 2011; Whitney et al., 2010) and may result in overall heterosis (also known as hybrid vigor): a higher fitness of hybrids as compared to their parents (Schwarz et al., 2005; Chen, 2013; Bernardes et al., 2017). Heterosis can stabilize polymorphisms by conferring a fitness advantage to hybrids and thereby favor the maintenance of hybridization either through the improved exploitation of novel ecological niches or the masking of recessive deleterious mutations. Therefore hybrid incompatibility acts to avert ongoing hybridization while heterosis favors the maintenance of hybrids.

One example of the simultaneous action of hybridization-averse and hybridization-favoring forces is found in a hybrid population of *Formica polyctena* and *F. aquilonia* wood ants in Finland (Kulmuni et al., 2010; Kulmuni and Pamilo, 2014; Beresford et al., 2017). Here, it has been reported that hybrid (haploid) males do not survive to adulthood, whereas (diploid) females have higher survivorship when they carry many introgressed alleles as heterozygotes (i.e., heterozygous for alleles originating from one of the parental species in a genomic background otherwise from the other parental species). Thus, a combination of hybrid incompatibility and heterosis seems to dictate the dynamics of the population in a ploidy-specific manner: hybrid haploid males suffer a fitness cost while diploid hybrid females can have a selective advantage over parental ones. Here, the differences in ploidy create an apparent sexual conflict (sensu Arnqvist and Rowe, 2005) between haploid males and diploid females, because their fitness landscapes (i.e., the complex relationship between genotypes and fitness created via hybrid incompatibility and heterozygote advantage) are different. This conflict is absent if the same rugged fitness landscape occurs in diploid autosomes.

When both hybridization-averse and hybridization-favoring forces are acting, the long-term resolution of a hybridizing population is difficult to foresee: will hybridization eventually result in either complete speciation or extinction of one of the populations involved? Alternatively, can it represent an equilibrium maintained stably on an evolutionary time scale? Furthermore, will the probability of these outcomes depend on ploidy? In other words, is one of these outcomes more probable when interacting genes are found on a “haplodiploid” X/Z chromosome than when they exist on a “diploid” autosome?

We here develop and analyze a population-genetic model of an isolated hybrid population in which both hybridization-averse and hybridization-favoring forces are acting, and we study the evolutionary outcomes in both haplodiploid and (fully) diploid genetic systems. The rich dynamics of the haplodiploid model can result in four possible evolutionary stable states depending on the strength of heterozygote advantage *versus* hybrid incompatibility, the strength of recombination, and the degree of assortative mating. This includes a case of symmetric coexistence (where all diversity is maintained) in which both alleles can be maintained despite the segregating hybrid incompatibility, and in which long-term hybridization is favored. We find that the dynamics differ between haplodiploid and diploid systems and that, unlike in previous models of sexual conflict in haplodiploid populations (Kraaijeveld, 2009; Albert and Otto, 2005), the conflict is not necessarily resolved in favor of the females. Indeed, a compromise may be reached at which the average fitness of females is decreased to rescue part of the fitness of males. Moreover, evaluation of the model using the data from the natural hybrid population suggests that, under the assumption of an equilibrium, the

Finnish ant population may represent an example of compromise between male costs and female benefits through asymmetric coexistence. We discuss our findings with respect to the long-term effects of hybridization, the potential for speciation in haplodiploid versus diploid species, and with respect to their relevance for X- or Z-linked alleles in diploid individuals.

## Materials and Methods

### The model

We model an isolated haplodiploid or diploid hybrid population with individuals from two founder populations  $P_+$  and  $P_-$ . Note that throughout the manuscript, we preferentially refer to (sub-)populations rather than species; in those instances in which we use the term ‘species’ it is in order to emphasize that the two populations have diverged sufficiently for (potentially strong) hybrid incompatibility to exist. We assume discrete generations and consider two loci, **A** and **B**. Each locus has two alleles, the ‘+’ allele ( $A_+$  or  $B_+$ ) inherited from population  $P_+$  and the ‘-’ allele ( $A_-$  or  $B_-$ ) inherited from population  $P_-$ . We refer to ‘hybrids’ as individuals that carry two alleles from each of the two parental populations and cannot be assigned to either parental background. We refer to ‘introgressed’ individuals as those genotypes for which three of the four alleles are from the same parental population; these genotypes are identical to those produced by hybridization followed by backcrossing. We ignore new or recurrent mutation and genetic drift. Thus, male and female populations are of effectively infinite size; selection modifies the relative abundance of the different haplotypes/genotypes but not the number of individuals (soft selection). The life cycle is as follows (Fig. 1; see also Table 1 for a list of model parameters); consistent with the recursions defined below, we begin the life cycle at the adult stage:

1. mating, either randomly or via genotype matching with assortment strength  $\alpha$  as detailed below;
2. recombination (in diploid individuals) at rate  $\rho$ ;
3. viability (or survival) selection, where heterosis is modeled as a heterozygote advantage,  $\sigma$ , and hybrid incompatibility is modeled as a fully recessive negative epistasis,  $\gamma_1$  and  $\gamma_2$  (further details are provided below and in Figure 2).

### Viability selection

The fitness landscape described here (Fig. 2) is inspired by the situation observed in Finnish *Formica* ants (Kulmuni et al., 2010; Kulmuni and Pamilo, 2014; Beresford et al., 2017). There, the authors discovered heterosis in the diploid females but recessive incompatibilities expressed in the haploid males. This creates a situation in which the same alleles that are favored in heterozygous females are selected against in hybrid haploid males and homozygous hybrid females. In the haplodiploid genetic system, males possess only one copy of each locus so they cannot be heterozygous and, therefore, cannot experience heterozygote advantage (Fig. 2(b)). Therefore, the fitness landscape with heterozygote advantage and recessive hybrid incompatibility expresses itself as an apparent sexual conflict when sexes differ in ploidy, as in haplodiploids or for X/Z chromosomes.

In our model, selection for heterozygous individuals is multiplicative with respect to the number of heterozygous loci: introgressed individuals with one heterozygous locus have fitness  $1 + \sigma$ , whereas diploid hybrid individuals that are heterozygous at both loci have

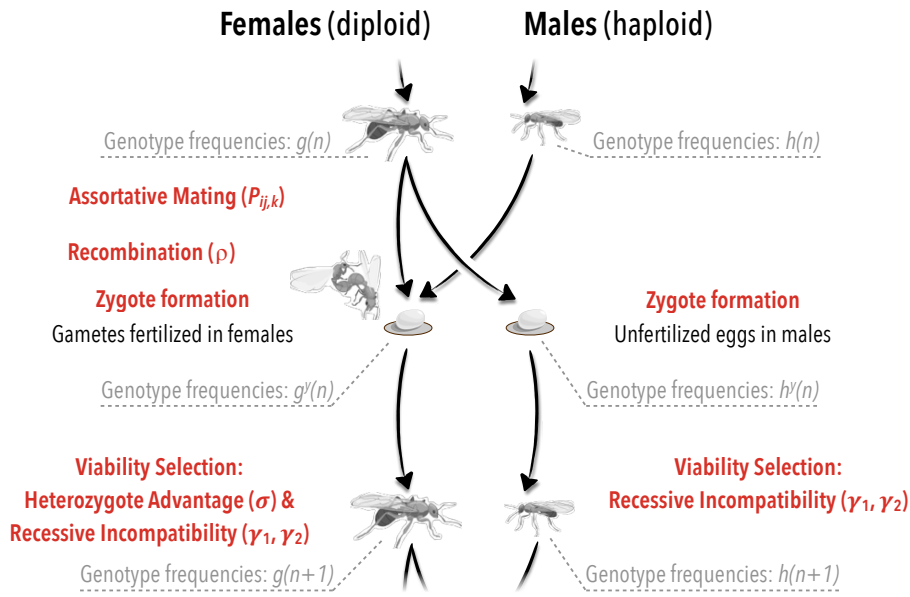


Figure 1: Illustration of the haplodiploid life cycle and its parameterization

Table 1: List of model parameters.

Symbol	Parameter	Limits
$\sigma, \omega$	<b>Strength of heterozygote advantage</b> , resulting in fitness $\omega = (1 + \sigma)$ or $\omega^2 = (1 + \sigma)^2$ of introgressed or double heterozygous diploid hybrids, respectively.	$\omega - 1 = \sigma > 0$
$\gamma_1, \gamma_2$	<b>Strength of fully recessive negative epistasis</b> , resulting in fitness $(1 - \gamma_1)$ for $A_+B_-$ homozygous diploid hybrids and $A_+B_-$ hybrid haploid males, and $(1 - \gamma_2)$ for $A_-B_+$ homozygous diploid hybrids and $A_-B_+$ hybrid haploid males.	$0 \leq \gamma_1, \gamma_2 \leq 1$
$\rho$	<b>Recombination rate</b> between locus <b>A</b> and <b>B</b> .	$0 \leq \rho \leq 0.5$
$\alpha$	<b>Strength of assortment</b> via genotype matching, where $\alpha = 0$ represents random mating, $\alpha > 0$ represents assortative mating among conspecifics, and $\alpha < 0$ represents assortative mating between heterospecifics.	$-1 \leq \alpha \leq 1$

survivorship  $(1 + \sigma)^2$  (Fig. 2(a)). Finally, the recessive epistatic incompatibility parameter  $\gamma_1$  acts on individuals homozygous or haploid for the  $A_+B_-$  haplotype, and  $\gamma_2$  acts on individuals homozygous or haploid for the  $A_-B_+$  haplotype (without loss of generality, we assume  $\gamma_1 \geq \gamma_2$ ). Thus, epistasis in this model can be asymmetric, reflecting, for example, two Dobzhansky-Muller incompatibilities of different strength that have accumulated at a negligible recombination distance between the same chromosome pairs. Note that when  $\gamma_1 = \gamma_2 = 1$ , haploid hybrid males and homozygous hybrid zygotes are produced but do not survive to adulthood and that the classical case of a single Dobzhansky-Muller incompatibility is recovered when  $\gamma_2 = 0$ .

### Assortative mating

Prezygotic isolation via assortative mating is an important mechanism that could mediate the detrimental effects to the population caused by the co-occurrence of heterozygote advantage and epistasis modeled here. In the Finnish wood ant population that inspired our model (Kulmuni and Pamilo, 2014), almost all egg-laying queens collected had been inseminated by males of the same genetic group, indicating that prezygotic isolation barriers are likely operating to result in assortative mating. In this case, assortative mating could arise via choosiness of mating partners, via genotype-dependent development times, or via other post-mating prezygotic mechanisms. We implemented assortative mating via genotype matching (reviewed in Kopp et al. (2017)), where the proportion of matings depends on the genetic distance between two mating partners (and their respective frequencies in the population). We define the genetic distance between the genotypes of a mating pair as the average Hamming distance, i.e. the number of differences between 2 aligned sequences of characters, between all possible pairs of haplotypes with one partner from each sex. We use quadratic assortment (e.g., De Cara et al., 2008), which results in assortative mating without costs of choosiness but with sexual selection. The mating probability of a pair of male and female genotypes,  $P_{ij,k}$  depends on the genetic distance between the two mates, the choosiness of the female, and the abundance of the different haplotype and genotypes as detailed below.

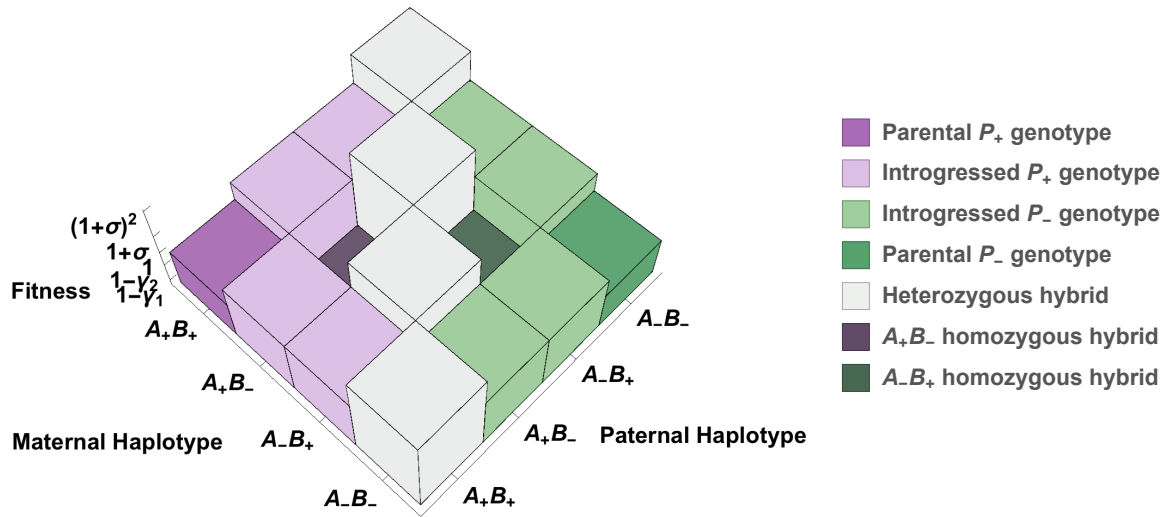
### Mathematical modeling and analysis

In a given generation  $n$ , the frequencies of the male and female adults are given by  $h_k(n)$  and  $g_{ij}(n)$ , respectively, with  $i$  and  $k$  indicating the haplotype received maternally and  $j$  the one of paternal origin. Without loss of generality, we assign index  $i = 1$  to haplotype  $A_+B_+$ , index  $i = 2$  to haplotype  $A_+B_-$ ,  $i = 3$  to haplotype  $A_-B_+$  and,  $i = 4$  to  $A_-B_-$ . Below, we describe the modeled life cycle (illustrated in Fig. S1) which determines how frequencies change from one generation to the next.

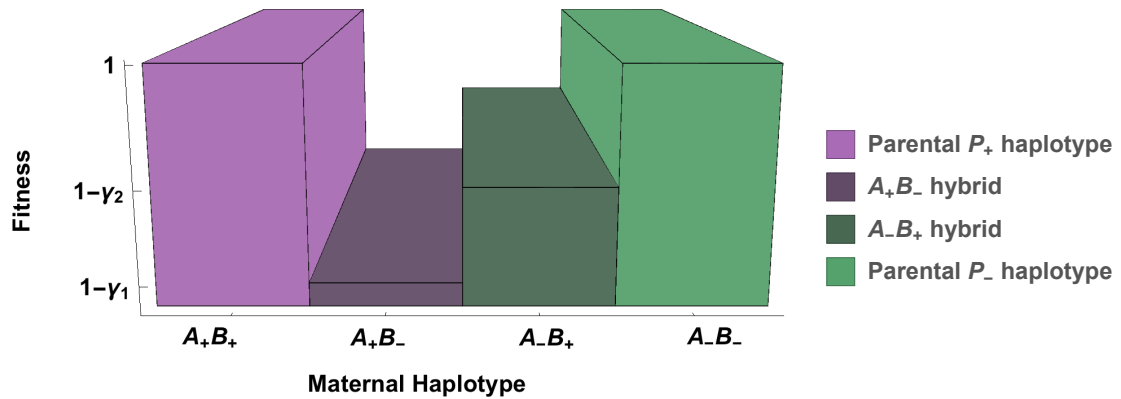
1. As detailed in figure 1 the first step of the life cycle is the mating between two individuals. The mating probability between an  $ij$  female and a  $k$  male is given by:

$$P_{ij,k}(n) = \frac{(1 - \alpha \frac{d_{i,k} + d_{j,k}}{2}) g_{ij}(n) h_k(n)}{\sum_i \sum_j \sum_k (1 - \alpha \frac{d_{i,k} + d_{j,k}}{2}) g_{ij}(n) h_k(n)} \quad (1)$$

with  $d_{i,k}$  the Hamming distance between two haplotypes. Note that for  $\alpha = 0$ , this simplifies to random mating and thus becomes equivalent to the dynamics described in Supplementary material (S7).



(a) Fitness Landscape for Diploid Individuals



(b) Fitness Landscape for Haploid Males

Figure 2: Three-dimensional fitness landscapes for the (a) diploid and (b) haploid genotypes. Panel a) corresponds to females in the haplodiploid model and all individuals in the diploid model. Individuals heterozygous at both loci (heterozygous hybrids) reside on a high fitness ridge (in white), whereas individuals homozygous at both loci (homozygous hybrids) suffer from reduced fitness due to negative epistasis. Panel b) shows the fitness landscape for haploid individuals (i.e. males) in the haplodiploid model. This landscape is identical to a transect from Panel a) for genotypes homozygous at both loci.

2. The next step is the formation of the zygote. Recombination happens only in females. We denote the frequency of newly born females as  $g_{ik}^y(n+1)$ .

$$\begin{cases} g_{ik}^y(n+1) = \frac{1}{2} \sum_{j=1}^4 (P_{ij,k}(n) + P_{ji,k}(n)) - \frac{\rho}{2} \Delta_k(n) & \text{if } i \in \{1, 4\} \\ g_{ik}^y(n+1) = \frac{1}{2} \sum_{j=1}^4 (P_{ij,k}(n) + P_{ji,k}(n)) + \frac{\rho}{2} \Delta_k(n) & \text{if } i \in \{2, 3\} \end{cases} \quad (2)$$

with  $\Delta_k(n) = P_{14,k}(n) + P_{41,k}(n) - P_{23,k}(n) - P_{32,k}(n)$ .

Males are composed from unfertilized female gametes, which have undergone recombination. The frequencies of newborn males are given by  $h_k^y(n)$ :

$$\begin{cases} h_k^y(n_y) = \frac{1}{2} \sum_{j=1}^4 (g_{kj}(n) + g_{jk}(n)) - \frac{\rho}{2} \tau(n) & \text{if } k \in \{1, 4\} \\ h_k^y(n_y) = \frac{1}{2} \sum_{j=1}^4 (g_{kj}(n) + g_{jk}(n)) + \frac{\rho}{2} \tau(n) & \text{if } k \in \{2, 3\} \end{cases} \quad (3)$$

with  $\tau(n) = g_{14}(n) + g_{41}(n) - g_{23}(n) - g_{32}(n)$ .

3. Individuals of both sexes are under viability selection. The frequencies of male and female adults of the next generations are given by

$$h_k(n+1) = \frac{w_k^m h_k^y(n)}{\sum_{k=1}^4 w_k^m h_k^y(n)} \quad (4)$$

with  $w_i^m$  the fitness of haplotype  $i$  in males and:

$$g_{ij}(n+1) = \frac{w_{ij}^f g_{ij}^y(n)}{\sum_{i=1}^4 \sum_{j=1}^4 w_{ij}^f g_{ij}^y(n)} \quad (5)$$

where  $w_{ij}^f$  denotes the fitness of the  $ij$  genotype. Note that there are no parental effects:  $w_{ij}^f = w_{ji}^f$ ; we maintain the distinction only for modeling convenience.

The complete recursion for females is obtained by substituting  $g_{ij}^y(n)$  by its expression given in (2) in (5) and  $P_{ij,k}(n)$  by (1). The complete recursion for males is given by substituting  $h_i^y$  by its expression given in (3) in (4). For  $\alpha = 0$ , the detailed recursion is given in Supplement (S7). Note that we use a different point of the life cycle (the gamete frequencies) as this is more easily tractable due to the reduced number of variables.

The diploid model can be obtained by applying equations (2) and (5) to males as well, with the corresponding relevant substitutions.

For the analysis, we focus on the equilibrium of the system defined by:

$$\forall \{i, j, k\} \in \{1, 2, 3, 4\}^3, g_{ij}(n+1) = g_{ij}(n) \text{ and } h_k(n+1) = h_k(n). \quad (6)$$

These equilibria can either be obtained by solving the system of equations presented above numerically, or by focusing on some of the known and potentially biological relevant equilibria, like fixation of a given haplotype. The stability of the equilibria is then obtained by computing the Eigenvalues of the Jacobian matrix at the focal equilibrium. If the absolute value of all Eigenvalues are below 1, the equilibrium is locally stable. For a more detailed explanation, see Otto and Day (2007, Chap. 7). We use this method to derive necessary and sufficient conditions for the existence and stability of the different evolutionary outcomes.

## Simulations

Derivations, simulations, and data fitting were performed in *Mathematica* (v 10.4.1.0; Wolfram Research, Inc., 2016). To enable complete reproducibility of the results, we provide an Online Supplement that documents all steps of the analysis as well as the code used for simulations and figures. Equilibrium genotype frequencies were obtained numerically when possible, or based on simulations until the differences between genotype frequencies of two consecutive generations were smaller than  $10^{-8}$  (or stopped after  $10^5$  generations without convergence).

## Fitting the model to a natural ant population

To compare our model with data from the natural, hybridizing Finnish ant population, we estimated the different genotype frequencies of parental *F. polyctena*-like and *F. aquilonia*-like individuals from the data. Assuming that the natural population is at equilibrium, we fit the data (Table S2) to the model by calculating the sum of squared differences between the observed data and predicted equilibrium frequencies. Complete details of data estimation and model fitting are given in the Supplementary Methods and Supplementary Results.

## Results

In this section, we describe the dynamics of a hybrid population under our model, with a particular focus on quantifying the differences between the haplodiploid and the diploid model. Two parameter domains are of particular interest:

1. The case of free recombination and strong epistasis (i.e., large  $\gamma_1, \gamma_2$ ) most likely resembles that of the natural ant hybrid population that inspired the model. Here, the hybrid incompatibility loci are located on different chromosomes, and epistasis is strong enough to erase a large fraction of male zygotes during development.
2. The case of low recombination is most relevant for the effects of a fitness landscape with epistasis (i.e., a “rugged” landscape) in X or Z chromosomes. Here, epistasis could arise, for example, through interactions between regulatory regions and their respective genes.

## Evolutionary scenarios

Below, we describe four different types of evolutionary stable states (i.e., equilibrium scenarios) of the model, which represent long-term solutions to the opposing selective pressures of the hybridization-averse force of recessive negative epistasis and the hybridization-favoring heterozygote advantage. The population will attain these equilibria if no further pre- or post-zygotic barriers or other functional mutations appear. Next, we provide various necessary and sufficient analytical conditions for these scenarios. Figure 3 illustrates the potential equilibria by means of phase diagrams.

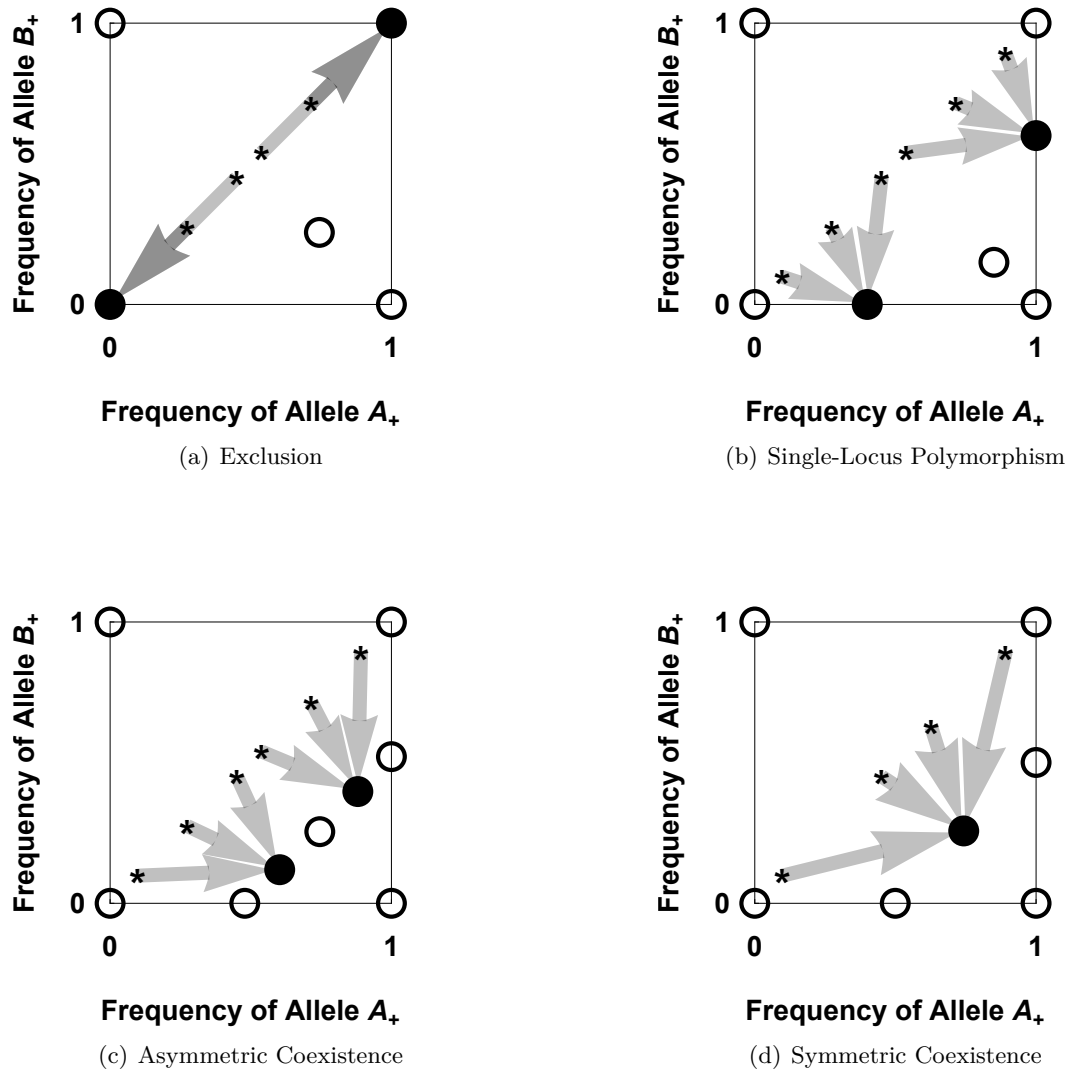


Figure 3: Phase-plane diagrams illustrating possible evolutionary scenarios in the haplodiploid model. The filled black dots show locally stable equilibria and the empty dots show unstable ones. The gray arrows show the basin of attraction starting from secondary contact scenarios (black asterisks on the line at  $p_{B_+} = p_{A_+}$ ). Panel (a) illustrates exclusion: There are 2 external locally stable equilibria, each corresponding to the fixation of a parental population haplotype. (Here,  $\sigma = 0.02$ ,  $\gamma_1 = 0.9$ ,  $\gamma_2 = 0.11$ ,  $\rho = 0.5$ , and  $\alpha = 0$ .) Panel (b) represents a single-locus polymorphism. Only one locus is polymorphic, leading to the maintenance of the weaker of the two incompatibilities (the  $A_-B_+$  interaction). (Here,  $\sigma = 0.009$ ,  $\gamma_1 = 0.11$ ,  $\gamma_2 = 0.002$ ,  $\rho = 0.5$ , and  $\alpha = 0$ .) Panel (c) corresponds to asymmetric coexistence. Two internal equilibria are locally stable, with one allele close to fixation. This scenario minimize the expression of the strongest interaction  $A_+B_-$ . (Here,  $\sigma = 0.03$ ,  $\gamma_1 = 0.11$ ,  $\gamma_2 = 0.0013$ ,  $\rho = 0.5$ , and  $\alpha = 0$ .) Panel (d) shows symmetric coexistence. Frequencies of alleles  $A_-$  and  $B_-$  are symmetric around 0.5, with  $p_{B_+} = 1 - p_{A_+}$ . This scenario maximizes the formation of female heterozygous hybrids. (Here,  $\sigma = 0.09$ ,  $\gamma_1 = 0.3$ ,  $\gamma_2 = 10^{-4}$ ,  $\rho = 0.5$ , and  $\alpha = 0$ .)

## Exclusion

The *exclusion* scenario corresponds to the hybrid population becoming identical to one of the two parental populations, either  $P_+$  or  $P_-$ , and the other parental population being therefore excluded. It occurs when both alleles from one of the founder subpopulations are purged, leading to a monomorphic stable state of the population (Fig. 3(a)). In this case, the initial frequency of  $A_+B_+$  versus  $A_-B_-$  individuals mainly determines the outcome (i.e., the population is swamped by the majority subpopulation). As a rule of thumb, this outcome is observed when recombination is frequent and when the hybridization-averse force of negative epistasis is strong as compared with the hybridization-favoring heterozygote advantage ( $\gamma_1, \gamma_2 \gg \sigma$ ).

With regard to the apparent sexual/ploidy conflict in the haplodiploid model, exclusion can be interpreted as a victory of the haploid males because all polymorphism is lost and no low-fitness hybrid males are produced. Conversely, since all polymorphism is lost, diploid females “lose” in this case and neither high-fitness introgressed (i.e., those individuals carrying only one ‘foreign’ allele) nor highest-fitness heterozygous hybrid females are produced. As discussed below, exclusion is never a possible outcome in the diploid model, in which there are no differences in ploidy.

## Single-locus polymorphism

A *single-locus polymorphism* occurs when one allele is purged from the population but the other locus remains polymorphic at equilibrium (Fig. 3(b)). Because this is possible for either of the two loci, two such equilibria exist simultaneously, which are reached depending on the initial haplotype frequencies. This outcome is observed when recombination is frequent, epistasis is asymmetric ( $\gamma_1 \neq \gamma_2$ ), and heterozygote advantage is small ( $\gamma_1 \gg \sigma$ ). Like asymmetric coexistence below, this case represents a compromise between the hybridization-averse and hybridization-favoring forces of negative epistasis and heterozygote advantage, and is reached by maximizing the number of introgressed individuals of one founder subpopulation.

In the haplodiploid model, this scenario can be seen as a haploid-dominated compromise. Since one locus is fixed, one epistatic interaction has disappeared and few low-fitness hybrid males are produced. In females, high-fitness introgressed female frequencies are maximized but, since one locus is fixed, the highest-fitness heterozygous hybrid female genotypes are no longer available.

The single-locus polymorphism is never stable in the diploid model, i.e., when the ploidy difference is removed from the model. In a diploid population that resides transiently at single-locus polymorphism, a rare mutant at the second locus will always begin as heterozygote and therefore reap the advantage of being a heterozygote hybrid long before it suffers the epistatic cost of being a homozygote hybrid.

## Asymmetric coexistence

“*Asymmetric*” *coexistence* occurs when all four haplotypes remain in the population and the frequency of introgressed individuals of one founder subpopulation is maximized (Fig. 3(c)). Because this can be achieved in two ways, two possible equilibria reside off the diagonal line  $p_B = 1 - p_A$  (where  $p_A$  and  $p_B$  denote the allele frequencies of the ‘-’ allele at the respective locus), and the initial contribution of different haplotypes determines which equilibrium will be attained. Like the single-locus polymorphism, this equilibrium represents a compromise between hybridization-averse and hybridization-favoring forces that is reached by maximizing the number of introgressed individuals. Our simulations demonstrate that

this scenario is rarely present in haplodiploids, and it generally involves asymmetric epistasis and intermediate-strength heterozygote advantage.

In the haplodiploid model, asymmetric coexistence can be seen as a compromise that is dominated by the diploids. Unlike in the single-locus polymorphism scenario, both loci are polymorphic and some double-heterozygous hybrid females are produced. But, unlike the symmetric coexistence scenario described below, females are not victorious over males because such high-fitness hybrid females are produced only at low frequencies.

### Symmetric coexistence

*Symmetric coexistence* occurs when a locally stable equilibrium exists on the diagonal  $p_B = 1 - p_A$ , such that the number of heterozygous hybrids is maximized (Fig. 3(d)). Our notion of “symmetric” refers to the total fraction of alleles from the  $P_+$  and  $P_-$  founder populations segregating at equilibrium, which is equal in this case. Here, prolonged hybridization is a mutual best-case scenario for both populations. This equilibrium is most likely when recombination is weak or when the hybridization-favoring force of heterozygote advantage is strong as compared with the hybridization-averse negative epistasis ( $\sigma \geq \gamma_1, \gamma_2$ ). In the haplodiploid model, symmetric coexistence represents a victory for the diploids, because they maximize their own fitness without regard to the production of unfit hybrid haploids.

The four evolutionary stable states described above usually result in either a single, globally stable equilibrium (in the case of symmetric coexistence) or a bistable system, in which two locally stable equilibria exist. In rare cases and close to bifurcation points, we observe cases of tristability, which are further described in Figure S2.

### Stability analysis of the model

Although the model dynamics are too complex to derive general analytical solutions, we were able to perform stability analyses for specific cases, which yield information about the general behavior of the model. In the following, our use of ‘>’ and ‘<’ does not necessarily imply strict inequalities; we merely did not explicitly study the limiting cases. For ease of notation, we refer to heterozygote advantage in terms of  $\omega$  below; recall that  $\omega = 1 + \sigma$ .

### Conditions for symmetric coexistence when epistasis is lethal

We begin by describing the equilibrium structure when epistasis is lethal, i.e.  $\gamma_1 = \gamma_2 = 1$ ; this case may resemble that in the natural ant population, in which most hybrid males do not survive to reproduce. For the haplodiploid model, we obtain a full analytic solution of the identity, existence and stability of equilibria. Here, only two outcomes are possible: symmetric coexistence and exclusion (Fig. 4(a)). As necessary and sufficient criterion for exclusion, we obtain

$$\rho > \frac{\omega^2 - 1}{\omega^2}. \quad (7)$$

Thus, exclusion is only possible if heterozygote advantage is not too strong, and if recombination is breaking up gametes sufficiently often to significantly harm the haploid males.

For the diploid model, we can show that no boundary equilibrium is ever stable; asymmetric and symmetric coexistence are the only two possible outcomes. Although it was not possible to perform a stability analysis on the internal equilibria, we were able to propose a

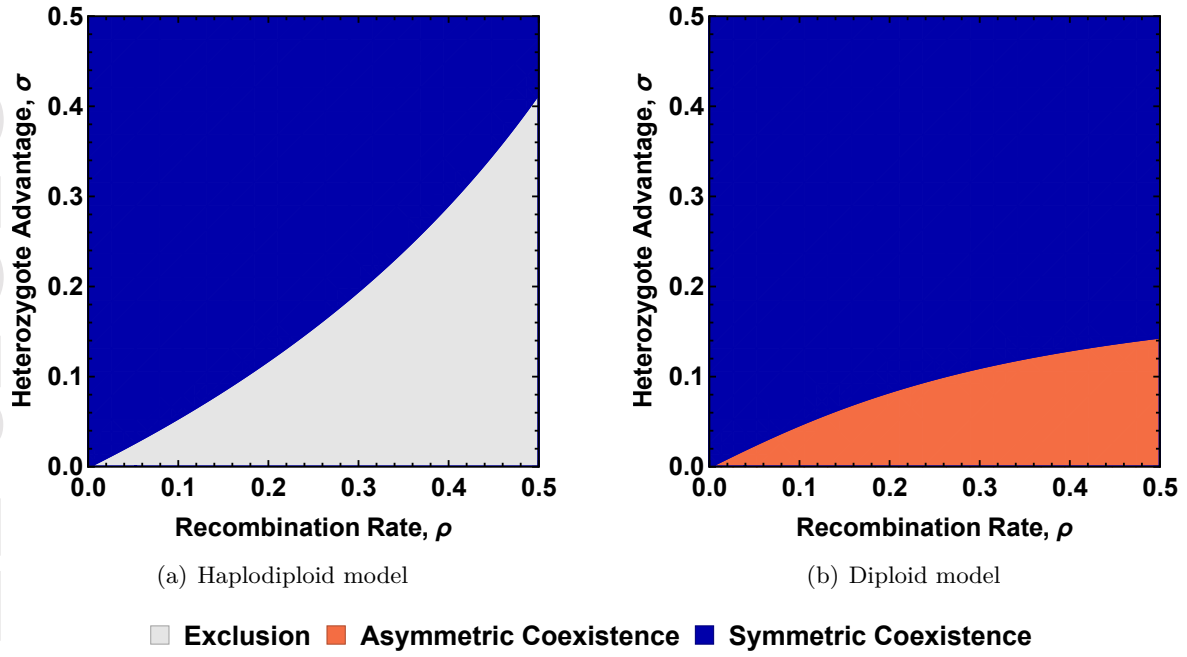


Figure 4: Symmetric coexistence can be locally stable if the heterozygote advantage,  $\sigma$ , is strong enough to compensate for recombination breaking up the parental haplotypes. Here we assume that epistasis is symmetric and lethal ( $\gamma_1 = \gamma_2 = 1$ ). Panel (a) is an illustration of the condition for haplodiploids given in equation (7) and panel (b) of equation (8) for diploids.

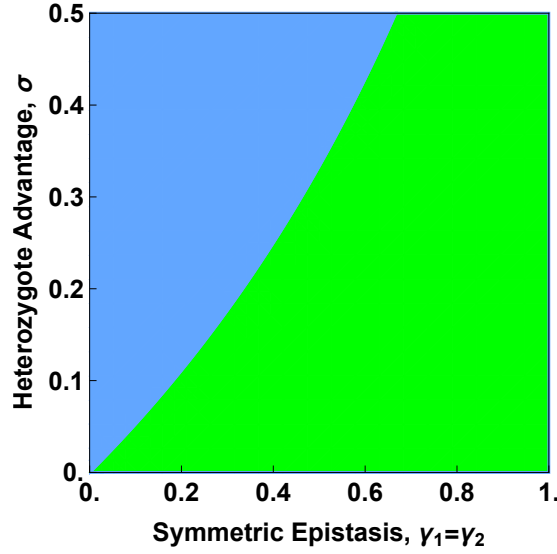
condition for asymmetric coexistence, which has been evaluated numerically:

$$\rho > \frac{(\omega^2 - 1)(2\omega^4 - 6\omega^3 + \omega^2 + 6\omega - 2)}{\omega^2(2\omega^2 - 4\omega + 1)(2\omega^2 - 3)} + 2\sqrt{\frac{(\omega - 1)^5(\omega + 1)^2(\omega^3 - \omega^2 - 3\omega + 1)}{\omega^4(2\omega^2 - 4\omega + 1)^2(2\omega^2 - 3)^2}}. \quad (8)$$

Although this expression is not very telling, its illustration in Figure 4(b) demonstrates how different this criterion is from that of the haplodiploid model. In the diploid model, males and females evolve on the same fitness landscape. Therefore, both males and females benefit from heterozygote advantage. This reduces the influence of the hybrid incompatibility on the optimal location of the population in genotype space, which thereby makes asymmetric coexistence less likely. Indeed, a heterozygote advantage of  $\omega - 1 = \sigma > \approx 0.14$  is sufficient to ensure symmetric coexistence for all recombination rates, whereas in the haplodiploid model,  $\sigma > \sqrt{2} - 1 \approx 0.41$  is necessary for symmetric coexistence independent of the recombination rate.

### General stability conditions in the haplodiploid model

Using the results derived for the case of lethal epistasis, and by means of critical examination of the existence and stability conditions that we were able to compute analytically, we arrived at several illustrative conjectures delimiting the evolutionary outcomes in the haplodiploid model when epistasis is not lethal ( $\gamma_1, \gamma_2 \neq 1$ ). These were all confirmed by extensive numerical simulations (see Mathematica Online Supplement). Note that assortative mating was not considered here.



■ Exclusion or Tristability ■ Symmetric Coexistence

Figure 5: In haplodiploids, symmetric coexistence requires that heterozygote advantage,  $\sigma$ , is strong enough to both compensate for recombination such that the condition in equation 10 is fulfilled (see also Fig. 4(a)), and to overcome the deleterious effects of epistasis, as expressed by condition 11 for symmetric epistasis.

Firstly, strong heterozygote advantage can always override the effect of epistasis. Specifically, if

$$\omega > \sqrt{2}, \quad (9)$$

the evolutionary outcome is always symmetric coexistence, regardless of the values of  $\gamma_1$  and  $\gamma_2$ . This is true not only for a single pair of interacting loci, but also for an arbitrary number of independent incompatibility pairs, because the detrimental effects caused by each incompatibility pair are eventually resolved independently (see also the section on multiple loci below). This result can be deduced from equation (7) for  $\rho = 0.5$  and therefore corresponds to an upper bound: if heterozygote advantage is very strong, recombination no longer affects the outcome.

Secondly, recombination is a key player to determine whether compromise or exclusion can occur. In particular,

$$\rho < \frac{\omega^2 - 1}{\omega^2} \quad (10)$$

is a sufficient condition for the observation of symmetric coexistence, independent of the strength and symmetry of epistasis. This makes intuitive sense, because hybrid incompatibility is masked until gametes are broken up by recombination.

Thirdly, for symmetric epistasis ( $\gamma_1 = \gamma_2$ ), there are three possible equilibrium patterns: symmetric coexistence, exclusion, and tristability of the two former types of equilibria. A necessary and sufficient condition for observation of anything but symmetric coexistence is

$$\omega < \sqrt{2} \quad \text{and} \quad \rho > \frac{\omega^2 - 1}{\omega^2} \quad \text{and} \quad \gamma_1 = \gamma_2 > \frac{2(\omega - 1)}{\omega}. \quad (11)$$

If the recombination rate  $\rho$  and the epistatic effects  $\gamma_1, \gamma_2$  are very close to this limit, there is tristability; if they are far away, there is exclusion (cf. Fig. 5).

Finally, for asymmetric epistasis ( $\gamma_1 \neq \gamma_2$ ), the dynamics display the whole range of possible evolutionary outcomes: symmetric coexistence, asymmetric coexistence, single-locus polymorphism, exclusion, as well as tristability of exclusion *and* symmetric coexistence, and single-locus polymorphism *and* symmetric coexistence. The local stability criterion for the stability of the monomorphic equilibria (i.e., the criterion for exclusion, or tristability of exclusion and symmetric coexistence) is

$$\omega < \sqrt{2} \quad \text{and} \quad \rho > \frac{\omega^2 - 1}{\omega^2} \quad \text{and} \quad \gamma_2 > \frac{2(\omega - 1)}{\omega}. \quad (12)$$

Thus, if epistasis is strong as compared with heterozygote advantage, no degree of asymmetry is sufficient to promote a compromise between males and females (i.e., single-locus polymorphism or asymmetric coexistence). In fact, we observe the following necessary (but not sufficient) condition for a single-locus polymorphism:

$$\omega < \sqrt{2} \quad \text{and} \quad \rho > \frac{\omega^2 - 1}{\omega^2} \quad \text{and} \quad \gamma_1 > \frac{2(\omega - 1)}{\omega} \quad \text{and} \quad \gamma_2 < \frac{2(\omega - 1)}{\omega}. \quad (13)$$

Hence, only a tight balance between the selective pressures of epistasis and heterozygote advantage in combination with asymmetry of the hybrid incompatibility promotes a long-term equilibrium with compromise.

## An extension to multiple loci

### Incompatibilities involving four loci

Above, we have demonstrated that recombination is an essential player when determining whether exclusion or coexistence is the long-term outcome in the haplodiploid dynamics. In order to see how our results change in the (biologically relevant) case of multiple hybrid incompatibilities, we implemented the dynamics for four loci. Given the complexity of the system, we considered only lethal incompatibilities, i.e.  $\gamma_i = 1$  for all interactions  $i$ . With this extension, we consider two scenarios. Firstly, in the “*pairwise*” case we consider pairs of independent hybrid incompatibilities, where we assume that the incompatible loci are located next to each other (locus **A** interacts with locus **B** at recombination distance  $\rho_{12}$ , and locus **C** with locus **D** at recombination distance  $\rho_{34}$ ), which leaves four viable male haplotypes ( $A_+B_+C_+D_+$ ,  $A_+B_+C_-D_-$ ,  $A_-B_-C_+D_+$  and  $A_-B_-C_-D_-$ ). Secondly, in the “*network*” case we assume that all loci interact such that only two viable male haplotypes exist  $A_+B_+C_+D_+$  and  $A_-B_-C_-D_-$ . In both cases, heterozygote advantage is defined as before, now acting on all four loci multiplicatively.

Under this model, we derived the conditions under which exclusion (the purging of all foreign alleles resulting in a monomorphic equilibrium) is locally stable (cf. Mathematica Online Supplement). For the pairwise case, exclusion is stable only if heterozygote advantage is relatively weak:

$$\omega < \min \left[ \frac{1}{\sqrt{1 - \rho_{12}}}, \frac{1}{\sqrt{1 - \rho_{34}}} \right], \quad (14)$$

where  $\rho_{ij}$  is the recombination rate between neighboring loci  $i$  and  $j$ . Note that this is independent of the recombination rate between non-interacting loci, here  $\rho_{23}$ . If  $\rho_{12} = \rho_{34}$ , this expression is equivalent to equation 7 (Fig. 4(a)). Overall, this condition indicates that exclusion, which we define as the fixation of one of the parental haplotypes, is less likely with four interacting loci than with two. This is because the fate of the two pairs of incompatibilities is decided independently, and exclusion requires that both pairs of incompatibilities fix for the same parental haplotype.

For the network case, the condition for stability of exclusion (see also Fig. S3) is

$$\omega < ((1 - \rho_{12})(1 - \rho_{23})(1 - \rho_{34}))^{-\frac{1}{4}}. \quad (15)$$

In this scenario, exclusion is a more likely outcome with two pairs of incompatibilities than with one. This is because there are more unfit intermediate types in this scenario as compared with the pairwise model. Specifically in males, 14 out of the 16 possible haplotypes do not survive to adulthood. To compensate for this fitness cost, any alternative evolutionary outcome requires strong heterozygote advantage.

### Incompatibilities involving an arbitrary number of loci

From the results for two and four loci, we derived a conjecture that generalizes to an arbitrary number of loci. For the pairwise case, equation 14 can be generalized to

$$\omega < \min \left[ \frac{1}{\sqrt{1 - \rho_{ij}}} \right], \quad (16)$$

with  $i$  and  $j$  representing neighboring interacting loci. Note that this result holds only if interacting loci are next to each other on the same chromosome, or if all loci are unlinked (in which case it simplifies to  $\omega < \sqrt{2}$ ).

For the network case, equation (15) generalizes to

$$\omega < \left( \prod_{\substack{i=1 \\ j=i+1}}^{n-1} (1 - \rho_{ij}) \right)^{-\frac{1}{n}}, \quad (17)$$

with  $i$  and  $j$  neighboring loci and  $n$  the total number of loci in the network. Unlike in the pairwise case, the results for the network case do not depend on the genetic architecture (here, the ordering of loci along the genome).

We can therefore deduce that, for the pairwise case, exclusion becomes increasingly unlikely as the number of pairs of independent hybrid incompatibilities involved in the genetic barrier increases. Conversely, the opposite result is observed for the network case: more loci make exclusion a more likely outcome, but each additional interaction contributes less (cf. Fig. S3).

### Increased assortative mating counteracts recombination and heterozygote advantage

Increasing the strength of assortative mating,  $\alpha > 0$ , counteracts the hybridization-favoring effect of heterozygote advantage, because matings between individuals with the same genotype are more common under stronger, positive assortment. Under sufficiently large positive  $\alpha$ , exclusion is unavoidable. In general, increasing  $\alpha$  leads to less maintenance of polymorphism in the population (Fig. S4). Conversely, when  $\alpha < 0$ , which means that individuals prefer to mate with those whose genotype is most different from their own, polymorphism is more likely to be maintained in the population.

Also with assortative mating, recombination remains a key player in determining the evolutionary outcome. When  $\alpha < 0$  and recombination is small, symmetric coexistence is possible even in the absence of heterozygote advantage (i.e.,  $\sigma = 0$ ; Fig. S4). Indeed, under these conditions and assuming epistasis is very strong, (almost) all hybrid males are dead and only parental males survive. This ‘disassortative’ mating ( $\alpha < 0$ ) creates a bias for the

rare male haplotype. For example, if one female genotype increases in frequency, it will seek mainly the males of the other parental haplotype to reproduce with (which are currently rare, as their frequency is directly tied to the frequency of the females in the previous generation). This will increase their reproductive success, which leads to an increase of this haplotype frequency. Therefore, under this mate choice regime, we would observe a stable population composed almost exclusively of the  $A_+B_+$  and  $A_-B_-$  haplotypes.

### Differences between the haplodiploid and the diploid systems

As described above and illustrated in Figure 6, the resulting haplodiploid dynamics display a wider range of possible evolutionary outcomes than the diploid dynamics. Because both males and females profit from heterozygote advantage in the diploid model, polymorphism is always maintained; in other words, even the smallest amount of heterozygote advantage promotes the creation or maintenance of diversity in diploids (Table S3). Conversely, in the haplodiploid model, polymorphism can be lost either at one or both loci, resulting in a single-locus polymorphism or exclusion. Thus, alleles responsible for incompatibilities are more effectively purged in the haplodiploid model.

In the diploid model, a single-locus polymorphism is never stable: Assume locus  $A$  is polymorphic and locus  $B$  is fixed for allele  $B_+$ . Then, a new mutant carrying allele  $B_-$  will always have a selective advantage regardless of the genotype in which it first appears (Table S3). In contrast, in the haplodiploid model, this is no longer true as the mutant carrying allele  $B_-$  will have a much lower fitness in males when associated to allele  $A_+$ . Therefore, if the cost of generating this unfit haplotype in males overrides the advantage in females, and allele  $A_+$  is at high frequency, then invasion of the  $B_+$  mutant may be prevented, leading to the stability of the single-locus polymorphism.

When polymorphism is maintained at both loci at equilibrium (i.e., asymmetric and symmetric coexistence), epistasis creates associations between the compatible alleles which results in elevated linkage disequilibrium (LD). Recombination breaks the association between alleles, thus high recombination decreases normalized LD ( $D'$ , where  $D' = \frac{LD}{D_{\max}}$  (Lewontin, 1964); Fig. S5).  $D'$  increases with the strength of heterozygote advantage at low recombination rates because it maximizes the discrepancy between highly fit double-heterozygote females on the one hand that can, under low recombination rate, still produce many fit male offspring and introgressed females on the other, who are less fit and produce many unfit hybrid males

In Figure S6, we compare the normalized LD (i.e.  $D'$ ) between the haplodiploid and diploid models. When polymorphism is maintained at both loci in both the haplodiploid and diploid model, normalized LD is always larger in haplodiploids than diploids. The difference in normalized LD between haplodiploids and diploids is maximized for intermediate recombination rates, where recombination is strong enough to create unfit hybrid genotypes, but not efficient enough to break the associations that are generated. Due to the increased selection against hybrid incompatibility in haploid males in the haplodiploid model, the normalized LD is usually 2-3 times higher in the haplodiploid as compared with the diploid model.

Thus, the hybrid incompatibility leaves a statistical signature in a population, even if the population finds itself at an equilibrium. The increased association across the genome, exhibited if the interacting loci are on the same chromosome, may also result in an underestimate of the recombination rate. Although both the diploid and the haplodiploid models display the elevated LD signal, it is much more pronounced in the haplodiploid scenario. This is because only an eighth of the possible diploid male genotypes suffer the cost of the incompatibility as compared to half of the possible haploid male genotypes.

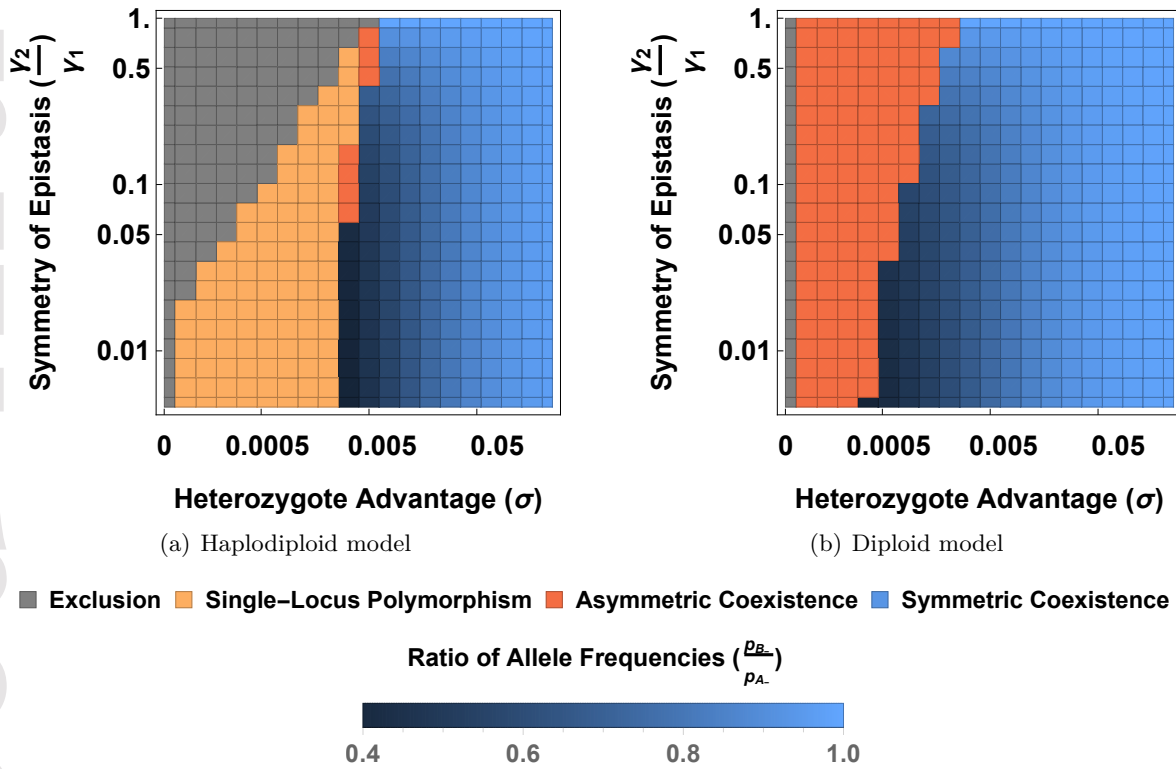


Figure 6: More evolutionary outcomes are possible in (a) the haplodiploid than (b) the diploid model. The y-axis shows the degree of asymmetry of epistasis, displayed as the ratio of the two epistasis parameters ( $\frac{\gamma_2}{\gamma_1}$ ) for a constant value of  $\gamma_1 = 0.01$ . For symmetric coexistence, the locally stable equilibrium can be at any point on the diagonal  $p_{B-} = 1 - p_{A-}$ , where  $p_{A-}$  and  $p_{B-}$  denote the allele frequencies of the  $-$  allele at the respective locus. Blue shading illustrates the location of the equilibrium at symmetric coexistence: darker shades correspond to a bigger disparity in allele frequencies. This is the case when the asymmetry of the two epistasis parameters is large (i.e. smaller values on the y-axis) because smaller values of  $\gamma_2$  favor the  $A_-B_+$  haplotype over the  $A_+B_-$  haplotype. (Here,  $\gamma_1 = 0.01$ ,  $\rho = 0.5$ ,  $\alpha = 0$ .)

## Discussion

Multiple recent studies have highlighted the pervasive nature of hybridization and its potential consequences for diversification and speciation (Abbott et al., 2013; Runemark et al., 2017; Montecinos et al., 2017). We here modeled the fate of a hybrid population in a scenario in which hybridization is simultaneously favored and selected against, inspired by a natural population of hybrid ants that simultaneously displays heterosis and hybrid incompatibility. In addition, both adaptive introgression and hybrid incompatibilities have been identified in natural systems (Heliconius Genome Consortium, 2012; Whitney et al., 2015; Corbett-Detig et al., 2013) and it is therefore likely that both processes may occur simultaneously during a single hybridization event. Furthermore, we were interested in comparing the long-term evolution of populations exposed to these opposing selective pressures under different ploidies (haplodiploid versus diploid), since it has been argued that haplodiploids might speciate more easily than diploids (Lohse and Ross, 2015). Finally, the comparison of ploidies can also be transferred to the case of diploid species with sex chromosomes, in which the described fitness landscape results in the diploid dynamics on the autosomes, and in the haplodiploid dynamics on the X/Z chromosome.

Our model considers a population in which heterozygote advantage and hybrid incompatibility act simultaneously on the same pair of loci, which creates a rugged fitness landscape with a ridge of high-fitness heterozygote genotypes, adjacent to which there are holes of incompatible double homozygotes (Fig. 2(a)). In haplodiploids, haploid males cannot profit from heterozygote advantage but suffer strongly from hybrid incompatibility (Fig. 2(b)). This results in a conflict of ploidies/sexes over the optimal location in the fitness landscape, because haploid males survive best if one parental haplotype is fixed whereas diploid females profit from maximum heterozygosity. Although females suffer from the same incompatibility as males, their presence is mainly masked in the diploid individuals because of the recessivity of the hybrid incompatibility. This is similar to Haldane’s rule (Charlesworth et al., 1987; Koevoets and Beukeboom, 2009).

### How ploidy matters

We found that, in the haplodiploid model, there exist four different stable outcomes of the conflict over hybrid status (Fig. 3): exclusion, where “males/haploids win”; symmetric coexistence, where “females/diploids win”; and two outcomes, single-locus polymorphism and asymmetric coexistence, where a compromise between male costs and female benefits is mediated by high frequencies of introgressed females. In fact, since low-frequency heterozygotes are favored both in males and in females in the diploid model, while only suffering the hybrid cost if introgressed alleles rise to high frequencies, exclusion and single-locus polymorphism never occur in the diploid model, which reduces the number of possible outcomes to asymmetric and symmetric coexistence. Therefore, consistent with Pamilo (1979); Pamilo and Crozier (1981); Patten et al. (2015), we found that introgression and maintenance of polymorphism, and thus long-term hybridization, are less likely in haplodiploids as compared to diploids.

Prior work has found that in haplodiploid species traditional sexual conflict tends to be resolved in favor of females because genes spend two thirds of their time in females (Albert and Otto, 2005). In our model, the co-occurrence of heterozygote advantage and hybrid incompatibility also creates an apparent sexual conflict that is caused by the difference in ploidy between the sexes. For several scenarios, we here derived the conditions for whether this conflict is resolved in favor of diploid females or haploid males. We find, that in addition to the strength of selection, recombination is a major player (cf. Fig. 4 and equation 12); only if recombination breaks up gametes, the hybrid incompatibility is expressed. With free

recombination, i.e., if the interacting genes are found on separate chromosomes, heterozygote advantage has to be very strong to counteract the hybrid incompatibility. We find that it has to be on the same order of magnitude as the strength of the incompatibility, but can be slightly lower in its absolute value. For example, heterozygote advantage with strength 41% is sufficient to result in symmetric coexistence even if the incompatibility is lethal (Fig. 4B). Thus, under consideration of absolute magnitude across the full parameter range, our results are consistent with prior work. However, reported cases and potential mechanisms of hybrid incompatibility indicate that large effects are feasible, whereas observed cases of heterozygote advantage or heterosis of large effect are relatively rare (Hedrick, 2012). Therefore, it may well be that under natural circumstances, the conflict modeled here may indeed be likely to be resolved via purging of at least one incompatible allele and thus in favor of males/haploids.

As expected in the presence of epistasis, we observed that linkage disequilibrium (LD) is elevated at all polymorphic stable states (i.e., for symmetric and asymmetric coexistence) both in the diploid and haplodiploid models, especially at intermediate recombination rates. This is particularly true for haplodiploids, which display about 2-3 times the LD of the diploid model with the same parameters. Transferred to the context of X/Z chromosomes, this is consistent with observations of larger LD on the X chromosome as compared with autosomes (Wall et al., 2002; Sandor et al., 2006; Li and Merilä, 2010). It has been argued that this is because selection is more effective on X-linked loci: recessive deleterious mutations are more visible to selection in haploid individuals (Charlesworth et al., 1987). However, a hybrid incompatibility accompanied by heterosis/heterozygote advantage as in our model may not be purged but create a continuous high-LD signal in an equilibrium population. This can potentially result in less efficient recombination and in underestimates of recombination rates on X chromosomes (because recombined individuals are not observed).

### **Generalization to multiple incompatibilities**

Exclusion remains a stable solution when we extend the model to multiple loci and incompatibilities. We describe an interesting difference between multiple independent pairs of incompatibilities, and multiple loci that all interact with each other: in the latter case, exclusion becomes increasingly probable because the number of viable males decreases. This scenario of higher-order epistasis has recently received attention with regards to speciation (Paixão et al., 2014; Fraïsse et al., 2014; Kulmuni and Westram, 2017), and it will be interesting in the future to identify molecular scenarios (for example, involving biological pathways) that could result in such incompatibilities. In contrast, exclusion becomes less likely in the case of independent incompatibility pairs, where each incompatibility has to be purged independently, and in the same direction, for exclusion to occur. Here, mechanisms that reduce the recombination rate, such as inversions, could potentially invade and tilt the balance towards coexistence and thus maintenance of polymorphism in the hybrid population. It is important to note that the independent purging of incompatibilities, which leads to a decreasing probability of exclusion with increasingly many incompatibility pairs, is only true in effectively infinite-sized populations. In small populations, we expect that exclusion becomes a more likely scenario, especially if lethal incompatibility pairs are present.

### **Model assumptions**

We chose a classical population-genetic modeling approach (Bürger, 2000; Nagylaki et al., 1992) to study how the co-occurrence of heterozygote advantage and hybrid incompatibility affect the long-term dynamics of a hybrid population. By treating the problem in a deterministic framework and considering only two loci throughout most of the manuscript, we greatly

oversimplify the situation in the natural population that inspired our model. However, at the same time this allowed us to gain a general insight, (often by means of analytical expressions), into how opposing selective pressures in genomes may be resolved, and to contrast these outcomes between haplodiploid and diploid systems. In addition to some obvious mechanisms at play in natural populations, which we ignore in our model (e.g., random genetic drift), some extensions of the model could be interesting to elaborate on in the future. For example, the ant populations represent networks of interacting nests with many queens per nest, but potentially different mating flight timing that depends, for example, on sun exposure in the spring. Thus, for the purpose of population-genetic inference of the evolutionary history (and potential evolutionary fate) of the hybrid ant population in Finland, it would be desirable to incorporate population structure, uneven sex ratios at birth, and sex-biased dispersal into the model, and obtain population-genomic data to infer evolutionary parameters.

### **Is the natural population at an equilibrium of asymmetric coexistence?**

Model fitting results (see Supplementary Methods, Results, and Discussion) are inconclusive about the fate of the natural ant population that inspired our model. Our results suggest that the natural population might be approaching an evolutionary outcome that allows a compromise between male and female interests; either as single-locus polymorphism or via asymmetric coexistence. In particular, our model is able to explain the unusual skew in the population, where *F. aquilonia*-like parental genotypes far outnumber *F. polyctena*-like genotypes (see Supplement). Furthermore, the high recombination rates and strong prezygotic mechanisms operating in the natural population (Kulmuni et al., 2010; Kulmuni and Pamilo, 2014), are consistent with a parameter domain in our models at which asymmetric coexistence can be stably maintained over a wide range of values of female hybrid advantage. More complex models, for example including more than two incompatibility loci, may be better able to explain the high frequencies of introgressed alleles as compared to parental females observed in the natural hybrid population. As argued in the Results, interactions at or between multiple loci should result in steeper differences of introgressed-allele frequencies across life stages than our model is able to produce.

### **Implications for hybrid speciation**

Our model illustrates how the co-occurrence of heterozygote advantage and hybrid incompatibility affects haplodiploid and diploid populations. We can hypothesize how these different outcomes may provide an engine to hybrid speciation, or which other long-term evolutionary scenarios we expect to arise. The case of exclusion, which is possible only in the haplodiploid model, will lead to loss of diversity in the hybrid population, and, in the two-locus case, should result in the reversion of the hybrid population into one of its parental species. However, if multiple pairs of interacting loci are resolved independently, they may be purged randomly towards either parent, which could result in a true hybrid species that is isolated from both its parental species (Buerkle et al., 2000; Butlin and Ritchie, 2013; Schumer et al., 2015). In fact, our finding that exclusion is less likely to occur in populations with multiple pairs of interacting loci may result from exactly this mechanism, but it is beyond the scope of this manuscript to explore this further.

The long-term fate of the population is less straightforward to anticipate in the case of polymorphic stable equilibria. For any of these, heterozygote advantage is strong enough to stabilize the polymorphism either at one or both loci. Without further occurrence of functional mutations, males (in the haplodiploid model) and double-homozygotes for the incompatible alleles will continue to suffer a potentially large fitness cost. Mechanisms that

could reduce this cost would be increased assortative mating or decreased recombination. However, neither of these would necessarily cause isolation from the parental species, unless they involved additional hybrid incompatibilities which isolate the hybrid population from its parental species. Alternatively, mutations that lower the hybrid fitness cost could invade, which would result in a weakening of species barriers and promote further introgression from the parental species. This indicates that any scenario in which polymorphic equilibria are stable may indeed be an unlikely candidate for hybrid speciation. Considering that such stable polymorphism (either as symmetric or asymmetric coexistence) is the only possible outcome in the diploid model, this results in the prediction that hybrid speciation would be more likely in a haplodiploid scenario. This is an interesting observation that is in line with other predictions that haplodiploids speciate more easily, that X/Z chromosomes are engines of speciation (Lima, 2014), and that hybrid speciation is rare (Schumer et al., 2014).

### Relevance of the model for sex chromosomes

Haplodiploids and X/Z chromosomes have a similar mode of inheritance, where one sex carries a single copy of the chromosome, and the other carries two copies. Therefore, our results apply equally to cases of X-to-X or Z-to-Z hybrid incompatibilities (Lohse and Ross, 2015). Although haplodiploid systems do not include all of the unique evolutionary phenomena exhibited by sex chromosomes (Abbott et al., 2017), our results for haplodiploids are relevant for sex chromosomes. Our model predicts the long-term evolution of a population under the simultaneous influence of heterozygote advantage and hybrid incompatibility, and indicates the signatures that this type of fitness landscape could leave depending on whether it finds itself on an X chromosome or an autosome.

Firstly, the complex selection pressure imposed by the co-occurrence of heterozygote advantage and hybrid incompatibility manifests itself as an apparent sexual conflict on the X chromosome/in haplodiploids. This conflict is caused by the ploidy difference between the sexes. Here, the same fitness landscape that would be masked on an autosome and result in a stable polymorphism, creates a signal of sexually antagonistic selection on an X chromosome. Most importantly, this signal is created without the need for direct sexually antagonistic selection on single functional genes that have a sex-specific antagonistic effect. Thus, our model proposes an additional mechanism by which sex chromosomes can appear as hotspot of sexual conflict (e.g., Gibson et al., 2002; Pischedda and Chippindale, 2006).

Secondly, we find that purging of incompatibilities is more likely in the haplodiploid model, and thus on X/Z chromosomes. This is consistent with the faster-X theory (Charlesworth et al., 1987). However, only if recombination is strong enough, incompatibilities will become visible to selection and purged in the presence of heterozygote advantage. If they are not purged, they may persist as a long-term polymorphism, invisible to most empirical approaches, and confound population-genetic inference by creating signals of elevated linkage disequilibrium.

### Conclusion

Hybridization is observed frequently in natural populations, and can have both deleterious and advantageous effects. We here show how diverse outcomes can be produced even under a rather simple model of a single hybrid population, in which heterozygote advantage and hybrid incompatibility are occurring at the same time. Consistent with previous theory on haplodiploids and X/Z chromosomes, we found that incompatible alleles are more likely to be purged in a haplodiploid than in a diploid model. Nevertheless, our results suggest that long-term hybridization can occur even in the presence of hybrid incompatibility, and if there are

many incompatible pairs or many loci involved in the incompatibility. The evolutionary fate of the Finnish hybrid ant population that inspired our model is difficult to predict; further population-genetic analysis will be necessary to gain a more complete picture of its structure and evolutionary history.

## Acknowledgements

We thank Laura Cêtre for her work on a previous version of the model. We thank Pekka Pamilo, Bret Payseur, two anonymous reviewers, and the members of the Bank and Kulmuni labs for discussion of the manuscript. This research was supported by the Fundação Calouste Gulbenkian and in part by the National Science Foundation under Grant No. NSF PHY-1125915. JK was supported by the Human Frontier Science Program, Finnish Cultural Foundation, Academy of Finland (252411 to CoE in Biological Interactions).

## Data Accessibility

The complete documentation of all steps of the analysis is available as a Mathematica Online Supplement. Ant colony data is provided as Supplementary Table S1; genotype frequency data were obtained from Kulmuni and Pamilo (2014).

## Author Contributions

CB, JK, and RB designed research, AB and CB developed the models, AHG performed simulations and data analysis, all authors interpreted the results and wrote the manuscript.

## Bibliography

- Abbott, J., Nordén, A., and Hansson, B. (2017). Sex chromosome evolution: Historical insights and future perspectives. *Proceedings of the Royal Society of London B: Biological Sciences*, 284(1854):20162806.
- Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A. W., Parisod, C., Pfennig, K., Rice, A. M., Ritchie, M. G., Seifert, B., Smadja, C. M., Stelkens, R., Szymura, J. M., Väinölä, R., Wolf, J. B. W., and Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2):229–246.
- Albert, A. and Otto, S. (2005). Sexual selection can resolve sex-linked sexual antagonism. *Science*, 310(5745):119–121.
- Arnqvist, G. and Rowe, L. (2005). *Sexual conflict*.
- Bateson, W. (1909). Heredity and variation in modern lights. *Darwin and modern science*, pages 85–101.
- Beresford, J., Elias, M., Pluckrose, L., Sundström, L., Butlin, R., Pamilo, P., and Kulmuni, J. (2017). Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. *Molecular Ecology*, 26(15):4013–26.

- Bernardes, J., Stelkens, R., and Greig, D. (2017). Heterosis in hybrids within and between yeast species. *Journal of Evolutionary Biology*, 30(3):538–548.
- Buerkle, C. A., Morris, R. J., Asmussen, M. A., and Rieseberg, L. H. (2000). The likelihood of homoploid hybrid speciation. *Heredity*, 84(4):441–451.
- Bürger, R. (2000). *The mathematical theory of selection, recombination, and mutation*, volume 228. Wiley Chichester.
- Butlin, R. K. and Ritchie, M. G. (2013). Pulling together or pulling apart: hybridization in theory and practice. *Journal of Evolutionary Biology*, 26(2):294–298.
- Charlesworth, B., Coyne, J., and Barton, N. (1987). The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, 130(1):113–146.
- Chen, C., Zhiguo, E., and Lin, H.-X. (2016). Evolution and molecular control of hybrid incompatibility in plants. *Frontiers in Plant Science*, 7:1208.
- Chen, Z. (2013). Genomic and epigenetic insights into the molecular bases of heterosis. *Nature Review Genetics*, 14(7):471–482.
- Corbett-Detig, R. B., Zhou, J., Clark, A. G., Hartl, D. L., and Ayroles, J. F. (2013). Genetic incompatibilities are widespread within species. *Nature*, 504(7478):135–7.
- Crozier, R. and Pamilo, P. (1996). *Evolution of social insect colonies: Sex allocation and kin selection*. Oxford University Press, Oxford, UK.
- De Cara, M., Barton, N., and Kirkpatrick, M. (2008). A model for the evolution of assortative mating. *The American Naturalist*, 171(5):580–596.
- de la Folia, A., Bain, S., and Ross, L. (2015). Haplodiploidy and the reproductive ecology of arthropods. *Current Opinion in Insect Science*, 9:36–43.
- Dieckmann, U. and Doebeli, M. (1999). On the origin of species by sympatric speciation. *Nature*, 400(6742):354–357.
- Dobzhansky, T. (1936). Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, 21(2):113.
- Evans, J., Shearman, D., and Oldroyd, B. (2004). Molecular basis of sex determination in haplodiploids. *Trends in Ecology and Evolution*, 19(1):1–3.
- Fraïsse, C., Elderfield, J., and Welch, J. (2014). The genetics of speciation: are complex incompatibilities easier to evolve? *Journal of Evolutionary Biology*, 27(4):688–99.
- Gibson, J., Chippindale, A., and Rice, W. (2002). The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1490):499–505.
- Hedrick, P. W. (2012). What is the evidence for heterozygote advantage selection? *Trends in Ecology & Evolution*, 27(12):698–704.
- Heliconius Genome Consortium (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, 487(7405):94.

- Höllinger, I. and Hermisson, J. (2017). Bounds to parapatric speciation: A Dobzhansky–Muller incompatibility model involving autosomes, X chromosomes, and mitochondria. *Evolution*, 71(5):1366–1380.
- Jablonka, E. and Lamb, M. J. (1991). Sex Chromosomes and Speciation. *Proceedings of the Royal Society B: Biological Sciences*, 243(1308):203–208.
- Johnson, N. A. and Lachance, J. (2012). The genetics of sex chromosomes: evolution and implications for hybrid incompatibility. *Annals of the New York Academy of Sciences*, 1256(1):E1–E22.
- Knegt, B., Potter, T., Pearson, N., Sato, Y., Staudacher, H., Schimmel, B., Kiers, E., and Egas, M. (2017). Detection of genetic incompatibilities in non-model systems using simple genetic markers: hybrid breakdown in the haplodiploid spider mite *tetranychus evansi*. *Heredity*, 118(4):311.
- Koevoets, T. and Beukeboom, L. (2009). Genetics of postzygotic isolation and Haldane’s rule in haplodiploids. *Heredity*, 102(1):16–23.
- Kopp, M., Servedio, M. R., Mendelson, T. C., Safran, R. J., Rodríguez, R. L., Hauber, M. E., Scordato, E. C., Symes, L. B., Balakrishnan, C. N., Zonana, D. M., et al. (2017). Mechanisms of assortative mating in speciation with gene flow: Connecting theory and empirical research. *The American Naturalist*, 191(1).
- Kraaijeveld, K. (2009). Male genes with nowhere to hide; sexual conflict in haplodiploids. *Animal Biology*, 59(4):403–415.
- Kulmuni, J. and Pamilo, P. (2014). Introgression in hybrid ants is favored in females but selected against in males. *Proceedings of the National Academy of Sciences*, 111(35):12805–10.
- Kulmuni, J., Seifert, B., and Pamilo, P. (2010). Segregation distortion causes large-scale differences between male and female genomes in hybrid ants. *Proceedings of the National Academy of Sciences*, 107(16):7371–6.
- Kulmuni, J. and Westram, A. M. (2017). Intrinsic incompatibilities evolving as a by-product of divergent ecological selection: Considering them in empirical studies on divergence with gene flow. *Molecular Ecology*, 26(12):3093–3103.
- Lewontin, R. (1964). The interaction of selection and linkage. i. general considerations; heterotic models. *Genetics*, 49–67(1):49.
- Li, M.-H. and Merilä, J. (2010). Sex-specific population structure, natural selection, and linkage disequilibrium in a wild bird population as revealed by genome-wide microsatellite analyses. *BMC evolutionary biology*, 10(1):66.
- Lima, T. G. (2014). Higher levels of sex chromosome heteromorphism are associated with markedly stronger reproductive isolation. *Nature Communications*, 5:4743.
- Lohse, K. and Ross, L. (2015). What haplodiploids can teach us about hybridization and speciation. *Molecular Ecology*, 24(20):5075–5077.
- Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, 20(5):229–237.

- Montecinos, A. E., Guillemin, M.-L., Couceiro, L., Peters, A. F., Stoeckel, S., and Valero, M. (2017). Hybridization between two cryptic filamentous brown seaweeds along the shore: analysing pre- and postzygotic barriers in populations of individuals with varying ploidy levels. *Molecular Ecology*, 26(13):3497–3512.
- Muller, H. (1942). Isolating mechanisms, evolution and temperature. In *Biology Symposium*, volume 6, pages 71–125.
- Nagylaki, T. et al. (1992). *Introduction to theoretical population genetics*, volume 142. Springer-Verlag Berlin.
- Orr, H. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, 139(4):1805–1813.
- Otto, S. P. and Day, T. (2007). *A biologist's guide to mathematical modeling in ecology and evolution*, volume 13. Princeton University Press.
- Paixão, T., Bassler, K. E., and Azevedo, R. B. R. (2014). Emergent speciation by multiple Dobzhansky-Muller incompatibilities. *bioRxiv*, page 8268.
- Pamilo, P. (1979). Genic variation at sex-linked loci: Quantification of regular selection models. *Hereditas*, 91(1):129–133.
- Pamilo, P. and Crozier, R. H. (1981). Genic variation in male haploids under deterministic selection. *Genetics*, 98(1):199–214.
- Patten, M., Carioscia, S., and Linnen, C. (2015). Biased introgression of mitochondrial and nuclear genes: A comparison of diploid and haplodiploid systems. *Molecular Ecology*, 24(20):5200–5210.
- Pischedda, A. and Chippindale, A. K. (2006). Intralocus sexual conflict diminishes the benefits of sexual selection. *PLoS biology*, 4(11):e356.
- Presgraves, D. (2008). Sex chromosomes and speciation in *Drosophila*. *Trends in Genetics*, 24(7):336–343.
- Runemark, A., Trier, C. N., Eroukhmanoff, F., Hermansen, J. S., Matschiner, M., Ravinet, M., Elgvin, T. O., and Saetre, G.-P. (2017). Variation and constraints in hybrid genome formation. *bioRxiv*, page 107508.
- Sandor, C., Farnir, F., Hansoul, S., Coppieters, W., Meuwissen, T., and Georges, M. (2006). Linkage disequilibrium on the bovine x chromosome: characterization and use in quantitative trait locus mapping. *Genetics*, 173(3):1777–1786.
- Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, 323(5915):737–741.
- Schluter, D. and Conte, G. (2009). Genetics and ecological speciation. *Proceedings of the National Academy of Sciences*, 106(Supplement 1):9955–9962.
- Schumer, M., Cui, R., Rosenthal, G. G., and Andolfatto, P. (2015). Reproductive isolation of hybrid populations driven by genetic incompatibilities. *PLOS Genetics*, 11(3):1–21.
- Schumer, M., Rosenthal, G., and Andolfatto, P. (2014). How common is homoploid hybrid speciation? *Evolution*, 68(6):1553–1560.

Schwarz, D., Matta, B., Shakir-Botteri, N., and McPherson, B. (2005). Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature*, 436(7050):546–9.

Seehausen, O., Butlin, R., Keller, I., Wagner, C., Boughman, J., Hohenlohe, P., Peichel, C., Saetre, G.-P., Bank, C., Brännström, Å., Brelsford, A., Clarkson, C., Eroukhmanoff, F., Feder, J., Fischer, M., Foote, A., Franchini, P., et al. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, 15(3):176–192.

Servedio, M. and Noor, M. (2003). The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):339–364.

Song, Y., Endepols, S., Klemann, N., Richter, D., Matuschka, F.-R., Shih, C.-H., Nachman, M., and Kohn, M. (2011). Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Current Biology*, 21(15):1296–1301.

Suomalainen, E., Saura, A., and Lokki, J. (1987). *Cytology and evolution in parthenogenesis*. CRC Press, Boca Raton, Florida.

Wall, J. D., Andolfatto, P., and Przeworski, M. (2002). Testing models of selection and demography in *Drosophila simulans*. *Genetics*, 162(1):203–216.

Whitney, K., Randell, R., and Rieseberg, L. (2010). Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus*. *New Phytologist*, 187(1):230–239.

Whitney, K. D., Broman, K. W., Kane, N. C., Hovick, S. M., Randell, R. A., and Rieseberg, L. H. (2015). Quantitative trait locus mapping identifies candidate alleles involved in adaptive introgression and range expansion in a wild sunflower. *Molecular Ecology*, 24(9):2194–2211.

Wolf, D., Takebayashi, N., and Rieseberg, L. (2001). Predicting the risk of extinction through hybridization. *Conservation Biology*, 15(4):1039–1053.

Wolfram Research, Inc. (2016). Mathematica v. 10.4.1.0. Champaign, Illinois, USA. <https://www.wolfram.com>.