Inferring inter- and intrachromosomal Dobzhansky-Muller incompatibilities from imbalanced segregation of recombinant haplotypes in hybrid populations

Juan Li [1,2, 3,#], Molly Schumer[4], Claudia Bank[1,2, 3,#]

1. Institute of Ecology and Evolution, University of Bern, Bern, Switzerland
2. Gulbenkian Science Institute, Oeiras, Portugal
3. Swiss Institute for Bioinformatics, Lausanne, Switzerland
4. Department of Biology, Stanford University, United States

[#]To whom correspondence should be addressed: juan.li@iee.unibe.ch, claudia.bank@iee.unibe.ch

Abstract

Dobzhansky-Muller incompatibilities (DMIs) are a major component of reproductive isolation between species. DMIs imply negative epistasis, exposed when two diverged populations hybridize. Mapping the locations of DMIs has largely relied on classical genetic mapping, but these approaches are stymied by low power and the challenge of identifying DMI loci on the same chromosome, because strong initial linkage of parental haplotypes weakens statistical tests. Here, we propose new statistics to infer negative epistasis from haplotype frequencies in hybrid populations. When two divergent populations hybridize, the variance of two-locus heterozygosity decreases faster with time at DMI loci than at random pairs of loci. If two populations hybridize at near-even admixture proportions, the deviation of the observed variance from its expectation is negative, which enables us to detect signals of intermediate to strong negative epistasis both within and between chromosomes. When the initial proportion of the two parental populations is uneven, only strong DMIs can be detected with our method, unless migration reintroduces haplotypes from the minor parental population. We use the two new statistics to infer candidate DMIs from three hybrid populations of swordtail fish. We identify numerous new DMI candidates some of which are inferred to interact with several loci within and between chromosomes. Moreover, we discuss our results in the context of an expected enrichment in intrachromosomal over interchromosomal DMIs.

Introduction

Hybrids between diverged populations often suffer from a fitness disadvantage caused by genetic incompatibilities between two or more loci, called (Bateson-)Dobzhansky-Muller incompatibilities, (DMIs). As a result, hybrids may be less fit than their parental species, inviable, or sterile (reviewed in Orr and Presgraves 2000, Fishman and Sweigart 2018). DMIs are common within and between species (Presgraves 2010, Corbett-Detig 2013), often uncovered in crosses between diverged populations or closely related species (reviewed in Wu and Ting 2004, Maheshwari and Barbash 2011).

Despite their biological importance, methods for detecting DMIs in natural hybrid populations are plagued with low power and statistical problems. Current methods rely on a large sample size of hybrid offspring (i.e., thousands of individuals) or exceptionally strong selection, and are only designed to detect pairwise interactions. Examples of known hybrid incompatibilities reflect the biases expected from statistical power that is limited to detecting loci with strong selection against hybrids, e.g., in the case of

malignant melanoma that reduce viability in two pairs *Xiphophorus* species (reviewed in Orr and Presgraves 2000; Powell et al. 2020). In another study, researchers were able to map multiple non-independent incompatibilities contributed to male sterility in a house mouse hybrid zone (Turner and Harr 2014).

Another commonly used approach to infer DMIs from hybrid genomes is to search for unexpected statistical associations (linkage disequilibrium; LD) between physically unlinked regions of the genome. Studies of DMIs that use LD signals between physically unlinked loci (interchromosomal DMIs) have taken advantage of data from simulations and experiments (e.g., *Arabidopsis thaliana*, Simon et al. 2008, and yeast, Li et al. 2013), population genomics (e.g., *Drosophila melanogaster,* Corbett-Detig et al. 2013), and genomic data from hybrid populations (e.g., swordtail fish, Schumer et al. 2014, Schumer and Brandvain 2016). Though widely used, LD-based methods are known to have an unacceptably high false positive rate and low sensitivity, even when researchers attempt to account for confounders such as hybrid population demographic history (Schumer and Brandvain 2016, Satokangas et al. 2020).

Another shortcoming of LD-based statistics is that they are not well suited for detecting interacting loci within chromosomes (intrachromosomal DMIs) in the absence of highly accurate recombination maps. To detect intrachromosomal interactions with LD-based methods, researchers would need to know the exact recombination rates between loci before they could distinguish expected LD caused by physical linkage from excess LD caused by DMIs. The authors of this study are not aware of any existing method that infers both inter- and intrachromosomal DMIs from hybrid genomic data.

Although intrachromosomal DMIs have not been well-studied in the hybrid incompatibility literature, several lines of evidence suggest that they are at least as common as interchromosomal DMIs. First, from a mechanistic perspective, gene order is not random in the genome. Genes tend to cluster on the same chromosome when they are co-expressed, co-regulated, included in one pathway, or part of a protein-protein complex (reviewed in Hurst 2004). When sequencing and protein precipitation technology (high-C, micro-C) were combined to profile the physical interactions (e.g., interactions mediated by enhancers, insulators, cohesion binding sites etc.) between DNA sequences, researchers found that most of the interactions were intrachromosomal, manifesting as chromosome territories and topological associated domains (reviewed in Szabo et al. 2019). This genomic architecture of gene regulation suggests that we may expect DMIs within chromosomes may be more common than DMIs between chromosomes from first principles.

Second, from an evolutionary perspective, natural selection may favor reduced recombination between loci under divergent selection or involved in epistatic interactions (reviewed in Ortiz-Barrientos et al 2016). Introgression is more common in high recombination rate regions of the genome where linkage disequilibrium breaks down rapidly over physical distance (Brandvain et al. 2014, Schumer et al. 2018, Martin et al. 2019, Chaturvedi 2020). One potential consequence is that that high recombination rates disassociate incompatible loci more quickly, whereas tight linkage may facilitate speciation in the face of gene flow (gene coupling, e.g., Bay et al. 2017, Byers et al. 2020; inversions, e.g., Noor et al. 2001). In line with this argument, theoretical studies indicated that hybrid populations isolated from their parent populations are most likely to develop into isolated species (homoploid hybrid speciation) with intermediate recombination between DMI loci (Blanckaert and Bank 2018). Given the predicted

importance of genome organization in adaptation and speciation, tools to identify intrachromosomal DMIs are needed to understand the architecture of reproductive isolation between species.

Here, we present a new statistical approach to close this key knowledge gap and to sensitively infer both inter- and intrachromosomal DMIs. We report that the imbalanced haplotype frequencies caused by selection against DMIs can lead to a unique pattern of LD and heterozygosity. We capture these patterns in two statistics that we name $X(2)$ and $\Delta D2$. We show that sensitivity and specificity of detecting DMIs are high in certain demographic scenarios and apply our statistic to infer DMIs in three swordtail fish hybrid populations. Based on this analysis, we find that DMIs are widespread and complex across hybrid populations of swordtail fish.

Results

Reduced variance of two-locus heterozygosity indicates DMIs

Classical studies of multi-locus evolution revealed that the variance of multi-locus heterozygosity increases with linkage disequilibrium (reviewed in Slatkin 2008). In the presence of population structure, the normalized deviation of the variance of *n*-locus heterozygosity from its expected value, $X(n)$, is elevated (Sved 1968, Brown et al. 1980, Maynard Smith et al. 1993, see Methods). Here, we demonstrate that a pairwise DMI, where *n=2*, changes $X(2)$ in the opposite direction (i.e., it creates a lower-than-expected variance of two-locus heterozygosity) and can be inferred by identifying negative $X(2)$ (i.e., the two-locus version of $X(n)$) along the genome, within a hybrid population. Importantly, this statistic distinguishes LD that occurs due to physical linkage from the specific association of alleles as caused by selection against a DMI.

To first understand the dynamics of $X(2)$ in the simplest scenario, we used a haploid toy model of a DMI with recombination (Kimura 1965, see also Bank et al. 2012). Consider a haploid panmictic population with two biallelic loci, **A** and **B,** with alleles *a*, *A*, and *b*, *B*, where *aB* and *Ab* represent the parental haplotypes and a hybrid incompatibility exists between alleles *A* and *B*. We follow the haplotype frequency dynamics in a model with discrete non-overlapping generations, ignoring genetic drift and assuming that recombination occurs with probability *c*. We denote the fitness of the haplotypes by $w_{ab} = 1, w_{Ab} = 1+\alpha,\ w_{aB} = 1+\beta,\ w_{AB} = (1+\alpha)(1+\beta)(1+\gamma)$ (Figure S1). Throughout the paper, we assume that α and β are small positive numbers, representing a weak beneficial effect of *A* and *B* (following Blanckaert & Bank 2018), and that $-\gamma \gg \alpha, \beta$ (the incompatibility is much stronger than directional selection). We use this toy model to analyze the dynamics of $X(2)$ and $D2_{ik}$ (defined below) upon formation of a new hybrid population composed of a fraction f of *aB* parental individuals and a fraction of $1-f$ of *Ab* parental individuals (i.e., *f* denotes the initial admixture proportion).

The toy model highlights unique features of haplotype frequency dynamics in new hybrid populations in the presence of a DMI (see also Blanckaert and Bank 2018 for the corresponding diploid model). Recombinant haplotypes *ab* and *AB* appear from the second generation onwards. At first, their frequency sharply increases (Figure S2) due to recombination. Then strong selection against the incompatibility sets in, depleting the *AB* haplotype and giving the *ab* haplotype a marginal advantage because offspring of *ab* individuals never harbor the incompatibility (Figure S2). Thus, the strong incompatibility drives haplotype frequency dynamics after recombinants appear.

The deviation of the variance of two-locus heterozygosity from its expectation under linkage equilibrium in this model is $\Delta_{var} = 2\sum_i \sum_k ({g_{ik}}^2 - {p_i}^2{p_k}^2)$, where p_j denotes the frequency of allele j and g_{ik} denotes the frequency of haplotype ik, and where we define $D2_{ik} := {g_{ik}}^2 - {p_i}^2{p_k}^2$ for $i \in \{a, A\}, k \in \{b, B\}$ (see Methods). The normalized version of Δ_{var} is $X(2) = \Delta_{var}/exp(Var)$. In a newly formed panmictic hybrid population, $X(2)$ is initially positive because of the overrepresentation of parental haplotypes as compared with recombinants. Importantly, $X(2)$ becomes negative if the haplotype frequencies are imbalanced, for example if a lethal DMI causes absence of one recombinant haplotype and overrepresentation of the other. Focusing on the two possible recombinants in the scenario described above, this imbalance is captured by $\Delta D2 = D2_{AB} - D2_{ab}$.

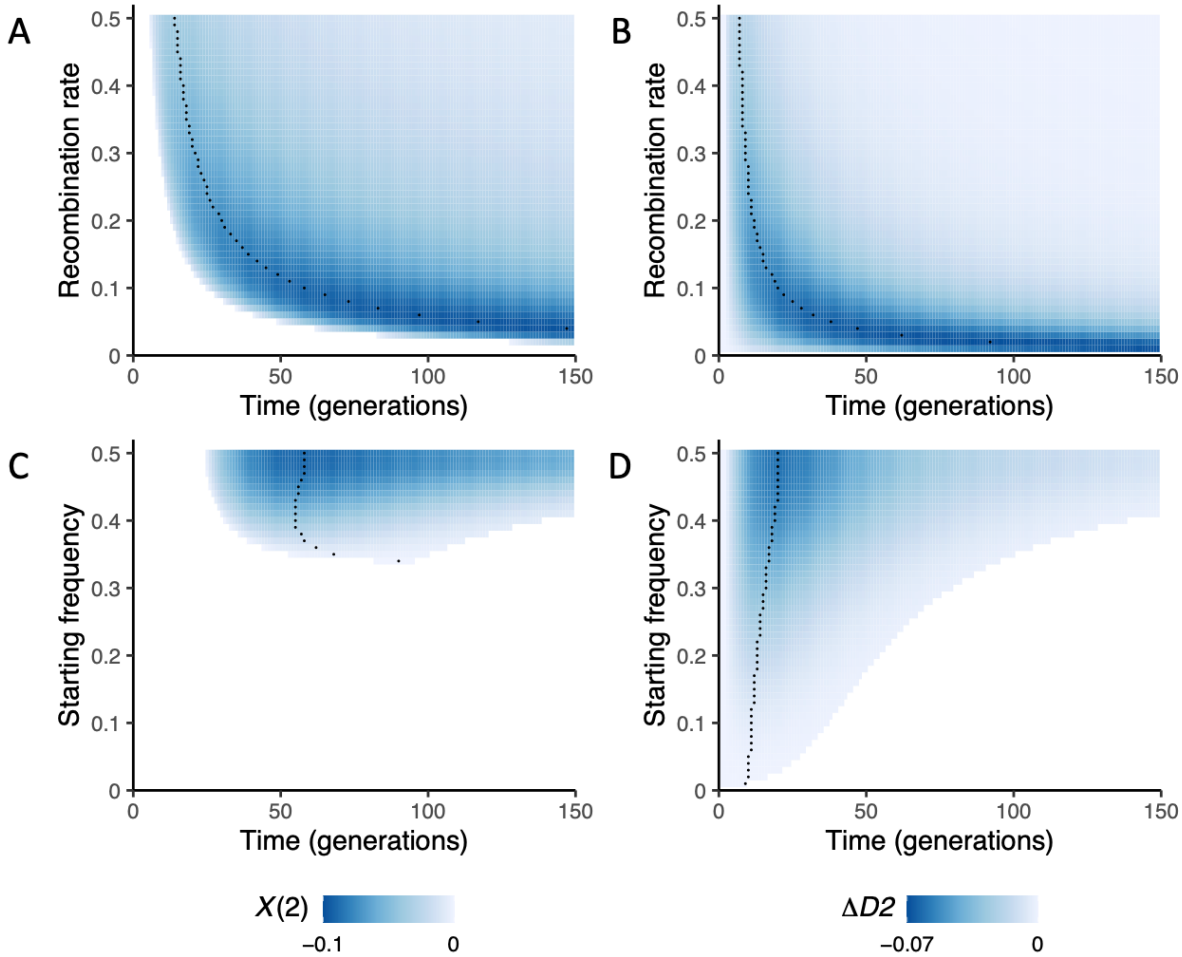

Figure 1. Dynamics of $X(2)$ and $\Delta D2 = D2_{ab} - D2_{AB}$ in the deterministic toy model for different recombination rates and admixture proportions. The model parameters are $\alpha = 0.001, \beta = 0.002, \gamma = -0.5$. Black dots indicate the time with the largest $-X(2)$ or -$\Delta D2$ (i.e., strongest DMI signal) in each trajectory. The values of $X(2)$ and $\Delta D2$ are represented by blue color intensity. A&B. Both $X(2)$ and $\Delta D2$ show more extreme values at intermediate recombination rates. Here, we assumed equal admixture proportions. C&D. Negative $X(2)$ and negative $\Delta D2$ show larger values when admixture proportions are more even. In these two panels, the recombination rate between the two DMI loci is 0.1.

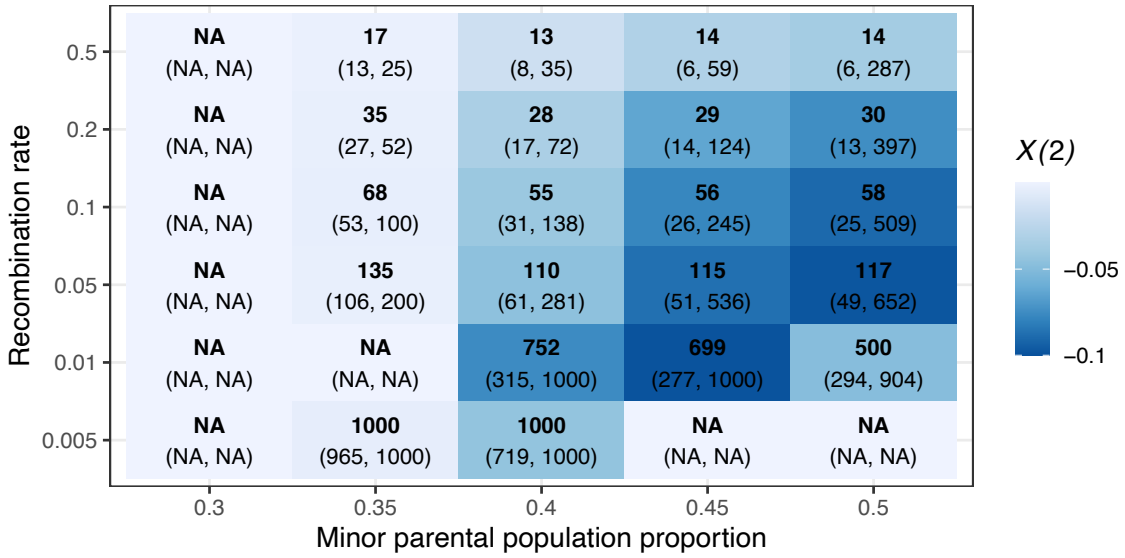

Figure 2. The time window of negative $X(2)$ for a DMI in the deterministic toy model for a range of admixture proportions (x axis) and recombination rates (y axis). Generations with the largest $-X(2)$ values are indicated in bold. Time windows for which negative $X(2)$ at DMIs is observed are shown in parentheses. Values of the largest $-X(2)$ observed over that interval are represented by blue color intensity. The model parameters used here are α= 0.001, β = 0.002, γ = -0.5. All trajectories were traced for 1000 generations. The DMI left the strongest signal in $X(2)$ for intermediate recombination rates, and the signal is not detected until later generations when recombination rates are lower and admixture proportions are more uneven.

Unique dynamics of $-X(2)$ at DMI loci

We compared the dynamics of the $X(2)$ statistic at DMI loci in scenarios where the strength of epistasis was strong ($\gamma = -0.5$) relative to scenarios without epistasis (i.e., $\gamma = 0$). Without epistasis and with weak direct selection for A and B ($\alpha = 0.001, \beta = 0.002$), both the recombinant frequencies and their $D2_{ik}$ values are expected to be close to even and independent of the admixture proportion, because recombinants are generated at equal proportions and because weak directional selection cannot drive significant imbalance in the recombinant haplotypes (Figure S3 A&B). Conversely, with a strong DMI, segregation of the two recombinants is imbalanced, as is evident from the $D2_{ik}$ trajectories of the two recombinants (Figure S3 C&D). Interestingly, without epistasis, $X(2)$ rapidly decreases and converges to zero over time (Figure S3 E&F). At equal admixture proportions, $X(2)$ for a DMI rapidly decreases to negative values and then slowly increases over time, reaching 0 when one of the interacting alleles is lost from the population. At intermediate recombination rates, $X(2)$ for a strongly selected DMI ($\gamma = -0.5$) attains its minimal value after around 60 generations (Figure S3 E). In populations in which admixture proportions are highly skewed, negative $X(2)$ values are never observed, even when DMIs are under strong selection (Figure S3 F). One exception to this observation is when the admixture asymmetry is counteracted by immigration from the minor parental population (discussed below). Overall, negative $X(2)$ can indicate a DMI once the recombinant imbalance caused by selection against the DMI overcomes the residual linkage of parental haplotypes.

The signal of negative $X(2)$ generated by selection against a DMI is most pronounced when the admixture proportions are even. In addition, -$X(2)$ and -$\Delta D2$ are larger at low and intermediate recombination rates (Figure 1 A&B). At the same time, the lower the recombination rate between the interacting loci, the longer it takes for recombinants to emerge, resulting in a later appearance of negative $X(2)$ (Figure 2). In theory, this phenomenon suggests that a DMI at any possible recombination distance will be eventually detectable with our approach. An exception is when two loci have vanishingly small recombination probabilities, such that direct selection and genetic drift undermine the signal of the DMI in $X(2)$, because the DMI is purged from the population (by losing one of the interacting partners) before detectable epistasis-induced recombinant imbalance appears. In practice, variation in recombination probabilities will change the time period during which a DMI can be detected.

Regardless of the recombination rate, our toy model suggests that detecting a DMI in an isolated hybrid population using $X(2)$ requires the initial admixture proportion to be near 0.5 (Figure 1 C&D, minor parental frequency 0.3-0.5). Intuitively, this is because the recombinant imbalance $\Delta D2$ (which is $D \cdot (g_{ab} - g_{AB} + p_a p_b - p_A p_B) \propto D \cdot (\gamma - \alpha\beta)$) reaches its largest value at equal admixture proportions since $D = g_{AB} \cdot g_{ab} - g_{aB} \cdot g_{aB}$ is maximal when $g_{Ab} = g_{aB} = 0.5$. As the recombinant imbalance becomes extreme, $-X(2)$ becomes larger.

We were interested in evaluating whether $-X(2)$ is diagnostic of epistatic selection or sensitive to selection in general. Without epistasis, two alleles *A* and *B* can also generate recombinant imbalance when they are under strong selection in the same direction (by means of Hill-Robertson interference; Felsenstein 1974). To evaluate the dynamics of $-X(2)$ in this scenario, we assumed that *A* and *B* are under strong negative selection and combine additively, such that haplotype *AB* is the least fit (but without epistasis). Here, recombinant imbalance results from selection differences between *AB* and *ab* (Figure S1), which result in negative $X(2)$ (Figure S4C) and $\Delta D2$. However, even with strong selection $X(2)$ remains negative only for a few generations at the very early stage of hybridization (before generation 10, Figure S4C). With a DMI, negative $X(2)$ persists for much longer. The stronger the DMI (i.e., the larger the epistasis parameter γ), the larger $-X(2)$ is observed (after generation 10, Figure 2 and S4A). Altogether, non-epistatic strong selection results in lower maximum values of $-X(2)$ and $\Delta D2$ that occur – and fade – much earlier than in the presence of a DMI (Figure S4).

In simulated data $-X(2)$ robustly indicates intra- and interchromosomal DMIs

To demonstrate that detection of intrachromosomal DMIs in genomic data is feasible, we used simulations with SLiM (Haller and Messer 2019, Figure 3). We simulated evolution of a chromosome in a hybrid population of 5000 individuals that is generated by a pulse of admixture between two parental populations at equal frequencies. Two interacting loci generating a DMI were located at 10 centimorgan (cM) and 20 cM such that the recombination rate between the two loci was 0.1. For these parameters, the deterministic trajectory indicated that a DMI should be detectable within a time window that spanned from generation 25 to 509, with the strongest signal (i.e., largest $-X(2)$) at generation 58 (Figure 2). Figure 3 shows the distribution of normalized *D,* a commonly used measure of linkage disequilibrium, (D', see Methods), $X(2)$ and $\Delta D2$ at generation 50. At generation 50, parental linkage has mostly broken down between pairs of distant loci, resulting in a peak at $D' = 0$. $D' \neq 0$ only within small genetic distances (Figure 3 A&D). Moreover, the distribution of $X(2)$ across all pairs of loci along the simulated chromosome has decreased to near 0 by generation 50 (resulting in a peak of $X(2)$ at 0, Figure 3B). In

general, negative LD ($D = g_{ab} \cdot g_{AB} - g_{Ab} \cdot g_{aB}$) correlates with positive $X(2)$ (Figure 3C). However, the selected pair of loci can be pinpointed by scanning for negative $X(2)$. Negative D' is also observed between the DMI loci (Figure 3D&E). Importantly, whereas the D' signal is consistent with increased LD – that may stem from residual linkage, population structure, or epistatic interaction (Figure 3D) – -$X(2)$ is specifically associated with unexpected recombinant imbalance ($\Delta D2$) caused by the DMI (Figure 3E&F).

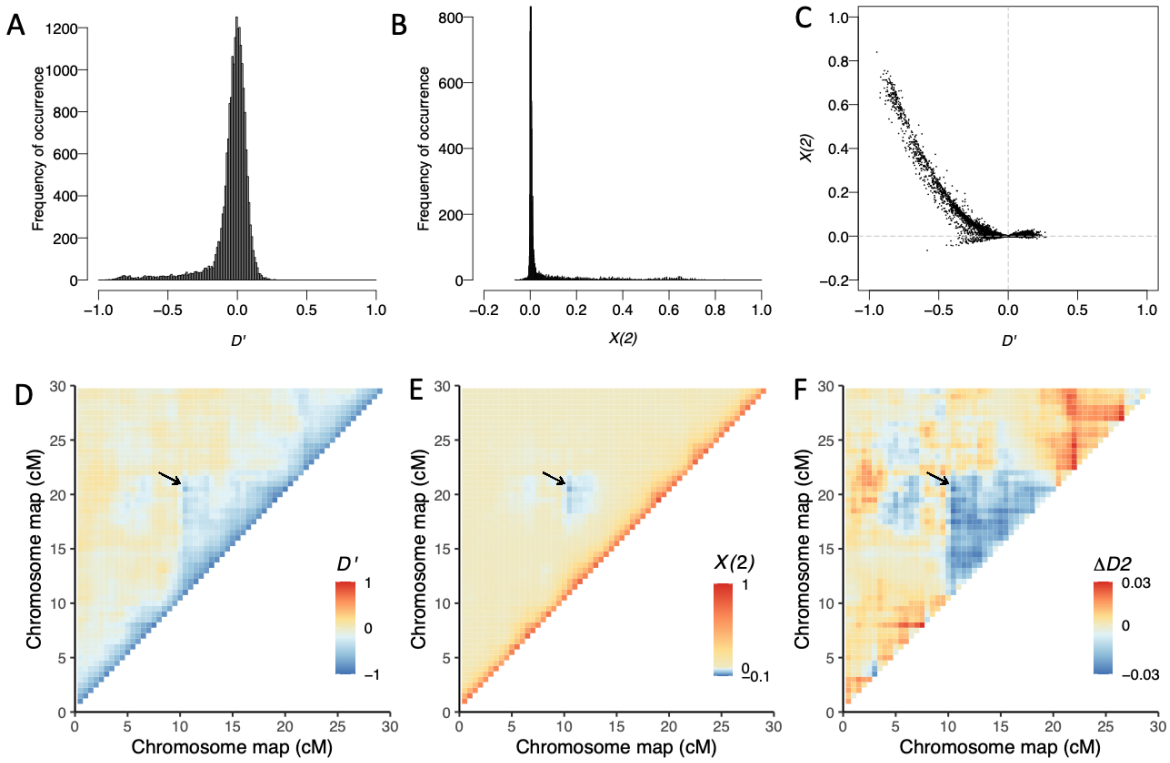

Figure 3. Negative $X(2)$ pinpoints an intrachromosomal DMI in simulated data. The distribution of $X(2)$, D', and $\Delta D2$ of a chromosome with a DMI pair at generation 50 under a Wright-Fisher model for one simulation run using SLiM. The direct selection coefficients are $\alpha = 0.001, \beta = 0.002$. The strength of epistasis is $\gamma = -0.5$. The interacting loci generating the DMI are located at 10 cM and 20 cM, such that the recombination rate is $c = 0.1$ between the two loci. A-C. The distribution of $X(2)$ and D' and their joint distribution within the chromosome. We observe negative $X(2)$ values and negative D' values at the two loci involved in the DMI and in flanking regions that are tightly linked to the DMI. D-F. Heatmaps of D', $X(2)$ and the $\Delta D2$ within the simulated chromosome. Only the DMI region and its 10 cM flanking regions are shown in D-F. The DMI is highlighted by a black arrow.

Since other metrics used to detect DMIs are sensitive to demographic parameters, we next used SLiM to explore whether the same was true for $X(2)$. We considered the effect of two demographic factors, population expansion and migration, on the sensitivity of the $X(2)$ statistic to DMIs. Natural hybrid populations may suffer from low fertility and high mortality in the early stages of hybridization, and the hybrid population may occupy a new niche or geographical area. During the period in which hybrid incompatibilities are purged, mean population fitness increases and the population may expand. To explore the possible effects of this phenomenon, we performed simulations assuming that the hybrid population would expand within 35 generations and arbitrarily assumed a 5% growth rate per generation.

We found that population expansion rarely affects the detection of a DMI (Figure 4) since it does not affect the relative genotype frequency dynamics.

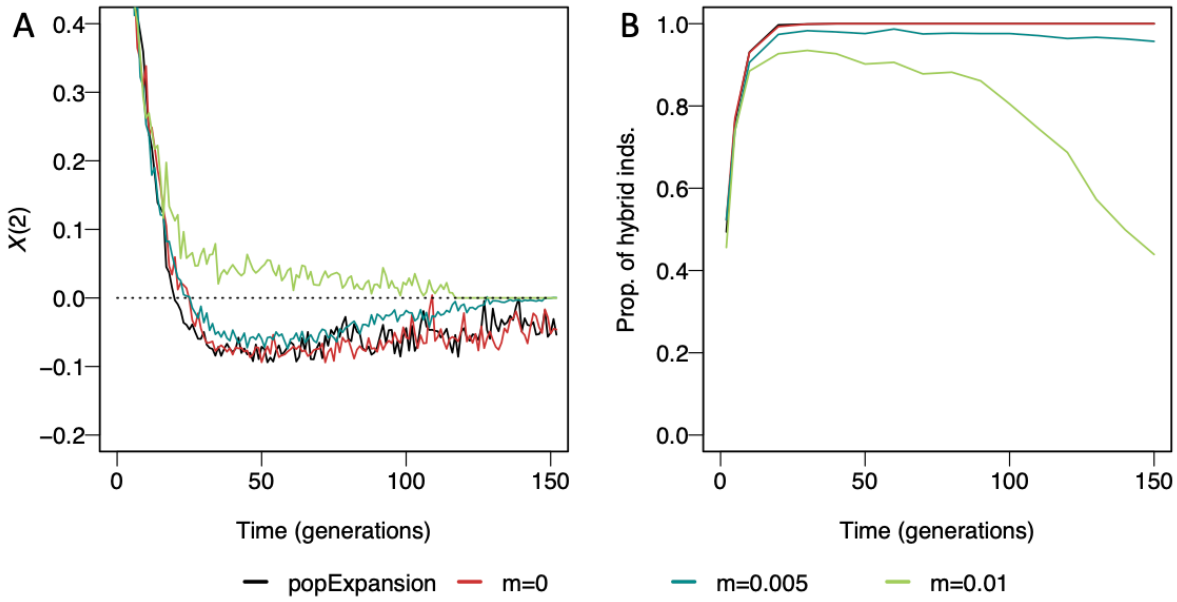

Figure 4. Effects of population expansion and migration from the minor parental population (migration rate, m) on dynamics of $X(2)$ for a DMI, and on the proportion of admixed individuals in a hybrid population for one simulation run using SLiM, assuming equal admixture proportions of the parental populations and a strong DMI ($\alpha = 0.001, \beta = 0.002, \gamma = -0.5$) under a Wright-Fisher model. The recombination rate between the interacting loci was 0.1. Trajectories were plotted from generation 2 onwards A. Lower migration rate and population expansion have a minor impact on $X(2)$ values at the DMI locus pair. Strong migration disrupts DMI detection by $X(2)$. B. High migration rates reduce the proportion of individuals in a population that are genetically admixed. Here, hybrids are defined as individuals that derived at least 10% of their genome from each parental population.

We next considered continual immigration from one of the parental species into the hybrid population. As shown in Figure 4, high migration rates reduce our ability to detect negative $X(2)$ when admixture proportions are even. However, depending on the specific migration scenario, migration can aid in the detection of a DMI. Generally, migration changes the haplotype frequencies and alters the time period during which $X(2)$ is negative (Figure S5). To quantify this effect, we studied three migration rates, no ($m = 0$), intermediate ($m = 0.005$) and strong ($m = 0.01$) migration from the minor parental population into the hybrid population, and two admixture proportions ($f \in \{0.3, 0.5\}$). If the migration rate is weak to intermediate ($m < 0.01$) our simulations suggest that it does not impede the occurrence of negative $X(2)$ in the presence of a DMI (Figure 4A). Importantly, we find that migration from the minor parent can increase the probability of detecting a DMI with $X(2)$. Whereas we had poor power to detect a DMI in populations with an admixture proportion of 0.3 without migration, power improves substantially with migration. If the minor parental population contributes most migrants, incompatible haplotypes will be continuously formed by recombination between resident and immigrant types. This can result in the DMI persisting in the hybrid population for a longer time period, which makes it possible to use $X(2)$ to detect

DMIs in scenarios of uneven admixture proportions (Figure S5). We note that immigration of the major parental population, in turn, is expected to reduce our ability to detect a DMI.

With strong migration ($m = 0.01$), the linkage of parental haplotypes remains strong due to the continuous reintroduction of these haplotypes. In such a scenario, $X(2)$ remains larger than 0 and thus makes it difficult to detect DMIs with this statistic (Figure 4A). Here, strong migration acts like strong selection for the immigrating parental two-locus haplotype, which hides the (still existing) recombinant imbalance. This will eventually result in one parental haplotype taking over the population (Figure 4B).

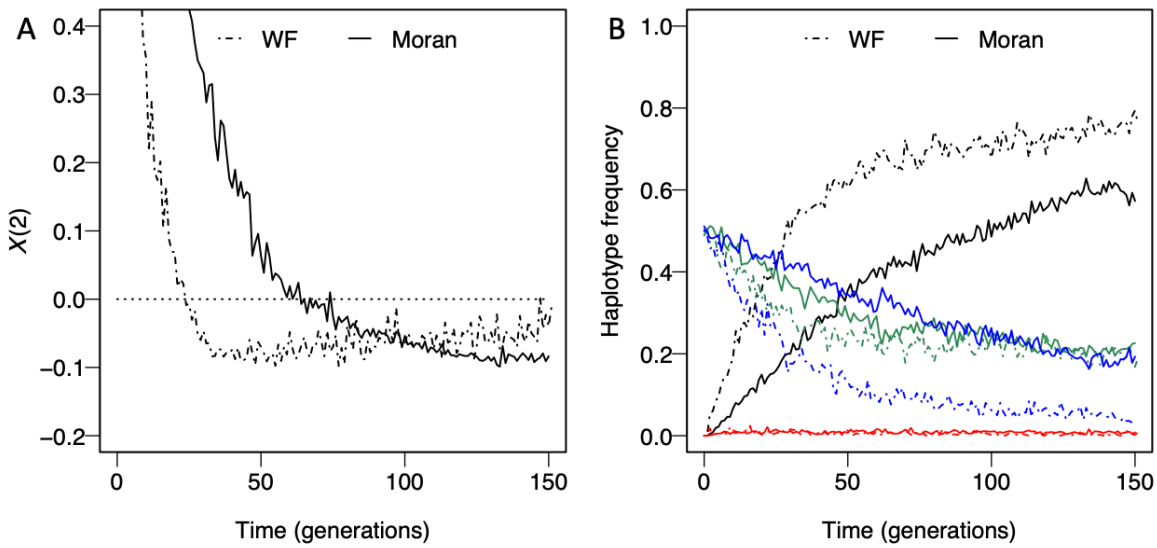

Figure 5. Dynamics of haplotype frequencies and $X(2)$ for a DMI under diploid Wright-Fisher (WF; dash-dotted line) and Moran (solid line) models for one simulation run using SLiM, assuming a population size of 5000. The model parameters were $\alpha = 0.001, \beta = 0.002, \gamma = -0.5$, with recombination rate $c = 0.1$ between the DMI loci. A. Negative $X(2)$ appears ~50 generations earlier under a Wright-Fisher model than under a Moran model. B. Haplotype frequencies change more quickly in the Wright-Fisher model than that in the Moran model. Lines indicate the haplotype frequencies of *AB* (red), *ab* (black), *Ab* (blue) and *aB* (green).

Next, to evaluate the performance of $-X(2)$ in a natural population with overlapping generations, we compared our results from Wright-Fisher simulations with those obtained under a Moran model. With overlapping generations, the emergence of negative $X(2)$ is delayed because recombinant imbalance builds up more slowly than in a model with non-overlapping generations (Figure 5A). At the same time, alleles *A* and *B* segregate in the population for much longer with overlapping generations than with non-overlapping generations (Figure 5B). This leads to a longer maintenance time of the interacting loci, resulting in a longer window in which $-X(2)$ can be used to detect the DMI.

Given these results, we next evaluated the sensitivity and specificity of $X(2)$ for DMI detection under a range of models using non-overlapping generations (Figure 6). We simulated four chromosomes, two of which contained the DMI loci and two of which evolved neutrally. We computed sensitivity as the

proportion of simulations in which the DMI loci showed $X(2) < -0.005$ & $D' < 0$, and specificity as one minus the proportion of simulations in which a pair of loci on the neutral chromosomes showed $(2) < -0.005$ & $D' < 0$. We considered six scenarios by combining two admixture proportions ($f = 0.5$ and $f = 0.3$) and three immigration rates (0, 0.005, 0.01) of the minor parental population. We also compared the performance of $X(2)$ to the G test which has previously been used to detect DMIs (Li and Zhang 2013). The recombination rate in these simulations was set to $c = 0.5$ because the G test can only be applied to unlinked loci.

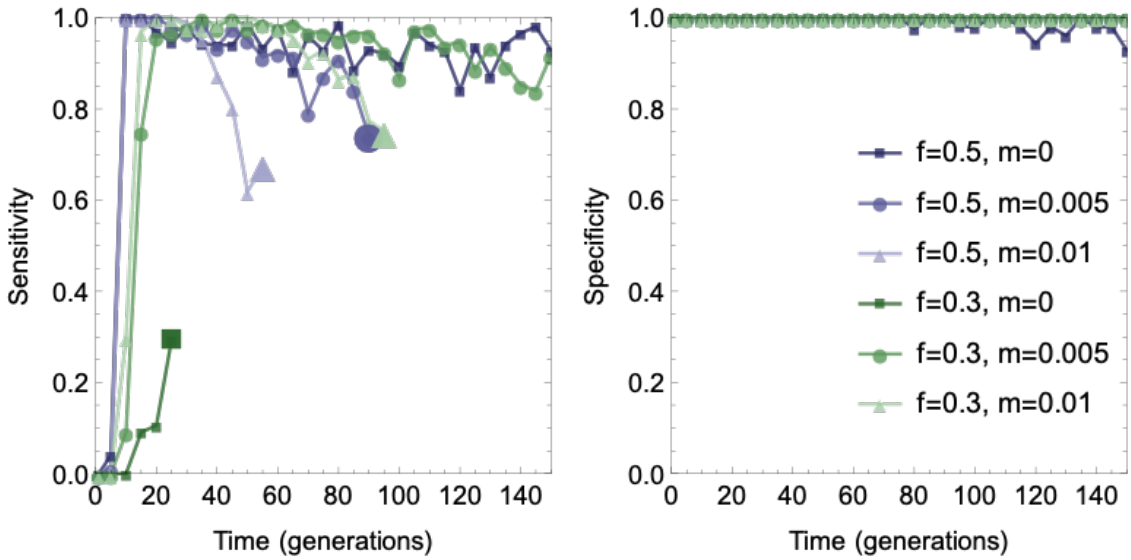

Figure 6. Sensitivity and specificity of negative $X(2)$ for detecting a DMI, computed from 100 simulations of each parameter combination using SLiM, assuming a Wright-Fisher population of 5000 individuals. We considered two admixture proportions, $f = 0.5$ and $f = 0.3$, and three migration rates: 0, 0.01, and 0.05. A strong DMI ($\alpha = 0.001, \beta = 0.002, \gamma = -0.5$) and no DMI ($\alpha = 0.001, \beta = 0.002, \gamma = 0$) were simulated under a Wright-Fisher model. The recombination rate between the DMI loci was *c=0.5*, and the population size was 5000. Here, any interaction with $X(2) < -0.005$ & $D' < 0$ was taken as a putative DMI. Trajectories end when one of the interacting loci became monomorphic (indicated by large plot markers). 300 individuals were sampled in each simulation. A. Sensitivity is high unless strong migration leads to swamping of the DMI alleles. B. Specificity is high under all tested scenarios when no DMI is present on the sampled chromosome.

In simulations of hybrid populations with non-overlapping generations, equal admixture proportions, and no or moderate levels of ongoing migration, $X(2)$ allows for detection of the DMI with high sensitivity and specificity (Figure 6, $f = 0.5$, $m = 0$ or 0.005). At high migration rates ($m = 0.01$), migration from the parental population dominates the haplotype frequency dynamics and the DMI cannot be detected with equal admixture proportions. In contrast, immigration can maintain the minor parental alleles when the admixture proportions are uneven ($f = 0.3$). With continuous input of the minor parental haplotype, the DMI remains in the population for longer, so the DMI can be detected based on negative $X(2)$ values before one of the parental haplotypes is lost. Thus, $X(2)$ can be used to detect DMIs in the case of uneven admixture proportions and intermediate migration rates, when immigration counteracts selection against the DMI in the minor parental population. In all scenarios explored here, $X(2)$ has very high specificity.

305 In 300 simulations of a null model of neutral evolution, $X(2)$ was always greater than -0.001, where
306 negative $X(2)$ was caused by genetic drift and random sampling of 300 individuals for the analysis.

307 We compared the performance of $X(2)$ with that of the G test, which was identified as one of the best
308 statistics to detect interchromosomal DMIs (Li and Zhang 2013). Figure S6 shows that the G test is highly
309 sensitive in the early stages of hybridization, but its sensitivity decreases after generation 20. There is also
310 an apparent tradeoff between sensitivity and specificity. Interestingly, unlike for the G test there is no
311 tradeoff between sensitivity and specificity for $X(2)$. We found that whereas significant hits using the G
312 test had negative $X(2)$, the opposite was not true, especially when the parental linkage has broken down
313 from generation 20 onwards (Figure 6, S6 and S8). Thus, $X(2)$ was sensitive to DMIs in scenarios in
314 which the G test was not.

315 Our analysis above relies on the use of phased genotyping data, which can be difficult to collect in
316 practice and can introduce inaccuracies via phasing errors. In contrast, pseudo-phased data is available for
317 many datasets, including the data analyzed below. In this case, only homozygous loci in each genome are
318 used for analyses where the phase can be inferred without computational phasing (e.g., genotyping data
319 reviewed in Davey et al. 2011). We calculated the sensitivity and specificity for the G test and $X(2)$ using
320 only homozygous genotypes and found that performance trends were similar to phased data with mildly
321 reduced sensitivities (Figure S8).

322 We next applied our detection method to simulated genomic data from a range of more general scenarios
323 with one and two DMIs. Here, positions of the interacting loci were distributed randomly across four
324 chromosomes (resulting in both inter- and intrachromosomal DMIs), epistatic coefficients were drawn
325 from a uniform distribution ($\gamma \in [-1, -0.001]$), and the direct selection coefficients α and β were drawn
326 from an exponential distribution with mean 0.001. We further simulated two neutral chromosomes
327 without DMI loci.

328 The proportion of pairs of loci that were incorrectly inferred as interacting partners (from here on referred
329 to as false positive rate) varied significantly between chromosomes with DMIs and neutral chromosomes.
330 For the neutral chromosomes, this proportion was almost 0 when $X(2) < -0.005$ was used as a
331 threshold (Table S2 & S3) for both phased and pseudo-phased data. However, on chromosomes with
332 DMIs, upwards of 2% of pairwise comparisons (from a total of $\binom{400}{2} = 79800$ pairs, of which 2 pairs are
333 true DMI pairs, corresponding to 0.005% of pairs) were identified as putative DMIs based on the $X(2)$
334 statistic (threshold $X(2) < -0.005$) although the two loci were not involved in a DMI. As expected, both
335 true and false positive rates decreased with a more stringent $X(2)$ threshold (Table S1-S3). Comparing to
336 simulations with a single DMI, the false positive rate increased noticeably but not excessively when two
337 DMIs were simulated (from ~1% to a maximum of 2.2% when $X(2) < -0.005$, Table S1 and S2),
338 indicating some interference between DMIs. When only homozygous loci (pseudo-phased data) were
339 used for the inference, the true positive rate remained similar, but the false positive rate was elevated to
340 3.6-6.1% (Table S2 and S3).

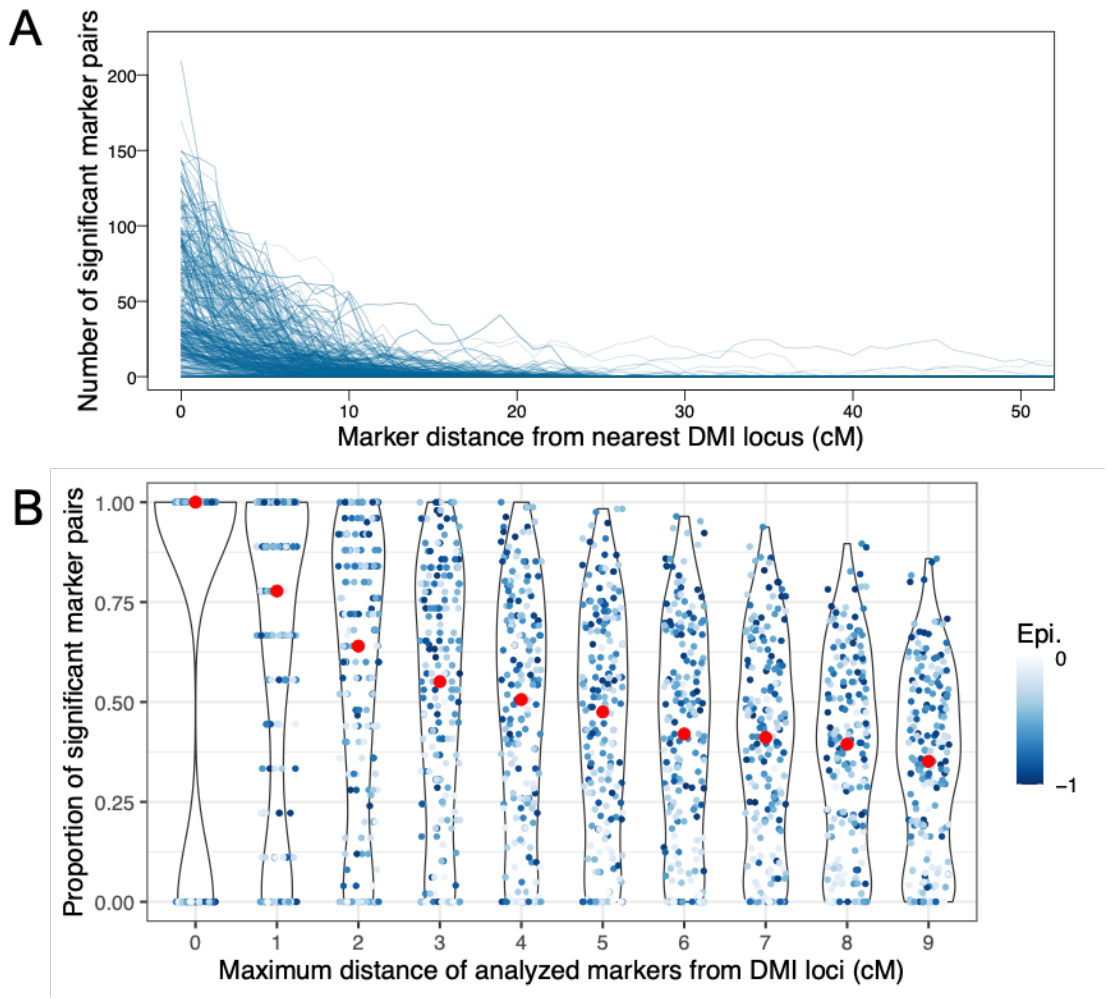

Figure 7. Propagation of the DMI signal into flanking regions in simulated data (sampled at generation 30, significance threshold: $X(2) < -0.005$ & $D' < 0$). Data resulted from 100 simulations using SLiM, in which two DMIs (i.e., four DMI loci) were randomly distributed on four chromosomes, thus resulting in inter- and intrachromosomal DMIs. The direct selection coefficients α and β were drawn from exponential distribution with mean = 0.001. The strength of epistasis was drawn from a uniform distribution, $\gamma \in$ [-1, -0.001], which are represented by the blue color intensity in panel B. To resemble the demography of the studied hybrid swordtail fish populations, the admixture proportion was set to $f = 0.3$ and the migration rate was $m = 0.01$. 200 individuals were sampled from each simulation to obtain the statistics. A. The number of significant marker pairs decreases rapidly with distance to the nearest true DMI locus. B. Proportion of significant marker pairs when statistics were computed for all pairs between two increasingly large regions around the true interacting DMI loci. Red dots indicate the median obtained from 100 simulations. Both weak and strong DMIs can be detected, but weak DMIs tended to leave a weaker signal that did not extend as far as the signal of strong DMIs.

False positives were most often caused by loci linked to DMIs (e.g., Figure 3E). Markers within 10 cM of DMI loci showed a large number of significant $X(2)$ pairs in simulations of two DMIs that were distributed randomly across four chromosomes (Figure 7A). In the most extreme case observed in the simulations, one DMI locus was involved in significant pairwise interactions with over 200 markers of a total of 400 markers (Figure 7A). This case involved strong epistasis for both DMI pairs ($\gamma \approx -0.75$),

360 and two DMI loci that were located very close to each other on the same chromosome (DMI 1 between
361 chromosome 1 (chr1) position 63 and chr4 position 100; DMI 2 between chr3 position 98 and chr4
362 position 92, with the extreme signal emitted from chr3 position 98). Thus, the extreme signal was likely
363 caused by interference between the two DMIs. At generation 30, the genomic window across which the
364 signal of interaction between two true DMI loci extends is large; the proportion of significant pairs when
365 only true DMI loci and their flanking regions were included in the calculation is high within 10 cM
366 (median ~35% vs. genomic median 1.1%, Figure 7B and Table S2), probably caused by hitchhiking or
367 sorting of large genomic blocks in hybrid populations due to the limited generation time and genetic drift.

368 In analyses of unphased genomic data, we can only rely on information from homozygous genotypes.
369 Therefore, we computed the same statistics for simulated data when only homozygous genotypes are
370 considered. Here, consistent with a higher false positive rate, the proportion of significant pairs between
371 DMI regions decreased more slowly with window size than when the complete data were considered
372 (Figure S9C; see also Figure 7B). We then applied the strict filtering criteria necessary to clean the
373 genomic data from natural hybrid populations before analysis (see below and Methods), for which a
374 homozygous frequency of >0.05 and a frequency of heterozygotes <0.6 were required. Because most
375 strong DMIs are (nearly) purged by generation 30, few DMI loci pass the filtering, and the signal
376 therefore changes such that the highest proportion of significant pairs is observed for genomic windows
377 that span several cM around the true DMI (Figure S9 B&C). Here, only the flanking regions of strong
378 DMIs (and not the DMI loci themselves) indicate the DMI. The proportion of significant pairs at up to
379 10cM distance from a true DMI locus is still higher than the genome-wide false positive rate (>35% vs.
380 genomic median up to 7.1%, Figure S9B and Table S3). Taken together, our analysis indicates that,
381 although greater precision could be obtained from phased data, our method shows good results for
382 pseudo-phased data.

383 **Using $X(2)$ *to detect putative DMIs in swordtail fish hybrid populations***

384 Two sibling species of swordtail fish, *Xiphophorus birchmanni* and *X. malinche*, were reported to
385 hybridize naturally in Mexico (Culumber et al. 2011, Schumer et al. 2014). Since these hybrid
386 populations formed recently and have been studied extensively, they represent an excellent opportunity
387 for using $X(2)$ and $\Delta D2$ to infer DMIs from natural populations. Three hybrid populations called CHAF,
388 TOTO and TLMC were sampled and genotyped in previous work (Table S4). All three hybrid
389 populations are estimated to have formed less than 100 generations ago: CHAF ~45 generations; TOTO
390 ~35 generations; TLMC ~56 generations (Schumer et al. 2014 and Powell et al. 2020). However, these
391 admixture times may be underestimated due to ongoing migration (Schumer et al. 2017, 2018, Powell et
392 al. 2020). The estimated proportion of the genome derived from the minor parental population was ~0.25-
393 0.3. The minor parental species for the CHAF and TLMC populations is *X. birchmanni* and for TOTO it
394 is *X. malinche*. All three hybrid populations are known to experience some level of ongoing migration
395 from the minor parental population (CHAF, ~5%; TOTO, ~3%; TLMC, <<1% (Schumer et al. 2017 and
396 Powell et al. 2020). Under the demographic parameters of these natural populations, we expect that
397 immigration counterbalances the asymmetric admixture proportions such that we can infer DMIs with
398 intermediate sensitivity and high specificity (Figure S7).

399 We used $X(2)$, $\Delta D2$ and D' to identify putative DMI pairs and to infer the most strongly depleted
400 genotypes in these three populations. To roughly count the DMIs in a fashion that is comparable between

401 populations, we sliced the genome into fragments of 1 Megabase (Mb), which we will refer to as a locus
402 in the following description. A putative DMI was identified if homozygous markers at this locus showed
403 $X(2) < -0.005$ and $D' < 0$ (see also Methods). Overall, the proportion of pairs of loci that fulfilled these
404 criteria was similar to the proportion observed in simulations of one or two DMIs (Table S3 and S5). To
405 increase our confidence in the candidates, we therefore performed bootstrapping to ensure that the locus
406 pair was consistently inferred to have $X(2) < -0.005$ and $D' < 0$ (Table S5, see methods).

407 In total, we detected 2203, 2039 and 1442 putative DMIs in the three hybrid populations CHAF, TLMC
408 and TOTO with false discovery rate <0.001 (15.2%, 12% and 6% of tests, respectively, Table S5). As
409 highlighted in the previous section, these numbers may include many genomic blocks that are close to
410 DMIs but not themselves carry a DMI locus (Figure 7 and S9, Table S5). Given the false positive rate
411 estimated from the simulations, we view this as a preliminary set of candidate DMIs for further
412 investigation. In each population, 200-400 bins (out of the total 711 bins, Table S5) were involved in at
413 least one putative DMI. If the interactions were randomly distributed across the genome, the occurrences
414 of the loci would fit a Poisson distribution ($\lambda = 2 \cdot$ number of interactions / number of loci involved in
415 interactions). In contrast to this expectation, loci with very small and very large numbers of occurrences
416 were greatly overrepresented among the DMI candidates (One-sample Kolmogorov-Smirnov test, $p <$
417 $2.2e - 16$, Figure S9). We speculate that the loci that appear in many putative DMI pairs are likely to be
418 true interaction partners in pairwise or complex DMIs, whereas many of the loci that occur only once or
419 twice may be noise caused by the true DMI loci. More than 99.5% of putative interactions involve a small
420 number of loci that are each involved in >10 interactions (54 in CHAF, 61 in TLMC and 66 in TOTO).
421 Among these putative interaction hubs, CHAF and TLMC share 8 loci, CHAF and TOTO share 2, and
422 TLMC and TOTO share 2 (Table S6). Altogether, these results strongly imply that a set of loci is
423 involved in complex interactions in the hybrid

424 Although our overall false positive rate is low, given the large number of tests performed in our
425 evaluation of the real data, we can be most confident in putative DMIs that are detected in multiple
426 populations. Across three hybrid populations we found 17 putative DMI pairs that were shared between
427 TLMC and CHAF, which are the two populations that have the same minor parental population (*X.*
428 *birchmanni*) and are therefore expected to share more detectable DMIs (Table S7). There were no shared
429 DMI pairs between the TOTO and the two other populations. Among all putative interaction partners, 14
430 loci were shared by three populations, out of which four were previously identified (Table S8). This set of
431 loci are exciting candidates for future work.

432 In comparisons between TLMC and CHAF, which share a minor parent species, two loci stood out with
433 respect to their many inferred interaction partners. Both chr20:23Mb (specifically, between 23Mb and
434 23.4Mb) and chr15:15M (specifically, between 15.6Mb and 16Mb) are putatively involved in 8
435 interchromosomal interactions. At both loci, alleles from *X. malinche* are incompatible with alleles at
436 several loci from *X. Birchmanni* (Table S7). Interestingly, there was no signal of an interaction between
437 these two loci. However, chr18:19Mb is putatively interacting with both chr20:23Mb and chr15:15Mb in
438 both TLMC and CHAF. *X. malinche* is the major parental population in TLMC and CHAF, leading to the
439 expectation that homozygous *X. malinche* alleles would be in the majority. In contrast to this expectation,
440 few alleles from *X. malinche* remained in these regions when combined with several *X. birchmanni*
441 backgrounds. In TOTO, where *X. malinche* is the minor parental population, many *X. malinche* alleles
442 were already lost from the population or are segregating at very low frequencies. This leads to a power

443 reduction for detecting *X. malinche*-specific interacting partners hinting why the two strong DMI
444 candidates may not be detectable in TOTO. For example, at locus chr20:23M, frequencies of
445 homozygous *X. malinche* genotypes were found at a median frequency of 0.004, which is much smaller
446 than our filtering criteria 0.05. No DMIs remained detectable in TOTO at locus chr15:15M, at which the
447 frequencies of homozygous *X. malinche* genotypes were found at a median frequency 0.009. Together,
448 our results point at chr20:23Mb and chr15:15Mb as strong DMI candidates at which *X. malinche* alleles
449 are deleterious in combination with certain (and possibly several) loci from *X. birchmanni*.

450 Out of all putative interactions identified in the swordtail hybrid populations, about 4% were
451 intrachromosomal (CHAF: 81/2203; TLMC: 67/2039; TOTO: 55/1442). Notably, none of the inferred
452 intrachromosomal interactions overlapped between any two populations. For example, Figure S11 shows
453 a putative intrachromosomal DMI in TLMC (chr20:9M hybrid and chr20:10M). In the $\Delta D2$ heat plot of
454 this genomic region, the candidate recombinant from TLMC (composed of *X. birchmanni* ancestry at
455 chr20:9M and *X. malinche* ancestry at chr20:10M) also showed a weak signal of elimination in CHAF.
456 However, in TOTO, no such recombinant imbalance is visible. The lack of overlap could stem from the
457 $X(2)$ signal of intrachromosomal DMIs appearing later than that of interchromosomal DMIs, combined
458 with the different timescales of hybridization in the three populations, differences in power, or true
459 differences in the frequency of intra versus interchromosomal DMIs.

460 In addition to comparing our candidates with a list of previously identified DMIs (Table S6 and S8), we
461 specifically analyzed genomic regions involved in a well-studied DMI that is known to induce melanoma
462 in hybrids. Powell et al. recently found a DMI between chr5:10-11Mb (*cd97*) and chr21:16.5-17.5Mb
463 (*Xmrk,* Powell et al. 2020). A previous QTL assay predicted that another gene, *cdkn2a/b* (~chr5:16-17Mb;
464 reviewed in Patton et al. 2010) interacted negatively with *Xmrk*. Previous work studying the interaction
465 between *Xmrk* and *cd97* has focused on CHAF, where melanoma frequency is high (interactions
466 generating melanoma appear to have been purged in the other hybrid populations). Interestingly, in
467 CHAF, it appears that both homozygous hybrid haplotypes are depleted in a part of this region. Thus, the
468 *Xmrk-cd97* interaction could be a hybrid incompatibility in which both recombinant haplotypes are
469 deleterious, in which case no negative $X(2)$ signal is expected (since the signal is driven by recombinant
470 imbalance as caused by classical DMIs). Regarding the potential DMI between *Xmrk* and *cdkn2a/b*, we
471 found negative $X(2)$ in all three hybrid populations that support this interaction (Figure S12, Table S9). In
472 CHAF, $X(2)$ of 49.5% of marker pairs was negative but only 0.2% of these passed the threshold $X(2) <$
473 -0.005. In TOTO, which is of intermediate age, 25.9% of all marker pairs showed $X(2) < -0.005$,
474 which is much higher than the expectation obtained for a DMI from bootstrapping (median=5.7%, Table
475 S5). This is consistent with our expectation of the signal of an intermediate-strength DMI that does not
476 display strong negative $X(2)$ until later generations.

477 **Discussion**

478 DMIs play an important role in maintaining reproductive isolation between species. Despite their crucial
479 importance, existing methods to identify DMIs are severely underpowered and susceptible to high false
480 positive rates. Here, we show that the deviation of the variance of heterozygosity from its expectation,
481 $X(2)$, can allow us to identify intra- and interchromosomal DMIs in hybrid genomes. Specifically, a
482 negative $X(2)$ statistic is indicative of the presence of a DMI, because epistatic selection transiently leads
483 to an imbalance of the recombinant haplotypes. A second measure that we develop, $\Delta D2$, quantifies this

484 recombinant imbalance more explicitly. After showing the expected sensitivity and specificity of these
485 statistics in simulations, we used $X(2)$, $\Delta D2$ and D' to infer candidate DMIs from three hybrid
486 populations of swordtail fish. Although the detection power with these statistics is strongly dependent on
487 the age and demographic history of the hybrid population, false positive rates across simulations of
488 various scenarios were low and are expected to be especially low among interactions detected in multiple
489 populations. Encouraged by this result, we found that natural hybrid populations of swordtail fish
490 harbored many potential DMIs, and our analysis detected known DMI pairs. Interestingly, putative DMI
491 pairs tended to include the same loci more often than expected by chance, which is suggestive of the
492 presence of interaction hubs in these data.

493 ***DMIs cause distinctive patterns of LD and variance of heterozygosity***

494 Linkage disequilibrium (LD) between physically unlinked loci has been frequently used to identify
495 candidate DMIs. However, LD is influenced by many factors including selection, population structure,
496 population differentiation, genetic drift, and epistasis. Sved first proposed the variance of heterozygosity
497 as a measure of multi-locus LD (Sved 1968). Generally, LD tends to raise the variance in heterozygosity
498 above its expected value at linkage equilibrium. This deviation from the expectation (called $X(n)$, where
499 n is the number of loci considered) was used to infer population structure within species, e.g., in natural
500 populations of *Hordeum spontaneum* and bacteria (Brown et al. 1980, Maynard Smith et al. 1993). These
501 studies assumed that the focal loci are independent, and that any disruption of independence would cause
502 LD and a positive deviation from the expected variance of heterozygosity (positive $X(n)$). In this paper,
503 we show that selection against DMIs generates an exceptional pattern of LD and $X(2)$, where LD is
504 strong but $X(2)$ becomes negative.

505 ***DMI inference with $X(2)$ does not rely on recombination rate estimates***

506 One advantage of our statistics is that they do not rely on knowledge of recombination rates, which are
507 often unknown or estimated with error and may also differ between the parental and the hybrid
508 populations. For example, the protein PRDM9, which is involved in hybrid incompatibility in mice
509 (Mihola et al. 2009), specifies the locations of recombination hotspots in many mammalian species
510 (reviewed in Smukowski and Noor 2011, Penalba and Wolf 2020). Traditional methods of DMI detection
511 are strongly affected by recombination rate variation, whereas in the case of $X(2)$ it solely affects the
512 temporal scale over which DMIs are detected. At the same time, incorporating recombination rate
513 information into DMI detection with $X(2)$ is an exciting direction for future work. That is because if the
514 recombination map was known, our statistics could be incorporated in a simulation-based framework to
515 estimate not only the location but also the strength of selection against the DMI (i.e., the size of the
516 epistasis coefficient γ).

517 ***Detection time windows differ between inter- and intrachromosomal DMIs***

518 The ability to detect a DMI based on negative $X(2)$ and $\Delta D2$ statistics is transient, and the time window
519 during which a DMI is detectable with these statistics depends on the demographic history of the
520 population, as well as the strength of the incompatibility and the recombination rate between the two loci
521 involved (Figure 2). Specifically, the less recombination there is between two DMI loci and the weaker
522 the incompatibility, the longer it takes for their recombinant imbalance to be exposed. Moreover, since an
523 intrachromosomal (or weak interchromosomal) DMI is purged more slowly, it is detectable for a longer

524 time period, unless one of the interacting alleles is lost by other forces (such as genetic drift). This
525 interaction between population history and power to detect a DMI must be considered when interpreting
526 $X(2)$ results from real data. Interestingly, we find that certain migration scenarios improve power to
527 detect DMIs over long time periods.

528 ***No evidence for an enrichment in intrachromosomal DMIs in swordtail population data***

529 Two lines of evidence led us to hypothesize that intrachromosomal DMIs may be more prevalent than
530 interchromosomal DMIs. From first principles, systems and molecular biology have established that
531 genes in the same biological pathway tend to cluster physically in the genome, and DMIs may be more
532 likely to arise from disruption or modification of such pathways (reviewed in Hurst 2004). This predicts
533 an enrichment in intrachromosomal DMIs under any scenario of speciation. Secondly, evolutionary
534 theory has predicted that tightly linked DMIs are more likely to accumulate and be maintained in the
535 presence of gene flow (Bank et al., 2012). Thus, both factors should contribute to an enrichment (or
536 reduced purging) of intrachromosomal DMIs. In contrast, in the set of DMI candidates inferred from
537 empirical data in this paper, interchromosomal DMIs were slightly overrepresented as compared to
538 intrachromosomal DMIs (<4% intrachromosomal DMIs compared with the expectation of 4.2% given an
539 average chromosome length 29.2 ± 4.63 Mb; permutation test, $p < 0.02$). However, given so many
540 unknowns about the true architecture of selection, we are currently not able to conclude whether this is a
541 biological result, or a pattern caused by the limits of our inference method combined with the
542 demographic history of the populations studied here. Although our method is the first that can be applied
543 to detect intrachromosomal DMIs, it is limited to pairs of loci that are at a sufficiently large distance on
544 the chromosome such that recombination occurs with appreciable frequency, thus making the
545 recombinants visible to selection. Therefore, we are likely to miss local interactions, such as those
546 between neighboring genes. Moreover, intrachromosomal DMIs are detected on a different timescale than
547 interchromosomal DMIs, making the two classes not directly comparable. Altogether, our analysis does
548 not allow for a firm statement about the true proportion of intra- to interchromosomal DMIs.

549 ***The challenge of inferring and understanding complex DMIs***

550 We find based on simulations that our statistics are robust to various demographic scenarios such as
551 population expansions and overlapping generations. However, in the modeling and simulation parts of
552 this paper, we mainly study the dynamics of one or two DMI pairs. This is because considering multiple
553 or complex DMIs would generate an intractably large parameter space to explore. Thus, we implicitly
554 assume that DMIs are independent of other incompatible loci in the genome. However, our results
555 indicate that some DMI candidates are involved in many interactions. For example, one locus may
556 interact with multiple loci (e.g., Brideau et al. 2006, Tang et al. 2015). Moreover, DMIs generate fitness
557 costs in hybrids, but these costs can express themselves through different fitness-related phenotypes. This
558 makes it difficult to study DMIs in the same way one would study a quantitative trait. For example,
559 previous studies suggest that *cdkn2a/b* and *Xmrk* interact in some species of swordtail fish (reviewed in
560 Patton et al. 2010). However, Powell et al. 2020 recently showed that *Xmrk* interacts with a previously
561 unidentified locus (*cd97*) to cause melanoma in other swordtail fish species (Powell et al. 2020). Our
562 results also indicate an interaction beween *Xmrk* and a genomic block with *cdkn2a/b* in swordtail fish,
563 which might cause unfit phenotypes other than melanoma. Lu et al. 2020 reported that *rad3d*, a gene
564 upstream of *cdkn2a/b,* interacted with *Xmrk* mapped in hybrids between *X. maculatus* and *X. helleri*.

565 Taken together, this suggests the simultaneous presence of multiple hybrid incompatibilities caused by
566 *Xmrk*. This finding is not specific to the study system; recent GWAS and QTL mapping studies in a house
567 mouse hybrid zone also suggested multiple, non-independent genetic incompatibilities (Turner and Harr
568 2014, Turner et al. 2014). In addition to the presence of true hubs of hybrid incompatibilities, other types
569 of selection and interference could contribute to the signal at these loci. For example, there could be
570 interference between different DMIs (analogous to Hill-Roberson interference), or strong negative
571 interactions may cause local stratification of the genome (e.g., Pool 2015).

572 Moreover, it is important to note that the sensitivity of the $X(2)$ statistic is specific to the recombinant
573 imbalance caused by classical DMIs (Orr 1995, Gavrilets 1997), where only one recombinant haplotype
574 suffers from the incompatibility, whereas the other has similar fitness to the parental genotype. The
575 evolution of this type of hybrid incompatibility is easy to explain without invoking multi-locus
576 interactions (Orr 1995). However, hybrid populations may carry hybrid incompatibilities beyond classical
577 DMIs, where both recombinant haplotypes suffer from reduced fitness (e.g., Moran et al. 2021). These
578 tend to be less easily purged by gene flow (Bank et al 2012, Lindtke and Buerkle 2015, Blanckaert et al
579 2020), and they may cause distinct patterns in the genome such as inflated single-locus heterozygosity. To
580 date, we are far from understanding the relative prevalence of hybrid incompatibilities of different types
581 and their overall role during hybridization and speciation.

582 ***Concluding remarks***

583 In this paper, we take a step towards genome-wide identification of inter- and intrachromosomal DMIs,
584 providing a powerful new tool for researchers in speciation genetics. However, the complexity of
585 potential interactions in the genotype-fitness map of an organism is vast (Fragata et al. 2019), and DMIs
586 are only one specific kind of pairwise epistatic interaction that is characterized by the fitness deficit of
587 one recombinant genotype. Even if this type of (negative) epistasis between two loci is inferred, drawing
588 conclusions about the functional underpinnings of this epistasis is difficult. For example, Kryashimsky
589 (2021) recently presented theoretical models that demonstrate how epistasis can propagate across modules
590 of a metabolic network, with negative epistasis at lower levels possibly resulting in amplified negative
591 epistasis at a higher level, and positive epistasis at a lower level possibly turning into negative epistasis at
592 a higher level. These findings reinforce the expectation of wide-spread negative epistasis (including
593 DMIs) for fitness. Mapping and functionally understanding complex interactions and their consequences
594 for genome evolution will require new interdisciplinary approaches that combine the tools and expertise
595 from genomics, systems and molecular biology and population genetics in the future.

596 **Materials and methods**

597 ***A toy model illustrates DMI dynamics.***

598 We first consider a toy model of a single DMI pair, similar to Bank et al. 2012, that is purged from an
599 isolated hybrid population. Consider two divergent loci resulting in parental haplotypes *aB* and *Ab*
600 (Figure S1). Hybridization and recombination generate two recombinant haplotypes, *ab* and *AB*. We
601 assume a DMI between *A* and *B*, and that direct selection acts on the incompatible alleles *A* and *B*. We
602 denote the fitness of the haplotypes as $w_{ab} = 0, w_{Ab} = \alpha, w_{aB} = \beta, w_{AB} = \gamma$. Throughout the manuscript,
603 we usually set these coefficients to $\alpha = 0.001, \beta = 0.002, \gamma = -0.5$, representing weak direct selection
604 and a strong DMI. To gain an intuition of the purging process and the statistics, we first compare the

frequency dynamics with and without DMIs by assuming haploid hybrid populations with recombination in a deterministic model.

Dynamics of the variance of heterozygosity for a two-locus DMI

In the model described above, we denote g_{ik} as the frequency of haplotype ik, where $i \in \{A, a\}$ and $k \in \{B, b\}$. The allele frequency of allele i at locus **A** is p_i and the allele frequency of allele j at locus **B** is p_k. The expected heterozygosity h at a locus is $h_{\mathbf{A}} = 1 - {p_a}^2 - {p_A}^2$, $h_{\mathbf{B}} = 1 - {p_b}^2 - {p_B}^2$. The expected heterozygosity at two loci is the sum of the heterozygosity at **A** and **B**,

$$h = h_{\mathbf{A}} + h_{\mathbf{B}}.$$

We define the two-locus variance of heterozygosity as in Sved 1968, summing over the contributions of numbers of heterozygous loci (see Table S10), as

$$\begin{aligned}\sigma_{het} &= (0-h)^2\left({g_{Ab}}^2 + {g_{aB}}^2 + {g_{ab}}^2 + {g_{AB}}^2\right) + (1-h)^2(2g_{ab}g_{Ab} + 2g_{ab}g_{aB} + 2g_{AB}g_{Ab} + 2g_{AB}g_{aB}) \\ &\quad + (2-h)^2(2g_{aB}g_{Ab} + 2g_{ab}g_{AB}) \\ &= h_{\mathbf{A}} + h_{\mathbf{B}} - {h_{\mathbf{A}}}^2 - {h_{\mathbf{B}}}^2 \\ &\quad + 2\sum_i\sum_k({g_{ik}}^2 - {p_i}^2{p_k}^2) \\ &= h_{\mathbf{A}} + h_{\mathbf{B}} - {h_{\mathbf{A}}}^2 - {h_{\mathbf{B}}}^2 + 4\sum_i\sum_k p_ip_kD_{ik} + 2\sum_i\sum_k {D_{ik}}^2,\end{aligned}$$

where $D_{ik} = g_{ik} - p_ip_k$ is the LD for each haplotype.

If there is no LD, the alleles that segregate at different loci in an individual are independent, so the expected value of the variance of heterozygosity is $\sigma_2^{exp} = h_{\mathbf{A}} + h_{\mathbf{B}} - {h_{\mathbf{A}}}^2 - {h_{\mathbf{B}}}^2$.

The deviation of the variance of heterozygosity from the independent expectation is therefore

$$\Delta_2 = 4\sum_i\sum_k p_ip_kD_{ik} + 2\sum_i\sum_k {D_{ik}}^2.$$

Its relative deviation from the expectation is

$$X(2) = \frac{\Delta_2}{\sigma_2^{exp}},$$

which is the focal statistic of this paper. Similar to LD, $X(2)$ indicates the association between alleles at two loci. The association is created by recombination, genetic drift, and negative epistasis between alleles.

The LD of recombinants is $D = g_{ab} \cdot g_{AB} - g_{aB} \cdot g_{Ab}$.

Normalizing D to D' we obtain

$$D' = \frac{D}{D_{max}},$$

where

$D_{max} = min(-p_A \cdot p_B, -p_a \cdot p_b)$ when $D < 0$;

$D_{max} = max(p_A \cdot p_b, p_a \cdot p_B)$ when $D > 0$.

We use D' as classical measure of LD throughout the manuscript.

ΔD2 quantifies imbalanced segregation of recombinant haplotypes

We also rewrote the deviation of the variance of heterozygosity from its expectation in the form of haplotype frequencies (g_{ik}) and allele frequencies (p_i, p_k)

$$\Delta_2 = 2\sum_i \sum_k ({g_{ik}}^2 - {p_i}^2{p_k}^2)$$

We propose $D2_{ik} = {g_{ik}}^2 - {p_i}^2{p_k}^2$ as a measure of the imbalance of haplotype frequencies, which is a key indicator of a DMI.

In a deterministic model without selection, we expect the recombinant haplotypes to be generated at the same frequency after hybridization. It follows that $D2_{ab}$ and $D2_{AB}$ of recombinant haplotypes are expected to be equal initially, independent of the initial admixture proportions. When there is a DMI, the recombinant haplotype *AB* is depleted, and this expectation does not hold. We use the *D2* difference between the two recombinants to refer to DMI-indicating recombinant imbalance and from now on refer to this as $\Delta D2 \coloneqq D2_{AB} - D2_{ab}$ (Figure S3 A-D).

$D2_{ik}$ can also be written as linkage disequilibrium D times a factor, $D2_{ik} = (g_{ik} - p_i p_k)\,(g_{ik} + p_i p_k) = D \cdot (g_{ik} + p_i p_k)$. Therefore, we expect $D2_{ik}$ to be 0 when the population approaches linkage equilibrium (D approaches 0; Figure S4 A-D). The total deviation of $D2_{ik}$ from 0, $\sum_i \sum_k |D2_{ik}|$, is an indicator of linkage disequilibrium. The proportion of $\Delta D2$ over the sum of all ΔD_{ik}, $(D2_{AB} - D2_{ab})/\sum_i \sum_k |D2_{ik}|$, can also be written as $|1 - p_A - p_B|$, showing the relationship between the recombinant imbalance and the elimination of the two alleles that interact negatively in the DMI.

Simulation of a DMI under Wright-Fisher and Moran models

In our simulations, implemented in SLiM (version 3.5, Haller and Messer 2019), we assumed a population size of 5000 diploid individuals, with selection occurring at the juvenile stage, followed by random mating and recombination. For simplicity and computational speed, we ignored mutation (similar to Blanckaert & Bank 2018). At the beginning of the simulation, two fully diverged parental populations of adults are combined at admixture proportions f and $1 - f$, which mate randomly. When considering immigration, we assume that migrants from one parental population enter the hybrid population at the juvenile stage after selection. Direct fitness is multiplicative and F1 individuals do not suffer from the DMI, similar to the "recessive" DMI scenario in Turelli and Orr 2000 and Blanckaert & Bank 2018, where the DMI affects homozygous-homozygous genotypes ($w_{AB/AB} = (1+\alpha)^2(1+\beta)^2(1+\gamma)^4$) and homozygous-heterozygous genotypes ($w_{AB/aB} = (1+\alpha)(1+\beta)^2(1+\gamma)^2$ and $w_{AB/Ab} = (1+\alpha)^2(1+\beta)(1+\gamma)^2$).

666 To represent the demography of the studied swordtail fish populations that have overlapping generations,
667 we compared the DMI dynamics and the difference in the time period in which DMIs can be detected
668 between a non-overlapping generation model (Wright-Fisher) and an overlapping generation model
669 (Moran), both of which operate in discrete time. In the Moran model, mortality rates for each age
670 determined the age structure [0.2, 0.2, 0.0, 0.0, 0.0, 0.0, 0.25, 0.5, 0.75, 1.0] from age 0 to 9 in every
671 generation. Age 0 is the new generation of juveniles. Probability of death for each individual was
672 determined according to these mortality rates and individuals that experienced mortality were replaced by
673 juveniles after selection.

674 ***Sensitivity and specificity***

675 Before applying the statistics to real data, we used simulations under a Wright-Fisher model with variable
676 admixture proportions and migration rates to estimate the sensitivity and specificity of our method in
677 detecting DMIs (Figure 2 & S5). We designed these simulations to correspond to realistic conditions for
678 its application to swordtail fish data. According to empirical estimates, the migration rate *m* ranges from
679 <<0.01 to ~0.05 in the natural hybrid populations studied here. To roughly explore these dynamics in our
680 simulations, we included no ($m = 0$), intermediate ($m = 0.005$) and strong ($m = 0.01$) migration from
681 minor parental populations, and two admixture proportions $f \in \{0.3,0.5\}$. We used the proportion of
682 successful DMI detection under our model to estimate sensitivity, or the true positive rate. Specificity was
683 estimated by the proportion of DMIs that we failed to detect under a null model of neutral evolution.
684 Since $X(2)$ was always larger than -0.001 in the 300 simulations we performed on neutrally evolving
685 pairs, DMI detection criteria were set to $X(2) < -0.001$ & $D' < 0$. $D' < 0$ ensures that the depleted
686 haplotype is a recombinant but not a parental haplotype. The population size was 5000 in each
687 simulation, and the recombination rate was set to 0.5 f or each pair of loci. We simulated each scenario
688 100 times and randomly sampled 300 individuals for further analysis. Sensitivity and specificity were
689 calculated from generation 10 onwards.

690 ***False and true positive rates with one and two DMI pairs***

691 To generalize the performance analyses, we simulated genomes with one DMI or two DMIs on two or
692 four chromosomes (chromosomes with DMIs are referred to as DMI chromosomes below). In the same
693 genome, two additional neutral chromosomes were simulated to estimate the false positive rate. One
694 hundred markers were simulated on each chromosome, and the population size was set to 5000. The
695 position of the DMI loci were distributed randomly on the DMI chromosomes such that DMIs could be
696 either inter- or interchromosomal. Epistatic coefficients were drawn from a uniform distribution ($\gamma \in$
697 $[-1, -0.001]$)). Direct selection coefficients were drawn from an exponential distribution (mean=0.001,
698 α, β). Due to the large parameter space, we only considered two demographic scenarios, equal admixture
699 proportions without migration ($f = 0.5, m = 0$) and uneven admixture proportions with high migration
700 rates ($f = 0.3, m = 0.01$). We simulated each scenario 100 times, such that 100 DMIs in the one-DMI
701 scenario and 200 DMIs in the two-DMI scenario could be used to determine the true positive rates. We
702 also studied the distribution of the total number of retained two-locus haplotypes when only homozygous
703 loci were considered for varying sample sizes ($n \in \{100, 200, 300\}$) in the two-DMI model. At least 200
704 individuals had to be sampled to obtain >50 homozygous-homozygous haplotypes (Figure S12). The
705 mean and 95% confidence interval of the false positive rate were estimated by resampling 2000
706 interactions from all pairwise statistics of 100 simulations 1000 times.

Fish data availability, quality control and bias control

Previous work has described our approach for local ancestry inference in *X. birchmanni* x *X. malinche* hybrids in detail (Powell et al. 2020, Schumer et al. 2020). Ancestry informative sites were defined based on a combination of low and high coverage sequenced individuals from two allopatric *X. birchmanni* (N=150) and two allopatric *X. malinche* populations (N=33). The difference in sequencing effort reflects a large difference in genetic diversity and the expected number of segregating sites between the two species (*X. birchmanni* π per basepair is 0.1%, *X. malinche* is 0.03%). Based on observed allele frequencies at these sites, we excluded any ancestry informative site that had less than a 98% frequency difference between the two samples. We used observed frequencies at these ancestry informative sites, in combination with counts for each allele from low-coverage whole genome sequence data to infer local ancestry using a hidden Markov model (HMM) based approach. The expected performance of this approach was evaluated in simulations, on pure parental individuals not used in the definition of ancestry informative sites and on early generation F_1 and F_2 hybrids where small switches in ancestry can confidently be inferred to be errors. A subset of simulations also included genetic drift between the reference panels and parental populations that form the hybrid populations (Schumer et al. 2020). Based on these analyses we estimated error rates to be ≤0.2% errors per ancestry informative site in all tested scenarios (Powell et al. 2020, Schumer et al. 2020). Thus, given high expected accuracy of ancestry inference in *X. birchmanni* x *X. malinche* hybrids, we do not expect that errors will cause the patterns observed in the empirical data from natural populations.

For convenience in downstream analysis, posterior probabilities from the HMM for ancestry state were converted to hard calls. Sites with a posterior probability greater than >0.95 for a particular ancestry state (homozygous *X. malinche*, heterozygous, homozygous *X. birchmanni*) were converted to a hard call for that ancestry state. This resulted in 629112, 629582, and 629101 calls at ancestry informative sites for analysis across populations. Individuals were excluded from this dataset if they were nearly pure parental in their ancestry (e.g., <10% of genomes from either parental population).

To avoid biases, we filtered markers in the three populations independently using the same criteria. In each population, we retained a locus for analysis if its frequency of each homozygous genotype was >0.05, and if the locus was missing (or heterozygous) in fewer than 60% of individuals. Ancestry informative sites were thinned by LD ($r^2 > 0.9$) within a 10kb window, with 1-marker step size by PLINK (v1.9, Chang et al. 2015). The thinning resulted in 18,188, 7,244, and 7,415 markers for CHAF, TLMC, TOTO respectively.

We then sampled 700 markers and calculated the statistics (including $X(2)$, D', $\Delta D2$ and haplotype frequencies) pairwise in this data set. If $X(2) < -0.005$ and $D' < 0$ (Table S1-S3), we considered a pair of markers as a candidate DMI pair. To compare the DMI locations among bootstrap replicates and between hybrid populations, we next split the genome into 1Mb bins (based on the reference genome of *X. maculatus*; Schartl et al. 2013). This window size was inspired by our simulations, which showed that the signal of selection at DMI loci could extend to 1-2Mb (Figure S11). Any two bins that contained pairwise interaction candidates were finally considered a putative DMI. We repeated this procedure 20 times.

The proportion of putative DMIs in all pairwise comparisons was very similar to the expected false positive rate obtained from our simulations, and the proportion of putative DMIs was elevated when markers were assigned to 1Mb bins (Table S3 and S5). To be conservative, given this result, we used several approaches. We used the median proportion p of candidate bins with negative $X(2)$ over 20 bootstrap samples as a baseline false positive rate. We then computed the probability of sampling the same candidate interaction between bins (based on choosing different markers in each round) multiple times; this should be binomially distributed with parameter p. Using a binomial test, we determined k as the least number of times that an interaction candidate should be observed to be retained as a putative DMI, using a p value of 0.05 corrected for a false discovery rate of 0.001 (following the Benjamini-Hochberg procedure). We obtained values of k between 6 and 10 for the three hybrid populations (see Table S5). As a second approach to reduce false positives, we also limited our focus to interactions that were observed in multiple populations. Independent observation of interactions in two or more populations provides strong evidence that these interactions are not expected by chance.

Code availability

All simulation and analysis code that was developed to obtain the results presented in this paper will be deposited in an open archive upon publication of the manuscript and is currently available at https://github.com/banklab/DMI-inference-X2.

Acknowledgments

We are grateful for helpful discussions and comments on the manuscript provided by the IGC-FCUL-unibe journal club and Milan Malinsky. We thank Vitor Sousa for extensive feedback and discussions and Alexandre Blanckaert for simulation code and discussions. We thank Kali Tal for her editorial suggestions. CB and JL are grateful for support by EMBO Installation Grant IG4152 and by ERC Starting Grant 804569 - FIT2GO. CB and MS are grateful for support by the HFSP Young Investor Grant RGY0081/2020. MS is also supported by Sloan, Searle, and Pew fellowships. Calculations were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern.

References

Bank C, Bürger R, Hermisson J. The limits to parapatric speciation: Dobzhansky–Muller incompatibilities in a continent–island model. Genetics. 2012;191(3):845–863.

Bay RA, Arnegard ME, Conte GL, Best J, Bedford NL, McCann SR, et al. Genetic coupling of female mate choice with polygenic ecological divergence facilitates stickleback speciation. Current Biology. 2017;27(21):3344–3349.

Blanckaert A, Bank C, Hermisson J. The limits to parapatric speciation 3: Evolution of strong reproductive isolation in presence of gene flow despite limited ecological differentiation. Philosophical Transactions of the Royal Society B. 2020;375(1806):20190532.

Blanckaert A, Bank C. In search of the Goldilocks zone for hybrid speciation. PLoS genetics. 2018;14(9): e1007613.

Brandvain Y, Kenney AM, Flagel L, Coop G, Sweigart AL. Speciation and introgression between Mimulus nasutus and Mimulus guttatus. PLoS Genet. 2014;10(6): e1004410.

Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang X, Barbash DA. Two Dobzhansky-Muller genes interact to cause hybrid lethality in Drosophila. science. 2006;314(5803):1292–1295.

Brown A, Feldman M, Nevo E. Multilocus structure of natural populations of Hordeum spontaneum. Genetics. 1980;96(2):523–536.

Byers KJ, Darragh K, Fernanda Garza S, Abondano Almeida D, Warren IA, Rastas PM, et al. Clustering of loci controlling species differences in male chemical bouquets of sympatric Heliconius butterflies. Ecology and evolution. 2021;11(1):89–107.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4(1): s13742–015.

Chaturvedi S, Lucas LK, Buerkle CA, Fordyce JA, Forister ML, Nice CC, et al. Recent hybrids recapitulate ancient hybrid outcomes. Nature communications. 2020;11(1):1–15.

Corbett-Detig RB, Zhou J, Clark AG, Hartl DL, Ayroles JF. Genetic incompatibilities are widespread within species. Nature. 2013;504(7478):135–137.

Culumber Z, Fisher H, Tobler M, Mateos M, Barber P, Sorenson M, et al. Replicated hybrid zones of Xiphophorus swordtails along an elevational gradient. Molecular ecology. 2011;20(2):342–356.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nature Reviews Genetics. 2011;12(7):499–510.

Felsenstein J. The evolutionary advantage of recombination. Genetics. 1974;78(2):737–756.

Fishman L, Sweigart AL. When two rights make a wrong: the evolutionary genetics of plant hybrid incompatibilities. Annual review of plant biology. 2018; 69:707–731.

Fragata I, Blanckaert A, Louro MAD, Liberles DA, Bank C. Evolution in the light of fitness landscape theory. Trends in ecology & evolution. 2019;34(1):69–82.

Gavrilets S. Hybrid zones with Dobzhansky-type epistatic selection. Evolution. 1997;51(4):1027–1035.

Haller BC, Messer PW. SLiM 3: forward genetic simulations beyond the Wright–Fisher model. Molecular biology and evolution. 2019;36(3):632–637.

Hurst LD, Pál C, Lercher MJ. The evolutionary dynamics of eukaryotic gene order. Nature Reviews Genetics. 2004;5(4):299–310.

Kimura M. Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. Genetics. 1965;52(5):875.

Kryazhimskiy S. Emergence and propagation of epistasis in metabolic networks. Elife. 2021;10: e60200.

Li C, Wang Z, Zhang J. Toward Genome-Wide Identification of Bateson–Dobzhansky–Muller Incompatibilities in Yeast: A Simulation Study. Genome biology and evolution. 2013;5(7):1261–1272.

Lindtke D, Buerkle CA. The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. Evolution. 2015;69(8):1987–2004.

Lu Y, Sandoval A, Voss S, Lai Z, Kneitz S, Boswell W, et al. Oncogenic allelic interaction in Xiphophorus highlights hybrid incompatibility. Proceedings of the National Academy of Sciences. 2020;117(47):29786–29794.

Maheshwari S, Barbash DA. The genetics of hybrid incompatibilities. Annual review of genetics. 2011; 45:331–355.

823 Martin SH, Davey JW, Salazar C, Jiggins CD. Recombination rate variation shapes barriers to introgression across
824 butterfly genomes. PLoS biology. 2019;17(2): e2006288.

825 Mihola O, Trachtulec Z, Vlcek C, Schimenti JC, Forejt J. A mouse speciation gene encodes a meiotic histone H3
826 methyltransferase. Science. 2009;323(5912):373–375.

827 Noor MA, Grams KL, Bertucci LA, Reiland J. Chromosomal inversions and the reproductive isolation of species.
828 Proceedings of the National Academy of Sciences. 2001;98(21):12084–12088.

829 Orr HA, Presgraves DC. Speciation by postzygotic isolation: forces, genes and molecules. BioEssays.
830 2000;22(12):1085–1094.

831 Orr HA. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics.
832 1995;139(4):1805–1813.

833 Ortiz-Barrientos D, Engelstädter J, Rieseberg LH. Recombination rate evolution and the origin of species. Trends in
834 ecology & evolution. 2016;31(3):226–236.

835 Patton EE, Mitchell DL, Nairn RS. Genetic and environmental melanoma models in fish. Pigment cell & melanoma
836 research. 2010;23(3):314–337.

837 Penãlba JV, Wolf JB. From molecules to populations: appreciating and estimating recombination rate variation.
838 Nature Reviews Genetics. 2020;21(8):476–492.

839 Pool JE. The mosaic ancestry of the Drosophila genetic reference panel and the D. melanogaster reference genome
840 reveals a network of epistatic fitness interactions. Molecular biology and evolution. 2015;32(12):3236–3251.

841 Powell DL, García-Olazábal M, Keegan M, Reilly P, Du K, Díaz-Loyo AP, et al. Natural hybridization reveals
842 incompatible alleles that cause melanoma in swordtail fish. Science. 2020;368(6492):731–736.

843 Presgraves DC. The molecular evolutionary basis of species formation. Nature Reviews Genetics. 2010;11(3):175–
844 180.

845 Satokangas I, Martin S, HelanteräH, Saramäki J, Kulmuni J. Multi-locus interactions and the build-up of
846 reproductive isolation. Philosophical Transactions of the Royal Society B. 2020;375(1806):20190543.

847 Schartl M, Walter RB, Shen Y, Garcia T, Catchen J, Amores A, et al. The genome of the platyfish, Xiphophorus
848 maculatus, provides insights into evolutionary adaptation and several complex traits. Nature genetics.
849 2013;45(5):567–572.

850 Schumer M, Brandvain Y. Determining epistatic selection in admixed populations. Molecular Ecology.
851 2016;25(11):2577–2591.

852 Schumer M, Cui R, Powell DL, Dresner R, Rosenthal GG, Andolfatto P. High-resolution mapping reveals hundreds
853 of genetic incompatibilities in hybridizing fish species. Elife. 2014;3:e02535.

854 Schumer M, Powell DL, Corbett-Detig R. Versatile simulations of admixture and accurate local ancestry inference
855 with mixnmatch and ancestryinfer. Molecular ecology resources. 2020;20(4):1141–1151.

856 Schumer M, Powell DL, Delclós PJ, Squire M, Cui R, Andolfatto P, et al. Assortative mating and persistent
857 reproductive isolation in hybrids. Proceedings of the National Academy of Sciences. 2017;114(41):10936–10941.

858 Schumer M, Xu C, Powell DL, Durvasula A, Skov L, Holland C, et al. Natural selection interacts with
859 recombination to shape the evolution of hybrid genomes. Science. 2018;360(6389):656–660.

860 Seehausen O. Hybridization and adaptive radiation. Trends in ecology & evolution. 2004;19(4):198–207.

861 Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal FX, et al. Quantitative trait loci mapping in five new
862 large recombinant inbred line populations of Arabidopsis thaliana genotyped with consensus single-nucleotide
863 polymorphism markers. Genetics. 2008;178(4):2253–2264.

864 Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nature
865 Reviews Genetics. 2008;9(6):477–485.

866 Smith JM, Smith NH, O'Rourke M, Spratt BG. How clonal are bacteria? Proceedings of the National Academy of
867 Sciences. 1993;90(10):4384–4388.

868 Smukowski C, Noor M. Recombination rate variation in closely related species. Heredity. 2011;107(6):496–508.

869 Sved J. The stability of linked systems of loci with a small population size. Genetics. 1968;59(4):543.

870 Szabo Q, Bantignies F, Cavalli G. Principles of genome folding into topologically associating domains. Science
871 advances. 2019;5(4): eaaw1668.

872 Tang S, Presgraves DC. Lineage-specific evolution of the complex Nup160 hybrid incompatibility between
873 Drosophila melanogaster and its sister species. Genetics. 2015;200(4):1245–1254.

874 Turner LM, Harr B. Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and
875 Dobzhansky-Muller interactions. Elife. 2014;3: e02504.

876 Turner LM, White MA, Tautz D, Payseur BA. Genomic networks of hybrid sterility. PLoS Genet. 2014;10(2):
877 e1004162.

878 Wu CI, Ting CT. Genes and speciation. Nature Reviews Genetics. 2004;5(2):114–122.