

Inversions Can Accumulate Balanced Sexual Antagonism: Evidence from Simulations and *Drosophila* Experiments

Reviewed Preprint

v2 • August 27, 2024

Revised by authors

Reviewed Preprint

v1 • December 28, 2023

Christopher S McAllester , John E Pool 

Laboratory of Genetics, University of Wisconsin – Madison, USA

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

Abstract

Chromosomal inversion polymorphisms can be common, but the causes of their persistence are often unclear. We propose a model for the maintenance of inversion polymorphism, which requires that some variants contribute antagonistically to two phenotypes, one of which has negative frequency-dependent fitness. These conditions yield a form of frequency-dependent disruptive selection, favoring two predominant haplotypes segregating alleles that favor opposing antagonistic phenotypes. An inversion associated with one haplotype can reduce the fitness load incurred by generating recombinant offspring, reinforcing its linkage to the haplotype and enabling both haplotypes to accumulate more antagonistic variants than expected otherwise. We develop and apply a forward simulator to examine these dynamics under a tradeoff between survival and male display. These simulations indeed generate inversion-associated haplotypes with opposing sex-specific fitness effects. Antagonism strengthens with time, and can ultimately yield karyotypes at surprisingly predictable frequencies, with striking genotype frequency differences between sexes and between developmental stages. To test whether this model may contribute to well-studied yet enigmatic inversion polymorphisms in *Drosophila melanogaster*, we track inversion frequencies in laboratory crosses to test whether they influence male reproductive success or survival. We find that two of the four tested inversions show significant evidence for the tradeoff examined, with *In(3R)K* favoring survival and *In(3L)Ok* favoring male reproduction. In line with the apparent sex-specific fitness effects implied for both of those inversions, *In(3L)Ok* was also found to be less costly to the viability and/or longevity of males than females, whereas *In(3R)K* was more beneficial to female survival. Based on this work, we expect that balancing selection on antagonistically pleiotropic traits may provide a significant and underappreciated contribution to the maintenance of natural inversion polymorphism.

eLife assessment

This study attempts to solve long-standing puzzles about inversion polymorphisms in *Drosophila melanogaster* by invoking sexually antagonism and negative frequency dependent selection. While the idea developed here is a **valuable** contribution to the field, the description of the empirical work and the simulations remain **incomplete**, as they do not provide a full picture of what exactly has been done.

<https://doi.org/10.7554/eLife.93338.2.sa4>

1 Introduction

1.1 The Enigmatic Prevalence of Inversion Polymorphism

Inversions, DNA segments in reversed sequence order relative to the ancestor, have been a focus of population genetic study for their direct or indirect effects on recombination in natural populations for more than a century. In their most dramatic effect on meiosis, inversions generate aneuploid gametes when a crossover occurs between the interior of an inverted region and its non-inverted homolog (White 1973 [↗](#)). As aneuploidy is often fatal in embryos or severely deleterious, this may translate to a fecundity cost for a heterokaryotypic parent: a parent heterozygous for an inversion arrangement. However, depending on a species' reproductive biology, such reproductive fitness effects may be mitigated if the zygotes terminate early or before significant parental resource investment (Charlesworth 1994, Munasinghe and Brandvain 2023 [↗](#)). Alternatively, crossovers may be inhibited directly in heterokaryotypes around the inversion by mechanisms similar to crossover interference (Koury 2023 [↗](#)), avoided by achiasmy in one sex or all, or aneuploid gametes may be segregated to polar bodies in female meiosis (Sturtevant 1921 [↗](#); Sturtevant and Beadle 1936 [↗](#); Krimbas and Powell 1992 [↗](#); Coyne et al. 1993 [↗](#); Gong et al. 2005 [↗](#)). Regardless of the differences between these scenarios, effectively fewer recombinant offspring are produced in heterokaryotypes.

Inversions may also have direct effects on fitness when disrupting a gene, a regulatory element, or other genomic structure like a topologically associated domain (Lupiáñez et al. 2015 [↗](#); McBroome et al. 2020 [↗](#)). Given also that a potential effect of inversions in heterokaryotypic individuals is the deleterious generation of aneuploid gametes, it seems reasonable to expect inversion rearrangements to confer either deleterious or neutral fitness consequences in the absence of association with other variants. However, when a new inversion mutates, it samples a single haplotype from a population, which is likely to have some non-neutral fitness effect, even if mild. Greater natural variation in fitness within a population increases the likelihood of sampling haplotypes with exceptionally deleterious or beneficial associations of alleles, even when there are no other effects on fitness (Berdan et al. 2023 [↗](#)).

In contrast to the above predictions, a considerable number of well-studied inversion variants are maintained at an intermediate frequency within a population, presumably due to the indirect effect inversions have on reducing recombination and maintaining fit haplotypes that are themselves under selection (reviewed in Berdan et al. 2023 [↗](#)). The evolutionary forces maintaining both haplotypes are thought to be quite varied between populations and arrangements, but a common first hypothesis for a frequently sampled inversion is that local adaptation in one region or environment is balanced with migration from another (Kirkpatrick and Barton 2006 [↗](#); Charlesworth and Barton 2018 [↗](#)). This association may be maintained more strongly when the alleles have beneficial epistatic interactions, and assumed beneficial epistasis between loci is often paired with another mechanism of balancing selection operating at a locus to

explain inversion (Dobzhansky 1949 [↗](#), 1950 [↗](#), 1970 [↗](#); Schaeffer et al. 2003 [↗](#); Hoffmann and Rieseberg 2008 [↗](#)). However, unless migration rates are very high, local adaptation models predict inversions to be at high frequency in their favored geographic range but at low frequency elsewhere, and many inversions do not have sampled locations at which they reach very high frequency.

Alternately, chance associations of both an inversion and its standard arrangement counterpart with different recessive deleterious alleles may mimic the fitness profile of a single overdominant locus (sometimes termed associative overdominance), where heterozygotes carrying both main haplotypes experience none of the deleterious effects of a homozygous recessive across the region (Sturtevant and Mather 1938 [↗](#); Frydenberg 1963 [↗](#); Zhao and Charlesworth 2016 [↗](#); Gilbert et al. 2020 [↗](#)). However, recent theory and simulation suggests that this dynamic is uncommon unless an arrangement polymorphism is maintained at intermediate frequency under some other selective force, and the population scaled recombination rate between arrangements is low (Charlesworth 2024 [↗](#)).

Balanced polymorphic loci may be maintained in other ways, under various models and selective pressures. For example, multiple overdominant loci that are each most beneficial as heterozygotes may benefit from linkage maintaining complementary haplotypes (Sved 1968 [↗](#); Charlesworth and Charlesworth 1973 [↗](#); Charlesworth 1974 [↗](#)). Several recent studies have proposed that functionally overdominant loci may be more common than previously considered. Specifically, the dominance relation between alleles at a locus may change between different sexes, seasons, life history stages, or other trait or fitness challenge contexts, generally termed ‘reversal of dominance’, which in some cases can lead heterozygotes to have greatest average fitness (Rose 1982 [↗](#); Wittman et al. 2017; Grieshop and Arnqvist 2018 [↗](#); Connallon and Chenoweth 2019 [↗](#)). Other specific biological dynamics may predict the maintenance of inversion polymorphism in particular ways. For example, inversions may also link drivers and enhancers in meiotic drive systems (Babcock and Anderson 1996 [↗](#); Mroczek et al. 2006 [↗](#); Courret et al. 2019 [↗](#)), and these drive-associated inversions may persist at equilibrium frequencies in the presence of repressors (e.g. Bastide et al. 2022 [↗](#)).

In natural populations, it can be difficult to distinguish among these alternate hypotheses. They are not always mutually exclusive, and the selective pressures acting upon an inversion may vary across space and time. The frequency patterns of many inversions, including a number in *Drosophila melanogaster*, remain poorly explained.

1.2 Inversions in *Drosophila melanogaster*

D. melanogaster is a widely studied organism in many fields of biology, including evolutionary and population genetics. Multiple polymorphic inversions segregate at meaningful frequencies in natural populations of *D. melanogaster* (Lemeunier and Aulard 1992 [↗](#); Aulard et al. 2002 [↗](#); Kapun et al. 2016 [↗](#); Kapun and Flatt 2019 [↗](#)), but even in this model system, it is not entirely clear which population genetic processes are most responsible for shaping their frequencies. In this species, achiasmy in males (Morgan 1910 [↗](#); John et al. 2016 [↗](#)) and a lack of evidence for female fecundity costs for most inversions (Sturtevant 1921 [↗](#); Sturtevant and Beadle 1936 [↗](#); Coyne et al. 1993 [↗](#); Gong et al. 2005 [↗](#)) suggest that most paracentric inversions in *D. melanogaster* do not generate aneuploid gametes at any appreciable frequency. Further, genome editing tools have allowed for the creation of synthetic inversions, revealing that the tested structural rearrangements are not themselves responsible for the gene regulatory changes that are associated with natural inversions in *D. melanogaster* (Said et al. 2018 [↗](#)). Instead, given that breakpoints of common *D. melanogaster* inversions occur disproportionately in regions that confer greater recombination suppression (Corbett-Detig 2016 [↗](#)), *D. melanogaster* inversions would appear to be maintained primarily for the sake of maintaining linkage associations between distant variants.

D. melanogaster originated in southern-central Africa, in seasonally dry Miombo and Mopane woodlands, and (potentially after adapting to a human commensal niche) expanded out across much of Africa and into the Middle East around 13,000 years ago, and into Europe around 1,800 years ago (Sprengelmeyer et al. 2020 [DOI](#)). With an ancestral N_e of around 2 million (Sprengelmeyer et al. 2020 [DOI](#)), *D. melanogaster* populations have high levels of standing diversity and high population recombination rates. Many *D. melanogaster* inversions have geographic clines repeated between several colonized regions (Mettler et al. 1977 [DOI](#); Reinhardt et al. 2014 [DOI](#); Kapun et al. 2016 [DOI](#)), and inversions often show geographic differences in frequency that exceed genome-wide average differentiation between the same populations (e.g. Pool et al. 2017 [DOI](#)). It has therefore been suggested that local ecological adaptation may shape the distribution of *D. melanogaster* inversions (Kapun et al. 2016 [DOI](#); Kapun and Flatt 2019 [DOI](#)). Indeed, it is difficult to explain their geographic differentiation without some type of spatially-varying selective pressures.

Yet, it is unclear whether ecological adaptation is a sufficient explanation for *D. melanogaster* inversion frequency patterns, for two reasons. First, most of the common inversions become less frequent as one samples populations farther from the sub-Saharan ancestral range (or similar warm environments) into potentially more challenging, colder high latitude and altitude environments (Aulard et al. 2002 [DOI](#); Kapun et al. 2016 [DOI](#); Lack et al. 2016a; Pool et al. 2017 [DOI](#); Kapun and Flatt 2019 [DOI](#)) with *In(3R)Mo* representing an exception to this pattern (Kapun et al. 2014 [DOI](#); Kapun et al. 2016 [DOI](#)). As these inversions are all derived (being absent from closely related species), local adaptation theory would predict that they arose to protect local adaptation against ongoing migration, more likely in the novel environment if it is experiencing high levels of immigration (Kirkpatrick and Barton 2006 [DOI](#)). It is unclear why inversions would be needed to adapt to the species' ancestral environments, or why ancestral "standard" arrangements would have an ecological advantage in derived environments that are very different from the ancestral range.

Second, even in the locations where inversions would hypothetically be ecologically advantageous, they rarely surpass intermediate frequencies (Lemeunier and Aulard 1992 [DOI](#); Pool et al. 2017 [DOI](#); Kapun and Flatt 2019 [DOI](#); Sprengelmeyer et al. 2020 [DOI](#)). Under a local adaptation model, that observation would require a secondary explanation that resists fixation, such as very high levels of long distance gene flow or the fixation of recessive deleterious variants on inverted haplotypes (Ohta and Kimura 1970 [DOI](#); Zhao and Charlesworth 2016 [DOI](#); Gilbert et al. 2020 [DOI](#)). While migration may well limit inversion frequencies in some populations, we note that some putatively adaptive Single Nucleotide Polymorphism (SNP) variants do achieve much greater frequency differentials between populations (da Silva Ribeiro et al. 2022 [DOI](#)), and thus we would have to suppose that inversions are under relatively weaker selection to explain their lesser differentiation. Regarding the hypothesis of high recessive load, we note that *D. melanogaster* populations do generally contain substantial recessive genetic load (Greenberg and Crow 1960), and it has been speculated that associative overdominance between sampled backgrounds might boost inversion frequencies in *D. melanogaster* populations during founder events (Pool et al. 2012) or in isolated wilderness environments (Sprengelmeyer et al. 2020 [DOI](#)). However, it seems likely that in large human-commensal populations, *D. melanogaster* inversions would escape from linked deleterious variants through gene conversion or double crossover events with standard chromosomes, in light of both the age of most of these inversions (Corbett-Detig and Hartl 2012 [DOI](#)) and the very high population recombination rates of *D. melanogaster*. Homozygous lines for each common inversion also appear at appreciable frequency in inbred lab strains, demonstrating a lack of perfect linkage of the inversions with recessive lethal or infertile variants (e.g. Lack et al. 2016a).

While local adaptation, gene flow, and associative overdominance may all contribute to geographic variation in inversion frequency among *D. melanogaster* populations, it is not clear that these processes are collectively sufficient to explain observed patterns. Alternatively, the allele frequencies and diversity patterns of at least some of these inversions might be primarily shaped by some form of balancing selection acting on inversion-linked variation, in which the

balanced frequency is dependent on the environment (Kapun et al. 2023). Balancing selection on inversions has received increasing attention in recent literature (Wellenreuther and Bernatchez 2018; Faria et al. 2019). In *Drosophila*, inversions have been associated with balancing selection due to seasonal, temporally fluctuating selection (Machado et al. 2021). However, the potential for temporally varying selection to maintain stable balanced polymorphisms on its own is somewhat limited (Gillespie 1988). Further, most common inversions maintain relatively high polymorphic frequencies across broad geographic ranges without obvious shared seasonal selective pressures, from tropical and subtropical dry forests to temperate locations with harsh winter conditions (Lemeunier and Aulard 1992). Therefore, we predict that the frequencies of some of these inversions could be in part explained by a mechanism of balancing selection that would operate even within a temporally and spatially homogenous population. Below, we propose such a model in which alleles that contribute to competitive mating display simultaneously contribute antagonistically to survival to reproductive maturity, a tradeoff that offers a mechanism by which the life-history characteristics of a population could generate epistatic balancing selection.

1.3 Sexually Antagonistic Polymorphism in *D. melanogaster*

Among mechanisms of selection, sexual selection is both prevalent, and provides a simple mechanism for frequency dependence (O'Donald et al. 1997). Alleles that contribute to competitive mate choice systems may often follow a negative frequency dependence within the population, as display success depends on the relative quality of a male among local competitors, and the more common a display-enhancing allele is, the fewer competitors it succeeds over. Combinations of display-enhancing alleles may similarly gain a disproportionate benefit together, as male mating success in competitive display systems can be skewed towards a few successful individuals (Bateman 1948; Jones et al. 2002; Tatarenkov et al. 2008, though see Gowaty et al. 2012; Hoquet et al. 2020) such that the last few alleles to push display into the top quantile contribute a significantly greater fitness increase to the genotype than the first few display alleles. This interaction of alleles is a form of epistasis - a non-independent fitness interaction between loci - that is emergent and non-specific to the genetic identity of the variants. In other words, the alleles contribute independently to the generation of a trait, and the fitness of the individual is a nonlinear function of that trait (Blows and Brooks 2003; Rest et al. 2013; Otwinowski et al. 2018). This epistatic interaction suggests a potential for indirect selection on the recombination-suppressing effects of inversions, in order to maintain combinations of sexually antagonistic variants that have the greatest fitness when carried together. This tendency of selection to favor the reduction of recombination when the haplotypes present are at a fitness optimum is well studied for recombination modifiers including inversions, and has been termed the 'reduction principle' (Feldman 1972; Feldman and Balkau 1973; Feldman et al. 1980; Altenberg and Feldman 1987; Altenberg et al. 2017; see Altenberg and Feldman 1987 for a direct treatment of inversions).

While the emergent epistasis generated by a mate choice system may favor linkage of variants, alleles that contribute only to a sexually selected display are not expected to remain as balanced polymorphisms, rather they are expected to fix, albeit perhaps slowly as their fitness benefits diminish at high frequencies. However, if these alleles simultaneously contribute deleteriously to other components of fitness, they are pleiotropic (Stearns 2010), and given the sex-limited benefits of the display they would likely also be sexually antagonistic (Rowe et al. 2018), whereby traits increase the fitness of one sex but decrease fitness if found in the other. Pleiotropy can result from traits that are inherently physiologically coupled, from resource allocation restraints, or from the specific role or behavior of genes. Pleiotropy can further be sexually antagonistic when the traits in the tradeoff are differentially important to different sexes, which may be common given the differing fitness requirements of the sexes in a population (Cox and Calsbeek 2009; Connallon and Chenoweth 2019). It should be noted that a model of two pleiotropically determined traits each with potentially sex-specific fitness can be treated separately from a model where one trait has sexually antagonistic fitness effects.

D. melanogaster mating involves a male display performance and female acceptance or rejection. A number of experimental evolution studies (Rice 1992 [↗](#); Rice 1996 [↗](#); Audet et al. 2024 [↗](#)) and genetic studies (Innocenti and Morrow 2010 [↗](#); Cheng and Kirkpatrick 2016 [↗](#); Glaser-Schmitt et al. 2023 [↗](#)) have confirmed the prevalence of sexually antagonistic loci in *D. melanogaster*. Pleiotropic alleles that trade off between survival and male display quality represent one clear potential case of antagonism to consider for a model of balancing selection, and one highly relevant to the biology of *D. melanogaster*.

Given the relevance of sexual selection and sexual antagonism to *Drosophila* evolution, we propose that polymorphic *D. melanogaster* inversion and standard arrangements may help maintain haplotypes of alleles that specify opposite ends of a sexually antagonistic tradeoff. We first explore this model through forward simulation, to establish that it may in theory arise in a population and maintain inversion polymorphism over a significant number of generations. Then we examine the results of an experiment with an outbred laboratory population of *D. melanogaster*, to test whether common inversions show changes in frequency over the course of a generation in a manner consistent with a trade-off between survival and male reproductive success. Many such antagonistically pleiotropic scenarios are possible under different selective tradeoffs, and indeed some have been proposed for other model inversions with varying mechanisms of maintenance (Betrán et al. 1998 [↗](#); Pearse et al. 2019 [↗](#); Mérot et al. 2020 [↗](#); Pei et al. 2023 [↗](#); see Discussion), but we focus on a specific scenario of sexual antagonism here.

2 Results

2.1 A Model of Sexually Antagonistic Inversions

Here, we consider a population featuring a competitive mate choice system that results in a skewed distribution of reproductive success among males. Within this population, we model a genomic region containing a number of linked loci, each with alternate antagonistic, pleiotropic alleles that trade off between male display success and survival likelihood (of females or both sexes). Critically, we do not directly model any epistatic interactions among loci affecting survival or display traits, nor do we invoke dominance in our diploid model, so any such fitness dynamics emerge from the modeled population processes.

If mate choice is highly competitive, such that each female can choose amongst several or many males, the variance of offspring number among males (**Figure 1A** [↗](#)) may be much higher than for females - yielding a distribution of male reproductive output that follows a relatively convex trend, in which most males produce few offspring. Conversely, superior display quality can enable males to mate successfully much more often than other display phenotypes they compete against (**Figure 1A** [↗](#)). An increase in the display of a male with low quality results in a smaller change in number of offspring than an increase in display of a male with high quality (change ii in **Figure 1A** [↗](#)), due to the reproductive skew towards highly competitive individuals. So, there is an emergent non-additive fitness effect, where multiple variants increasing the reproductive rank of an individual together have a larger fitness benefit than the combined effects of each substitution separately - a positive, synergistic epistatic effect. This effect occurs whether the difference is substituting alleles at two loci, or substituting two homologous alleles at a locus in a diploid or polyploid, which would give an emergent dominance effect. Additionally, fitness conveyed by an allele favoring display quality is also frequency-dependent: since mating success depends on the display qualities of other males, the relative advantage of a display trait will be diminished as more males carry it, for both haploid and diploid populations (**Figure 1B** [↗](#), male fitnesses). We refer to this as a negative frequency dependent selection upon alleles favoring display, though the alternate alleles have a positive frequency dependence. We note that some authors instead use “negative frequency dependence” to specifically describe fitness functions in which any and all alternate alleles are of selective benefit when rare (e.g. Charlesworth and Charlesworth 2010 [↗](#)).

And so, for an antagonistic variant that benefits display and begins under net positive selection, the benefit will decrease as the variant rises in frequency. If a variant only conveys benefit to mate competition success, it is expected to fix given enough time, as even if its evolution is nearly neutral at high population frequency, selection will push the frequency back up if it drifts down (Hazel 1943 [↗](#); Dickerson 1955 [↗](#); Charlesworth and Hughes 2000 [↗](#)). However, if the variant is antagonistically pleiotropic for traits not involved in competition between males, such as survival in the face of climate, pathogens, or predators, these fitness costs may typically not be frequency dependent. If the benefit of an antagonistic display-favoring allele overcomes the pleiotropic cost when the variant is rare, its diminished benefit when at higher frequency may reverse its net benefit, while at some intermediate frequency, we predict that these fitness effects should be equal and an equilibrium frequency should be reached and maintained (**Figure 1B** [↗](#)). In the absence of other forces, this dynamic is expected to maintain the allele under balancing selection indefinitely.

We further predict that multiple such balanced pleiotropic loci should generate interesting emergent behavior. Because of the emergent synergistic epistasis from mate competition (see discussion of **Figure 1A** [↗](#) above), the haplotype with the greatest male display benefit should increase in frequency relative to other haplotypes until reaching an equilibrium at which its relative advantage in mate choice is reduced enough to be balanced by the pleiotropic, frequency-independent, non-epistatic survival cost (**Figure 1B** [↗](#)). The haplotypes and their linked alleles conveying intermediate phenotypes will have less than proportional benefits during mate choice due to the convex relationship between quality and success described above (**Figure 1A** [↗](#)), but a larger, proportional cost to survival against non-intraspecific survival challenges, giving them lower average fitness in the equilibrium state than either of the extreme haplotypes (**Figure 1C** [↗](#)). Therefore, recombination between the two extreme haplotypes will generate less fit haplotypes in the gametes, and a local modifier of recombination that reduces recombination and its associated costs (such as an inversion) will be favored by indirect selection (Altenberg and Feldman 1987 [↗](#); Otto and Lenormand 2002 [↗](#)). And so, selection will favor the association of these haplotypes with the alternate standard and inverted karyotypes, which generate fewer recombinant gametes from a heterokaryotypic parent. We further predict that in a population that is close to the balanced equilibrium state for such a pleiotropic trade-off, the population of adults will contain a higher frequency of survival-focused haplotypes than the population of zygotes, but the display-focused haplotypes will increase in frequency in zygotes relative to adults, in line with the life stage-specific fitness differences between these haplotypes.

While **Figure 1** [↗](#) illustrates a haploid scenario for the sake of simplicity, this model extends naturally to a diploid case. The main difference is that heterokaryotypic diploids – those heterozygous for the arrangement – would have intermediate trait values and be less fit than either homokaryotype, for the same selective reasons that an intermediate haplotype with two alleles of opposing effects at different loci is least fit in a haploid context. This is therefore, interestingly, an example of a scenario in which the two most common haplotypes in the population can exhibit underdominant fitness, and yet may still be maintained at a stable balanced equilibrium. By comparison, the simplest models of underdominant fitness predict an unstable equilibrium that disfavors polymorphism (Gillespie 2004, Charlesworth and Charlesworth 2010 [↗](#)), a characteristic shared with non-frequency-dependent models of disruptive selection. The cost of generating underdominant offspring might be compensated for by a strong co-directional dominance of the variants involved, such that heterokaryotypic individuals exhibit a phenotype similar to one of the homokaryotypes. However, dominance is not a necessary feature of our model, and it is not invoked in our simulations. Whereas, investigating a wider range of genetic architectures would be a potential direction for future study.

2.2 Simulation Results

We investigated the predictions of the above model of inversion-associated balanced sexual antagonism using a purpose-built forward simulation program, “Sexually Antagonistic Inversion simulator” (SAIsim; see Materials and Methods). This simulation program incorporates the

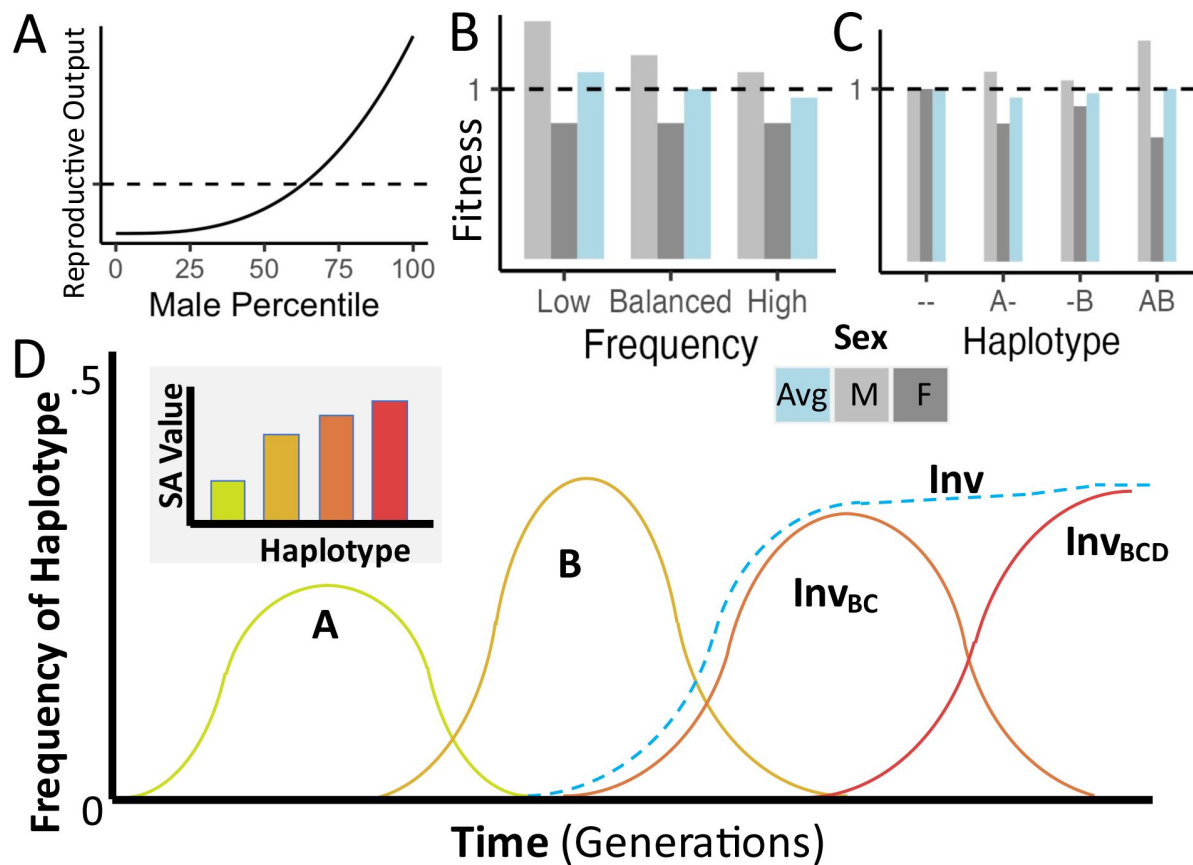


Figure 1

Conceptual representations of the proposed model of inversion-associated balanced sexual antagonism. (A) Hypothetical skewed distribution of male mating success, yielding a greater variance in reproductive success for males than for females in our model. (B) Fitnesses of display-favoring alleles under sexually antagonistic balancing selection, illustrating a pleiotropic variant that should rise in frequency when rare, but decline if frequency exceeds the balanced equilibrium value. (C) Fitnesses of display-favoring alleles at two haploid loci under the same model, illustrating synergistic epistasis between displaying-favoring alleles. (D) A hypothetical trajectory of four such mutations, in which B outcompetes A, an inversion links B and C to create a more strongly display-favoring haplotype, and then the addition of D to that haplotype further the accumulation of antagonistic variants, reaching an equilibrium frequency while displacing less extreme display-favoring haplotypes.

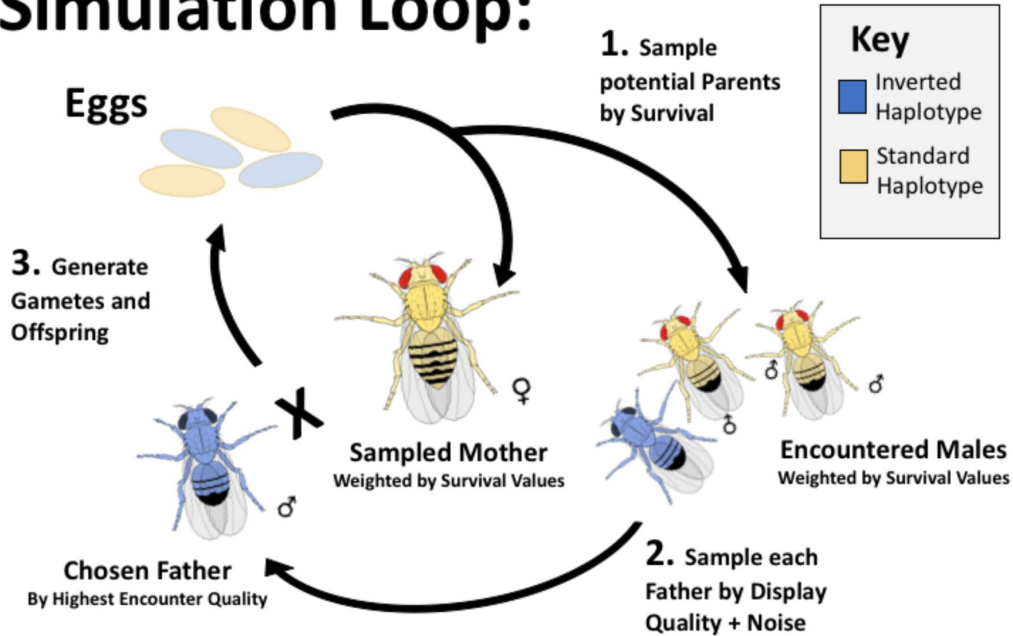
generation of new nucleotide and inversion polymorphisms by mutation, where each nucleotide mutation independently draws a positive or negative effect on survival and on male display. It also simulates crossing over and gene conversion, and inherently the effects of drift (Materials and Methods). SAIsim includes gene conversion between all homologous sites regardless of karyotype, and models crossing over in heterokaryotypic individuals from even-numbers of crossover events within the inverted region. It explicitly models mate selection, in that for each female randomly selected to reproduce, a set number of ‘encountered’ males are randomly selected, and the observed display value of each male for this competitive event is assigned based on the sum of the male’s genetic effects and an environmental noise effect resampled at every encounter. The male with the highest observed display value then succeeds in this mating (**Figure 2** [↗](#)).

With this simulator, we first confirmed that some individual variants with an antagonistic pleiotropy between survival (initially of both sexes) and male reproduction can generate stable selectively balanced polymorphisms *in silico*. We simulated a single locus in diploid populations of $N = 1000$, polymorphic for an allele of a specified survival cost and display benefit at an initial frequency of 0.5, with the alternate allele having no trait effect. One hundred males were sampled for each mating competition. After $20N$ generations, we calculated the proportion of simulations retaining polymorphism and the average final frequency of the variant across all simulations, revealing that there is a range of paired trait effects for which the mutation remained at intermediate frequency after $20N$ generations (**Figure 3** [↗](#)). Alleles with small effect sizes had only a narrow band of balanced selection, outside of which they were either fixed or lost, while large-effect alleles remained balanced across a larger area of simulated parameter space. While this balanced range is somewhat narrow, such variants may persist under balancing selection indefinitely in the absence of other forces, and many small-effect alleles in linkage blocks may confer larger effects in synergy. In a scenario where only females pay the survival cost, the balanced parameter space of antagonistic trait effects is broader and shifted somewhat (**Figure S1**). In a scenario where only males pay the survival cost, and variants are neutral in females, essentially no simulations retain polymorphism (**Figure S1**). The number of males sampled to assess mate competition for each offspring also has a significant effect, with larger values giving broader polymorphic ranges, though simulations with costs only in females show robust polymorphism even when only sampling two males for each competition (**Figure S1**).

We then tested the behavior of two linked loci, both polymorphic for an antagonistic and non-antagonistic allele. We identified two antagonistic variants previously found to undergo balancing selection separately, and performed simulations with varying genetic distances between them (**Figure 4** [↗](#)). Crossing over and gene conversion were simulated, with gene conversion occurring for each heterozygous variant at a probability of 10^{-2} , with random direction. per female meiosis per site, intended to be a conservative approximation of the number of gene conversion events affecting a given site in *D. melanogaster*, scaled to simulations of $N=1000$ (Comeron et al. 2012 [↗](#)). The stronger antagonistic variant (modelled at fixed position 0.225M) persisted in all simulations, in line with expectations, and so we focus on whether the weaker variant (at varying positions) also persists with it. As seen in the upper panel of **Figure 4** [↗](#), the weaker variant (pale red line) was retained at close linkage, as the haplotype containing both antagonistic variants acts as a single larger effect supergene. However, as the weaker variant was moved farther from the stronger variant, recombination increasingly generated unfit intermediate genotypes compared to the extremes of the reproduction-survival tradeoff, and the stronger variant outcompeted the weaker variant among these haplotypes, leading to the loss of the weaker variant.

We then examined the same spacings between antagonistic variants across a region containing a polymorphic inversion, with its left breakpoint close to the stronger variant, and the right breakpoint within the range of the positions in which the weaker variant was simulated (**Figure 4** [↗](#), lower panel). When the inversion was added, the weaker variant was then retained when the two antagonistic variants are relatively more distant, due to association of the weaker variant with the right-hand inversion breakpoint (pale red line), at which point this association also

Simulation Loop:



12

Figure 2

The layout of a single simulated generation. Simulated individuals each reach the reproductive stage with probability in proportion to the product of the survival effects of their alleles. To generate each offspring in the next generation, a surviving female is first randomly sampled (with replacement) to be the mother. Then a specified number of surviving males are randomly sampled to generate the pool of males encountered by the sampled female - without replacement for each pool and with replacement between pools. For each encounter pool, each male is assigned an observed display quality as the sum of the display effects of its alleles, plus a normally distributed noise effect for each encounter. The highest quality male in the encounter pool is selected as the father, and the offspring genome is generated. The female gamete is generated with crossover and gene conversion events. Odd numbers of crossovers in the interior of an inversion are resampled from the same parent, to model the biology of an organism that eliminates aneuploid gametes during oogenesis, or exhibits reproductive compensation. *D. melanogaster* males do not cross-over, and so males do not recombine in the simulations presented.

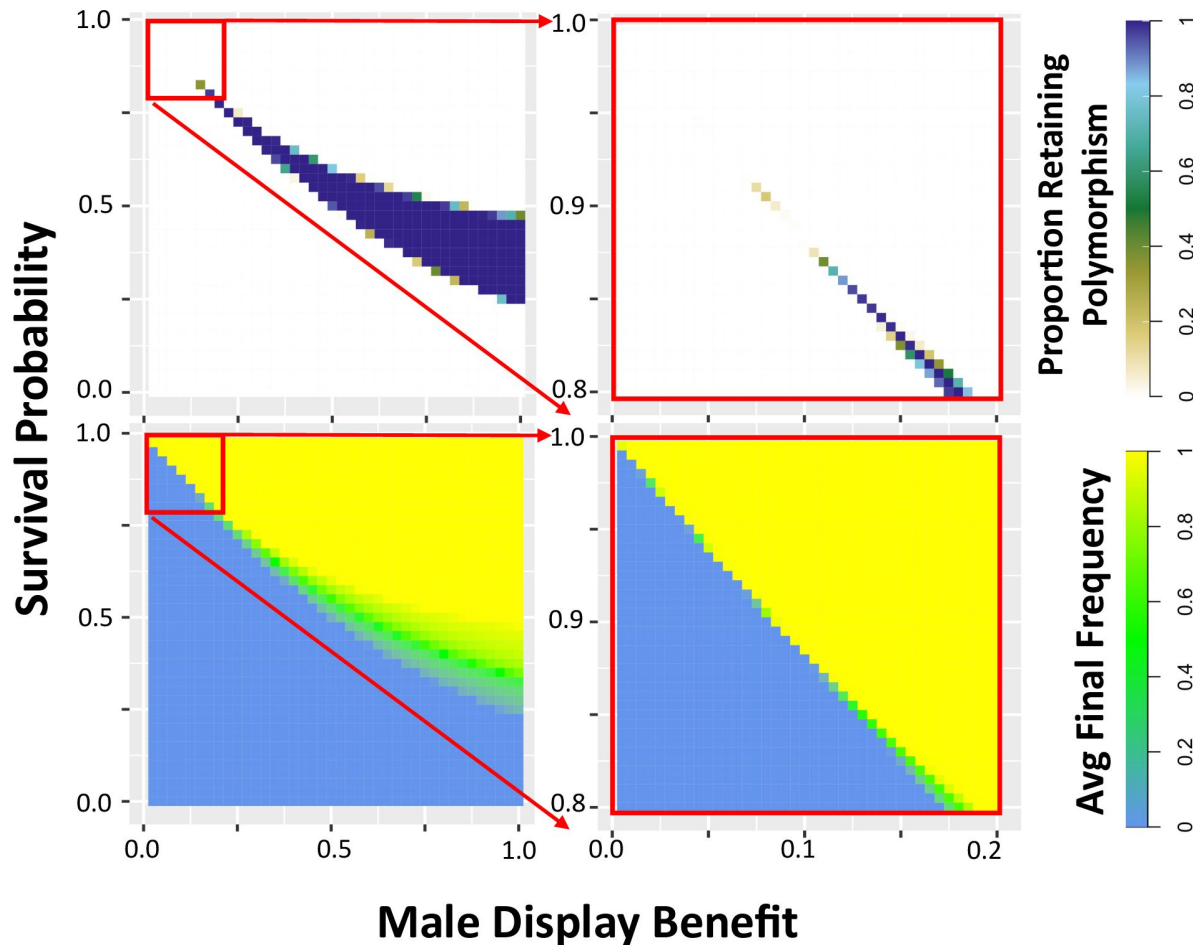


Figure 3

Simulations show that a balanced equilibrium frequency exists for certain sexually antagonistic variants. The proportion of simulations retaining polymorphism (top) and the average final frequency (bottom) are plotted for a single locus at which one antagonistic allele with the indicated effects reducing survival proportion (of both sexes) and increasing male reproductive quality score is initiated at equal frequency with an alternative allele that has no such effects. Larger effect sizes are plotted in the left panels, with values incrementing by 0.025, and smaller effect sizes (a subset of the larger) are plotted in greater detail in the right panels, with values incrementing by 0.005. Most parameters result in either all simulations retaining polymorphism (purple), or else all simulations losing polymorphism (white). However, simulations retaining polymorphism have a broad range of equilibrium frequencies (shades of green). For each parameter combination, 500 replicates were simulated for a diploid population of 1,000 individuals evolving for 20,000 generations, with further details as indicated in the Materials and Methods.

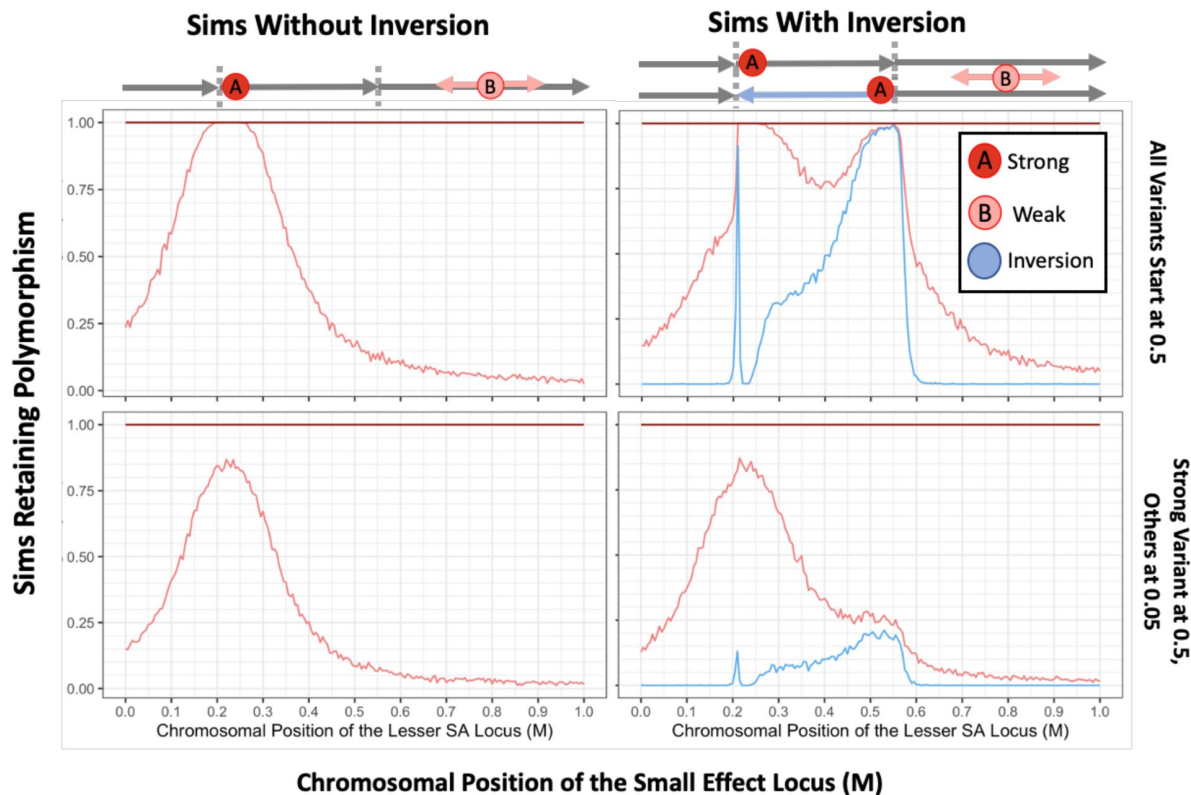


Figure 4

Simulated inversions may persist as polymorphisms when linked to sexually antagonistic, pleiotropic variants. In turn, the presence of inversions facilitates the accumulation of antagonistic variation. Rates of long-term persistence are shown for simulations with two defined antagonistic variants alone (left panels) and with the additional presence of a defined inversion (right panels). To focus the potential addition of a second linked variant to an antagonistic haplotype, we begin with the stronger variant already at an intermediate frequency in the population (0.5). The weaker antagonistic variant and the inversion either both start at 0.5 frequency as well (upper row) or both start at a lower frequency of 0.05 (lower row). The 1 Morgan chromosomal segment is diagrammed above the plots, with the inversion breakpoints marked (0.21M, 0.55M), chosen to include space across and beyond the inverted region. The proportion of simulations in which the second, variably-positioned weaker antagonistic variant (survival and reproductive values 0.86, 0.12) maintains polymorphism is shown (light red line), dependent on the recombination distance from a stronger allele (0.75, 0.3) at fixed position 0.225M, which always persists (dark red line at 1 across the top of each plot). In the right panels, the proportion of replicates in which the inversion is retained is also plotted (blue line). Values for each parameter combination were averaged over 1000 replicate simulations, for populations of 1000 diploid individuals run for $20N$ generations. Gene conversion occurs for each heterozygous variant at a probability of 10^{-2} per female meiosis with random direction. Crossover arrangements generating aneuploidy, where there are an odd number of crossovers within the inverted region of a heterokaryotypic parent, are resampled from the same parent to represent removal to polar bodies or reproductive compensation.

preserves the inversion itself in a polymorphic state (light blue line). While more stochastic, a weaker antagonistic variant near the middle of the inversion also appeared to maintain polymorphism more frequently in systems starting with an inversion than without, i.e. across approximate positions 0.30-0.45M in [Figure 4](#), in spite of its weaker linkage with the inversion itself due to double cross-over events. The inversion is not maintained when the weaker antagonistic variant is within the inversion but tightly linked to the stronger variant (see positions 0.22-0.25M); here there is little crossing-over for an inversion to suppress, and so the effect of the inversion approaches neutrality. Interestingly, there is a small window of weaker variant position in which the inversion maintains polymorphism despite the already tight linkage between SA variants, where the inversion seemingly does not modify recombination much (see positions 0.2-0.22M). Here, the weaker antagonistic variant is more tightly linked to the breakpoint than to the stronger variant in this range, and presumably benefits more consistently from the crossover reduction. The maintenance of the weaker variant drops more rapidly outside of the inversion than over similar distances in the simulation without inversions (positions 0-0.2M, bottom versus top panel). This seemingly counter-intuitive result may stem from crossover events that result in aneuploid gametes being resampled in our simulations, which effectively increases the crossover rate outside the inversion breakpoints, a behavior which should be biologically reasonable in taxa with aneuploid chromosome elimination or reproductive compensation. So, for a weaker variant outside the left inversion breakpoint, the presence of the inversion can increase the crossover rate between the two antagonistic variants, and thus reduce the probability of the weaker variant surviving by retaining linkage to the stronger one.

In further simulations, we began with a population initially free of variation, in which sexually antagonistic mutations arise stochastically. To assess differences between longer or shorter antagonistic regions, or higher or lower crossover rates, we modeled populations with parameters scaled to either 1 Mb or 10 kb, using estimates from [Comeron et al. \(2012\)](#) and [Huang et al. \(2016\)](#). The mutations had pleiotropic effects increasing reproductive quality score and decreasing survival proportion, each drawn sampled independently (Materials and Methods).

Depending on the effect scores drawn, a mutation could be broadly advantageous (potentially leading to a selective sweep), broadly deleterious, or potentially subject to balancing selection ([Figure 3](#)). These mutations arose at a rate equal to one thousandth of the scaled single nucleotide mutation rate given in [Huang et al. \(2016\)](#). Inversions (of random lengths and positions) also arose stochastically, at a hundredth of the rate of antagonistic variants. The frequency trajectories of antagonistic variants and inversions are tracked, and the simulation was run for 20N generations to allow near-equilibrium patterns to be established. These simulations demonstrated the potential of a population without initial balanced polymorphism or genetic variation to establish alternate karyotypes which link sexually antagonistic, pleiotropic alleles and an inversion, to maintain polymorphism in a balanced manner, and then to subsequently accumulate significant antagonistic character ([Figure 5A](#)), although the degree of antagonism accumulated depends strongly on the rate of crossing over (relative to the occurrence of antagonistic mutations; [Figure 5B](#)).

The simulations that successfully maintained balanced antagonistic chromosomes also had an interesting distribution of allele and genotype frequencies of the balanced haplotypes at equilibrium. Without any selection, the inversion frequency follows a pattern characteristic of the expected neutral site frequency spectrum ([Figure S2](#)). But strikingly, in parameter spaces where simulations with shared survival costs between sexes maintained polymorphism under balancing selection, inversion frequency distributions were tightly clustered around 0.25 and 0.75 among the zygote population ([Figure S2, 6A](#)). We found that this outcome was related to the genotype frequencies of reproducing adults. These simulations evolve toward strong antagonism, and it is predominantly males homozygous for the display-favoring haplotype (whether that is inverted or standard) who succeed in reproducing, and these males therefore contribute one copy of the display-favoring haplotype to each of their male and female offspring. However, this haplotype's

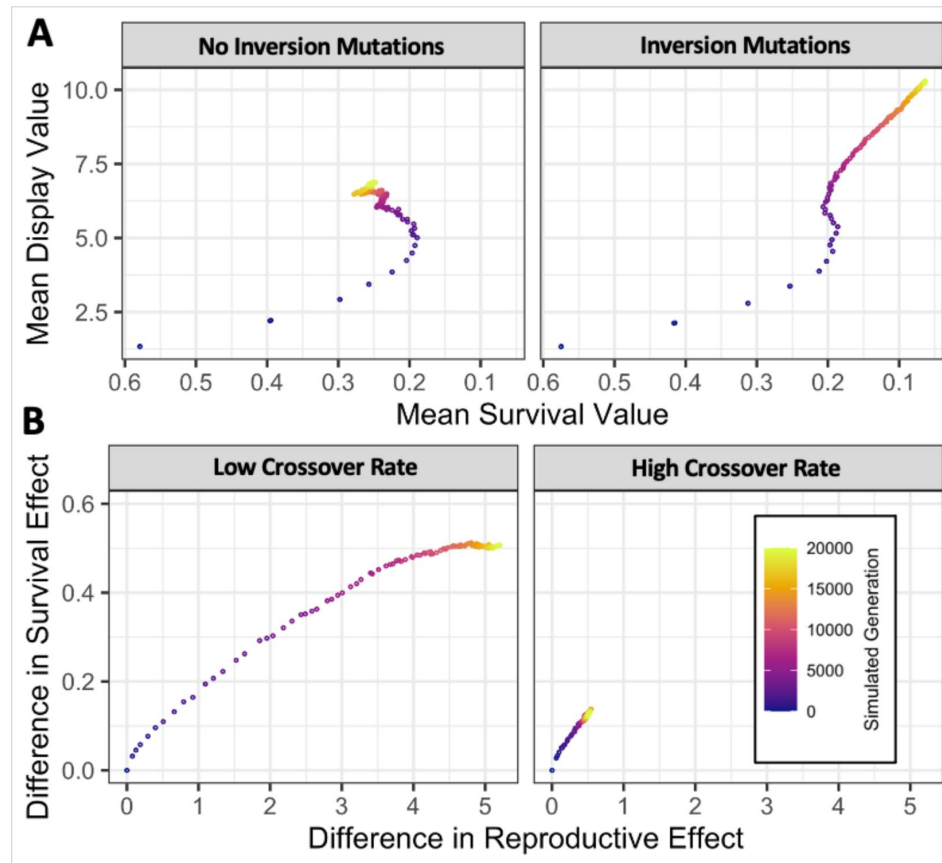


Figure 5

Simulated populations with stochastic mutation of new inversions and sexually antagonistic variants generate persistent inversion polymorphisms and accumulate linked sexually antagonistic variants. (A) In simulated populations with stochastic mutation of new variants having pleiotropic effects on survival and male display, levels of antagonism increased with time before stabilizing at an intermediate value (left panel). Whereas, simulations which also allowed the mutation of new inversions continued accumulating stronger antagonism throughout the simulated time interval (right panel). (B) In simulations which allowed the mutation of new inversions, a low crossover rate (left panel, 43.6 cM in females after scaling) allows populations to accumulate larger differences between the most common karyotype and all others than when the rate of crossover is particularly high (right panel, 43.6 M in females after scaling). Gene conversion occurs at each heterozygous variant with a scaled probability of 0.1295 per heterozygous variant in both scenarios. Values given are the mean of 1000 replicate simulations of 1000 diploid individuals run for 20N generations.

significant cost to viability and longevity would select for the heterokaryotypic females over the display-favoring homokaryotypes. Given that successful parents would predominantly be heterokaryotypic females and display-favoring homokaryotypic males, around three quarters of transmitted chromosomes have the display-favoring arrangement and one quarter does not, yielding the 0.25 and 0.75 zygote inversion frequencies seen in simulations. Although these are autosomal simulations, selection yields an inheritance pattern is similar to that of a ZW sex chromosome system, where the Z chromosome is homokaryotypic in the fathers, and heterokaryotypic with the W chromosome in the mothers.

The sex-specific fitness dynamics present at equilibrium under this model also result in population-wide genotype frequencies that depart from standard Hardy-Weinberg expectations. If the preponderance of matings are between heterokaryotypic females and homokaryotypic display-favoring males, then homokaryotypes for the survival-focused haplotype will rarely be produced (as confirmed in **Figure 6B**). We note that in cases where the standard arrangement favors male display, then especially if antagonism is strong, this model could explain an observed deficiency or absence of inversion homokaryotypes, even if the inverted arrangement was not associated with any recessive deleterious load. Further, heterokaryotypes may become disproportionately frequent after viability selection (**Figure 6B**). And still, this enrichment for heterokaryotypes reflects neither the existence of recessive deleterious load associated with each arrangement, nor overdominance with respect to viability. Hence, our model may offer another alternative explanation for deviations from Hardy-Weinberg genotype frequencies observed for polymorphic inversions in natural populations.

We further ran simulations in which the survival costs were not shared, but instead specific to females, representing a distinct, stronger form of sexual antagonism. In these simulations, we surprisingly recovered three different equilibrium arrangement states (**Figure 7**). 20.0% of simulations exhibited a state similar to the shared costs simulations presented in **Figure 6**, where we infer that fathers were homokaryotypic for a display-favoring arrangement, and mothers were heterokaryotypic and carried survival-focused arrangement that is observed in surviving but not reproducing males. 37.0% of simulations also developed two predominant arrangements, but with transmission patterns similar to an XY sex chromosome system rather than a ZW sex chromosome system. In these cases, we infer that the mothers were homokaryotypic for an X-like survival-focused arrangement, but the fathers were heterokaryotypic and carried a Y-like display-focused arrangement that was apparently lethal in females. 40.6% of simulations instead featured three intermediate frequency haplotypes: with both Y-like and W-like arrangements restricted to reproducing fathers and mothers respectively, and a third shared arrangement with intermediate phenotypic effects.

Specific properties of our simulations may have influenced the occurrence of these three stable states under the female-limited cost model. To reduce the computational complexity of both crossovers and arrangement-specific trait values, the simulations do not allow two inversions on the same individual chromosome, whether overlapping or not, though two chromosomes sampled from the population may have distinct inversions with overlapping breakpoints. This means that if two arrangements, each with an inversion, become the prevalent haplotypes and thereby remove all standard arrangement chromosomes from the population, there is no mutational path to generate a third arrangement. However, a portion of the two-state simulations still had the ancestral, non-inverted arrangement as one of the two haplotypes present, which could have been mutated to a third arrangement to establish the three-arrangement pattern. We are not sure what selective or mutational barriers may prevent transitioning between each state once established. One clear barrier is that it is difficult to generate a less extreme arrangement from a Y-like arrangement, as these can accumulate an indefinite number of mutations since males experience no costs.

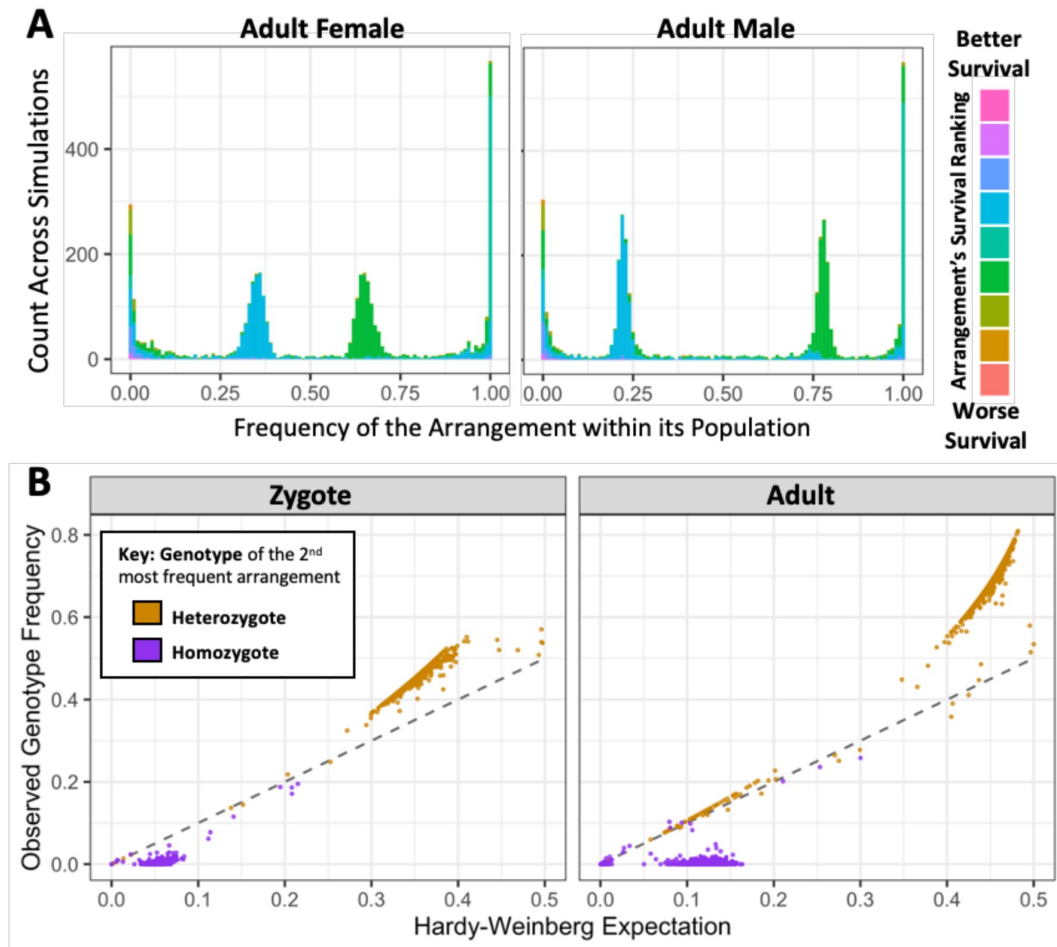


Figure 6

Antagonism-associated karyotypes reach predictable equilibrium frequencies in simulations with antagonistic and inversion mutations, when survival costs are shared between sexes. (A) Histograms of the number of arrangements at a specified frequency across all simulation replicates in females (left) and males (right), colored by how the average survival effect of that arrangement ranks compared to the others in its population. Ranks are normalized to be relative to the median arrangement value, to better account for rare arrangements and ties. These simulations reach an equilibrium with two predominant arrangements, with the more survival-focused arrangement around an overall frequency of 0.25 and more common in females, and the more male-competition-focused arrangement around a frequency of 0.75 and more common in males. (B) Because successful males tend to be homozygotes and successful females heterozygotes, the less frequent of the two major arrangements (which tends to favor survival) has homokaryotype frequencies far below the Hardy-Weinberg expectation, and there is an excess of heterokaryotypes. These genotype frequencies shift somewhat between zygote (left panel) and adult (right panel) populations as the genotype affects survival.

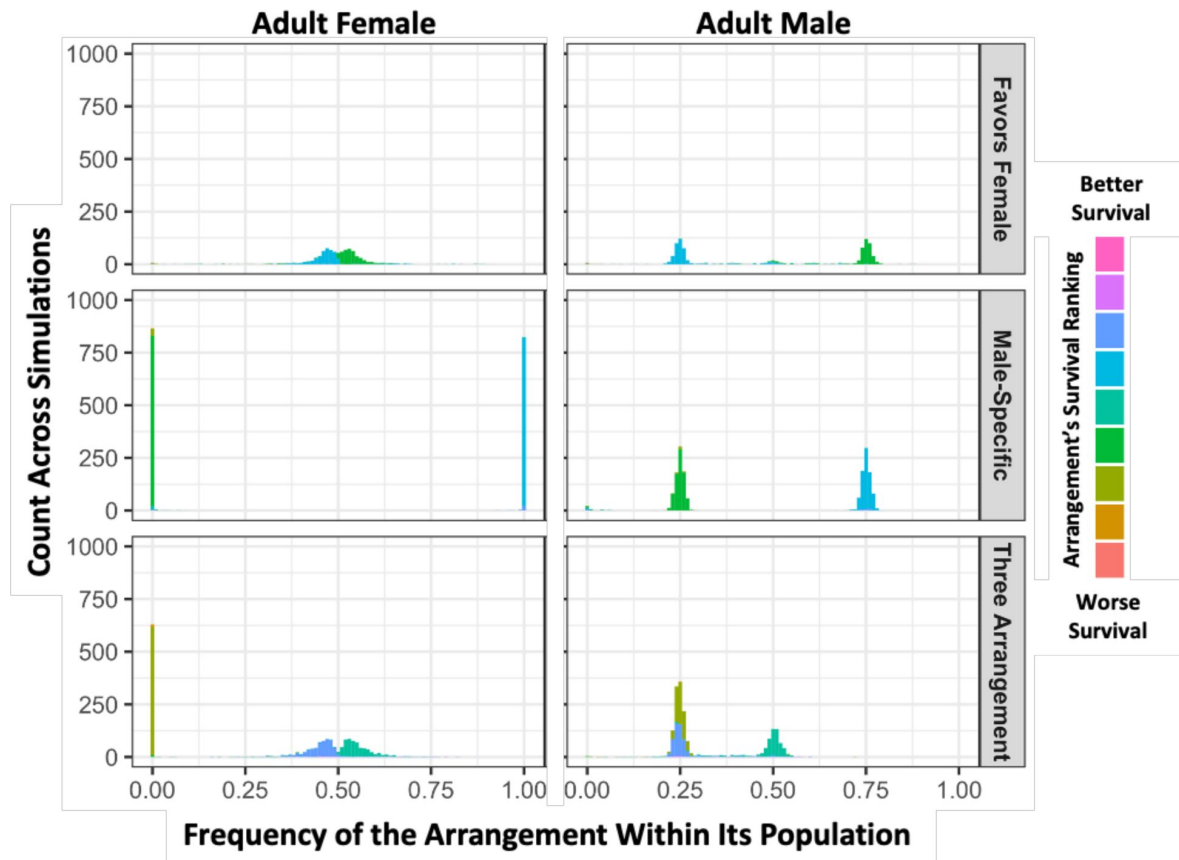


Figure 7

In simulations with female-limited survival costs, three distinct outcomes are observed involving either two or three balanced haplotypes. From simulations with randomly occurring antagonistic and inversion mutations, histograms of inversion frequencies across simulations for surviving females (left) and males (right) are partitioned into three categories based on sex-specific haplotype frequency outcomes. The “Favors Female” category (top) includes simulations in which there are exactly two arrangements with frequencies between 0.1 and 0.9 in both sexes (20.0% of replicates). The “Male-Specific” category (middle) includes simulations in which there are two arrangements with whole-population frequencies between 0.1 and 0.9, but one arrangement is absent in adult females (37.0% of replicates). The “Three Arrangement” category (bottom) includes simulations in which there are exactly three arrangements with whole-population frequencies between 0.1 and 0.9 frequency (40.6% of replicates). 2.3% of simulations did not fit in these categories and had only one predominant arrangement. As in [Figure 6](#), inversions are colored by how the average survival effect of that arrangement ranks compared to the others in its population, and ranks are normalized to be relative to the median arrangement value, to better account for rare arrangements and ties.

2.3 Empirical tests for inversion-associated tradeoffs

A specific motivation for developing and evaluating the model of inversion-associated, balanced sexual antagonism was to potentially better explain the pattern of inversion frequencies observed in *D. melanogaster*, in which multiple inversions reach some of their highest frequencies in populations from the ancestral range or with similarly tropical/subtropical climates. We therefore set out to test experimentally whether any of four inversions common in a Zambia population of this species show evidence for tradeoffs between male reproductive success and survival (in terms of viability and/or longevity).

We assembled an outbred population with moderate frequencies of each inversion to obtain karyotypically diverse males. We set up four crosses using these males along with females from one of four different inversion-free inbred Zambia lines. This crossing scheme is represented in [Figure 8](#). These parental crosses occurred in a single large population cage, with populations of at least 600 of each sex. We developed a PCR and next-generation sequencing-based assay to estimate inversion-associated SNP frequencies in each pool of males and in each corresponding pool of embryos. These SNPs (given in Table S1) were based on our own analysis of variants that showed perfect association with inversions among 197 Zambia genomes (Table S2; see Materials and Methods). We then implemented a resampling method to test for significant inversion frequency changes between parents and offspring (while accounting for the variance observed in technical replicates; see Materials and Methods), in order to determine whether males carrying each inversion were more or less successful in reproducing than the average male. We also compared frequencies between embryos and aged adults that eclosed either in the first or second half of the collection period, to investigate the relationship between inversions and viability/longevity. We calculated p-values from the resampling tests to assess the significance of each individual change in inversion frequency between relevant samples (Table S3). However, in accord with our primary question regarding fitness tradeoffs, we primarily focus here on joint tests of two opposing frequency changes, as further described below. We could therefore test the hypothesis that a given inversion had opposing influences on survival and male reproduction, which would provide evidence consistent with our model of balanced sexual antagonism, while also examining whether any such tradeoffs depend on maternal genetic background.

When we initially examined the joint effects of male reproduction and survival, we found that in a majority of cases (in 13 out of 16 inversion-line trials; binomial $P = 0.02$; [Figure 9](#)), a given inversion increased in frequency in adult offspring compared to the null expectation based on the parental male pool. The remaining three cases were all for *In(3L)Ok*. Ten of the 13 increases were primarily due to increase in frequency between embryo and adult samples, suggesting that in terms of viability and/or longevity, inversion heterokaryotypes often had higher fitness than standard homokaryotypes, as no offspring homokaryotypic for an inversion were generated by our crossing design. We note that our standard homokaryotypes should not have broad-scale homozygosity within the inverted region, but are instead heterozygous for unique standard-arrangement haplotypes from two independent Zambia strains, and that long genomic tracts of identity-by-descent between Zambia strains are quite rare in this large, ancestral-range population (Lack et al. 2016a). Hence, we do not expect that inversion heterokaryotypy is buffering against deleterious effects of inbreeding. Instead, this fitness advantage suggests either an advantage for heterokaryotypy in particular, or perhaps more likely, that multiple inversions harbor alleles that benefit survival under our experimental conditions (see Discussion).

When we separated the effects of male reproduction (based on father to embryo changes) and survival (embryos versus adults), we found that the effects of inversions on survival were generally consistent across maternal genotypes. Three inversions were associated with increased survival: *In(3R)K* was in all four crosses, while *In(2L)t* and *In(2R)NS* each showed elevated survival with all maternal strains except ZI418N ([Figure 9A](#)). In contrast, *In(3L)Ok* was associated with lower survival in each cross ([Figure 9A](#)). With regard to male reproduction, the trajectory of

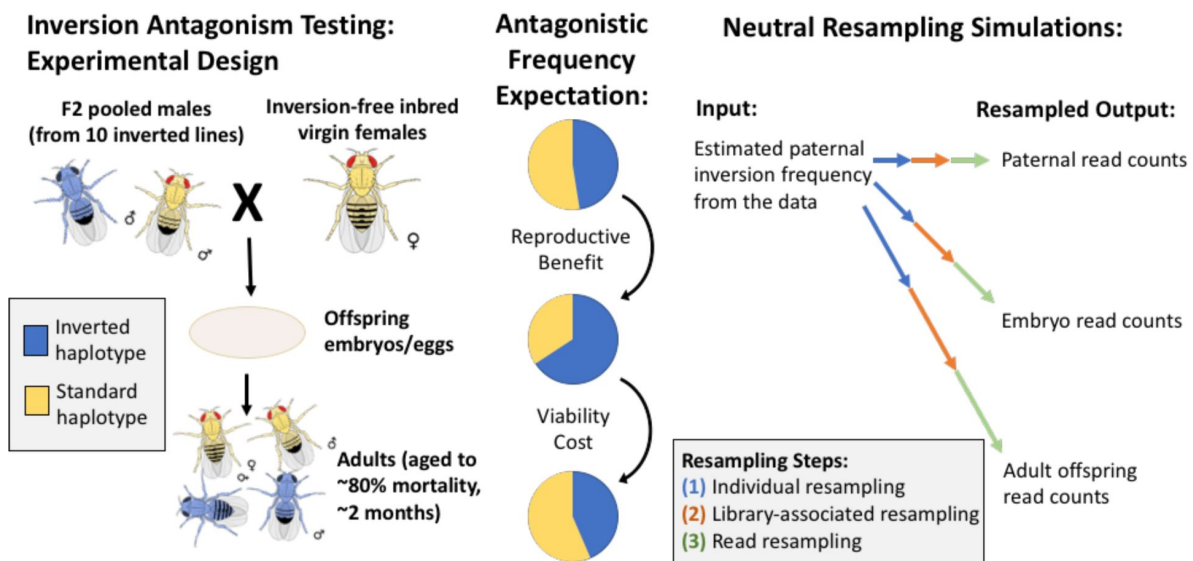
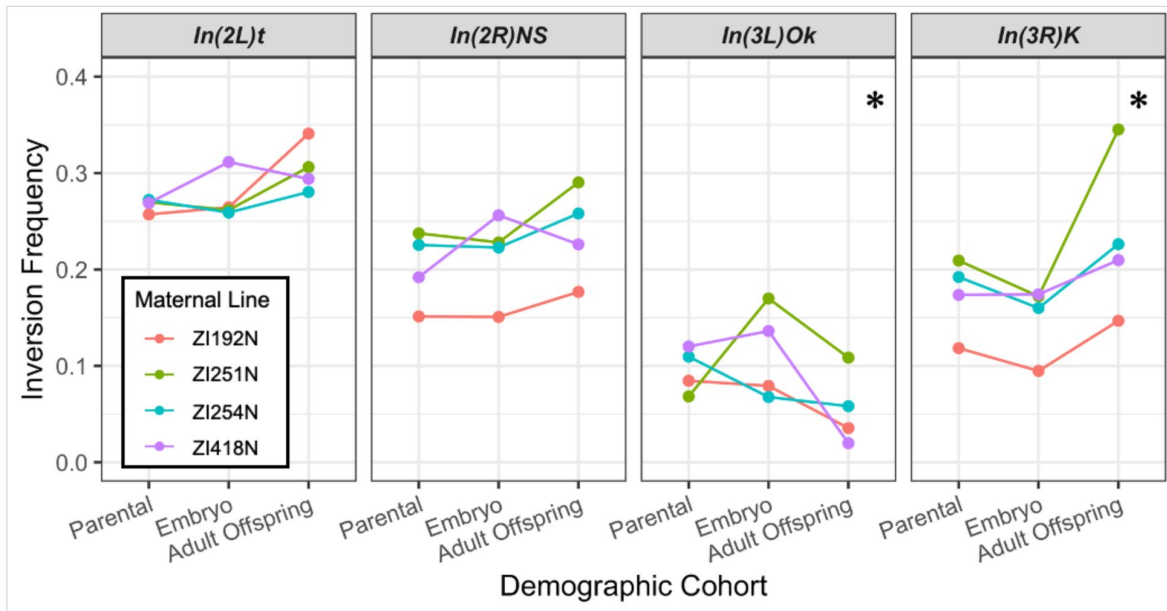


Figure 8

The layout and potential expectations of the laboratory evolution experiment. We cross outbred males from a high-inversion population to inbred, non-inverted females from a specific inbred line, and collect DNA samples from the fathers, embryo offspring, and aged adult offspring (left). We hypothesize from our antagonistic pleiotropy model that one or more of the four inversions present on separate chromosome arms in our experimental populations may experience opposing selection between fathers and embryos versus between embryos and aged adults (middle). To test the significance of observed frequency changes, we designed a resampling model to represent the sources of neutral variation expected in the experiment (right).

A Testing for tradeoffs between male reproduction and survival



B Inversions convey some sex-specific survival effects

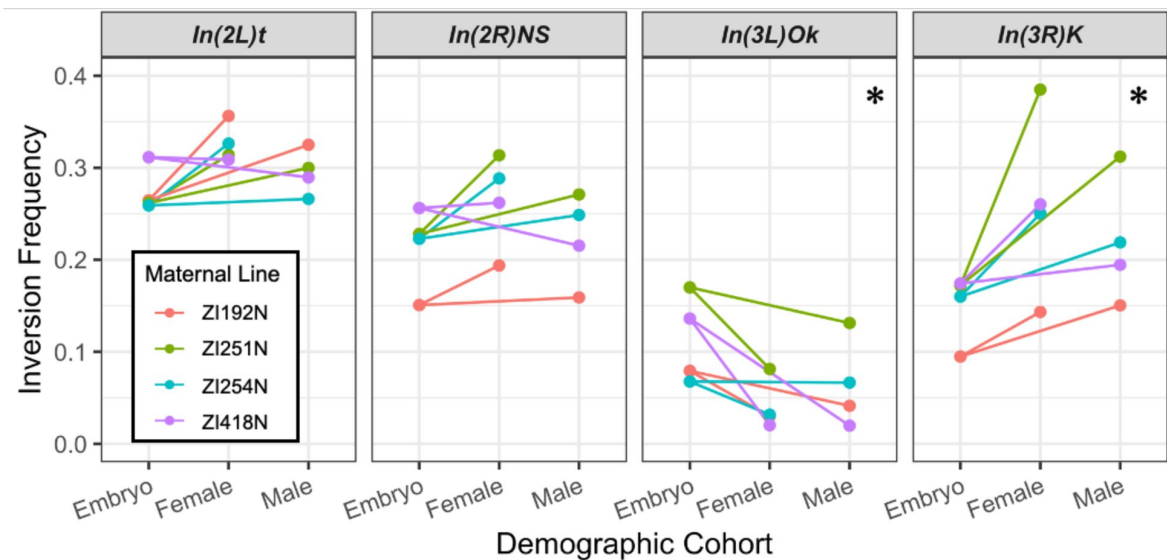


Figure 9

Some *Drosophila* inversions show evidence for antagonistic tradeoffs and sex-specific survival. (A) Testing for antagonistic fitness effects on reproduction and survival. Each inversion's frequency in parents, embryo offspring, and adult offspring are shown for each maternal inbred line. Parental frequency is taken as the paternal frequency divided by two to account for the inversion-free inbred mothers. Adult offspring frequencies are taken from the sum of estimated allele counts across all adult offspring cohorts. * indicates p -value < 0.05 for evidence of significant reversed frequency change from parental to embryo and from embryo to offspring, combined across all four maternal line crosses and corrected for multiple tests, when compared to neutral simulated experiments (details in Materials and Methods). (B) Testing for evidence of sex-specific survival effects. The changes in inversion frequencies between embryos and both female and male adult offspring are illustrated. * indicates p -value < 0.05 for significant differences in female versus male inversion frequency among adult offspring, combined across all four maternal crosses and corrected for multiple tests.

inversion frequencies from fathers to embryos to offspring was less consistent between different crosses (**Figure 9A**), implying that differences in male reproductive success may be dependent on the maternal genotype, which is consistent with past findings in *D. melanogaster* (Reinhart et al. 2015). Despite such variability, each of the four inversions had multiple frequency changes that departed strongly from null expectations (**Figure 9A**; Table S3), as might be expected if selection was present, but varied in strength and direction between crosses and life stages.

To test our central hypothesis of a tradeoff between male reproductive success and cumulative survival during development and adult aging, we evaluated the significance of each change in inversion frequency between the paternal and embryo cohorts, alongside the change in inversion frequency between the embryo and combined adult offspring cohorts. Among these results, *In(3R)K* had particularly consistent changes in frequency across replicate lines, with a decrease between parents and embryos, and then an increase between embryos and aged offspring (**Figure 9A**). When we assessed the likelihood of observing contrasting frequency shifts between life stages of this magnitude, integrating across the four maternal crosses and accounting for multiple test correction, we found that the *In(3R)K* pattern was significantly non-random ($p = 2.70 \times 10^{-3}$; Table S4; see Materials and Methods). The other inversion that gave reversed frequency shifts from cross-averaged data was *In(3L)Ok*, which unlike *In(3R)K*, was associated with (variably) higher male reproductive success and (consistently) lower survival. The frequency reversal observed for *In(3L)Ok* was also found to be statistically significant ($p = 2.97 \times 10^{-7}$). Wright-Fisher estimates of selection coefficients for these and all comparisons are given in Table S8.

We also tested for frequency differences between the different sexes of adult offspring and between the early-eclosing and late-eclosing cohorts of adult offspring. All four inversions displayed differences in frequency between female versus male offspring (**Figure 9B**), consistent with differential effects on sex-specific survival. Strikingly, *In(3L)Ok* and *In(3R)K* each displayed sex survival differences in alignment with their antagonistic effects implied above: *In(3L)Ok* reduced female survival more strongly than male survival ($p = 0.0450$), whereas *In(3R)K* favored the survival of females more than it did males ($p = 4.61 \times 10^{-3}$; Table S4; see Materials and Methods). The other two inversions, *In(2L)t* and *In(2R)NS*, each consistently favored female over male survival in all four maternal crosses, although their initial unidirectional cross-combined p values (0.0260 and 0.0126) were not significant after multiple test correction ($p = 0.282$ and 0.137). In contrast, we found no consistent effects of inversions on eclosion time (Figure S3). Estimated selection coefficients for each tested change based on a Wright-Fisher population approximation are available in Table S7; however these may overestimate inversion fitness effects since the survival terms reflect the probability of reaching advanced age. Further, there may be an influence of differential sex-specific mortality leading to greater frequency shifts in females. We collected a total of 1960 females and 3156 males, and if the survivors represent 20% of an original population with an equal sex ratio, then there was 84.7% female mortality and 75.3% male mortality. Overall, we conclude that *D. melanogaster* inversions have distinct and often context-dependent influences on male reproduction and on survival, with some evidence for antagonistic tradeoffs of the sort proposed in our model.

3 Discussion

Here, we have introduced a new conceptual model in which (1) sexually antagonistic variation is present and subject to balancing selection that maintains extreme genotypes favoring either male reproduction or survival, and (2) inversions are predicted to facilitate the aggregation of progressively more antagonistically differentiated haplotypes. We have tested this model using simulations, confirming the predicted efficacy of balancing selection and demonstrating that inversions indeed facilitate the construction of haplotypes containing multiple antagonistic

variants. Finally, we have deployed *Drosophila* lab experiments to uncover some evidence for inversion-associated tradeoffs between reproduction and survival, depending on the specific inversion and strains tested.

3.1 Inversions in *D. melanogaster*

We have presented results from a single generation laboratory evolution experiment, in which we crossed outbred male *D. melanogaster* carrying inversions *In(2L)t*, *In(2R)NS*, *In(3L)Ok*, and *In(3R)K* to four separate inbred female lines and tracked the frequency of the inversions in the initial males, embryonic offspring, and aged adults. For most inversions, the strength and direction of the tradeoff varied notably between crosses. This might suggest high variance in their phenotypic effects or potentially maternally-dependent variation in male reproductive-success, but is not distinguishable from experimental noise with only one replicate per maternal line. However, greater consistency was observed for *In(3R)K*, which was associated with lesser male reproductive success but greater survival. A significant tradeoff was also inferred for *In(3L)Ok*, which favored male reproduction (albeit variably) over survival. These results provide significant but measured evidence for the presence of sexually antagonistic autosomal inversions in this model species, while emphasizing the importance of genetic background.

As indicated in the Introduction, models like ours in which inversion frequencies are influenced by balancing selection on pleiotropic tradeoffs may help explain multiple aspects of observed *D. melanogaster* inversion frequencies. Adjacent models have been explored in flies, for example an inversion in *Drosophila buzzatii* that was found to trade off between faster development time and larger adults (Betrán et al. 1998 [↗](#)), and an inversion polymorphism found to trade off between larval survival and adult reproduction in the seaweed fly *Coelopa frigida* (Mérot et al. 2020 [↗](#)). Some other groups found to harbor antagonistic inversions include the rainbow trout (Pearse et al. 2019 [↗](#)) and the zebra finch (Pei et al. 2023 [↗](#)). These related systems are variously hypothesized to involve maintenance by overdominance (Mérot et al. 2020 [↗](#)), linked deleterious recessives (Pei et al. 2023 [↗](#)), or reversals of dominance (Pearse et al. 2019 [↗](#)), but these mechanisms are not mutually exclusive. This class of models inherently accounts for the intermediate maximal frequency of most *D. melanogaster* inversions in natural populations. To the extent that inversions favor male reproduction, the sexually antagonistic model we propose here could also help account for the predominantly autosomal locations of common inversions in this species, given that male-specific fitness effects have a weaker influence on the frequencies of X-linked variants since the X chromosome spends only one third of its time in males.

This model could also help explain the observed geographic variability of inversion frequencies, as a change in environment may change the fitness value of either pleiotropic effect of the alleles. If a population experiences a novel harsh environment with low survivability, then the population density may decrease, and the number of encountered competing males may decrease as well, reducing the relative fitness advantage of enhanced male display. If there is a genetic tradeoff between mating display and viability, the equilibrium frequency of an antagonistic inversion that favors the genetic extreme of display should be lower. Additionally, in hostile environments, the survival cost of display-favoring variants could be increased, which could further shift their equilibrium frequencies downward or render them not worth carrying at any frequency. These predictions could contribute to the observed negative relationships between some inversion frequencies and latitude or altitude (Kapun et al. 2016 [↗](#); Pool et al. 2017 [↗](#)). As a specific example, it is notable that in a high altitude Ethiopian environment with year-round cool weather, which hosts some of this species' most pronounced phenotypic evolution (Bastide et al. 2014 [↗](#); Lack et al. 2016b), inversions are virtually absent - occurring at the lowest frequencies of any analyzed population (Sprengelmeyer et al. 2020 [↗](#)). Consistent with such reasoning, a disadvantage of some inversions in challenging thermal environments (both cold and hot) was supported by a previous experimental evolution analysis that included *In(2L)t*, *In(2R)NS*, and *In(3R)K* (Kapun et al. 2014 [↗](#)), though they did find that *In(3R)C* consistently increasing in frequency in hot conditions and *In(3R)Mo* in cold. However, our experimental results include evidence that some of these

inversions, such as *In(3R)K*, may have a tradeoff opposite from our prediction - that the inverted arrangement is associated with survival success instead of reproductive success. Conversely, a recent genome-wide association study found *In(3R)K* heterokaryotypes to be associated with elevated male reproductive success compared to standard homokaryotypes (Yamamoto et al. 2024 [DOI](#)), underscoring that inversion-related tradeoffs may depend on genetic background and/or experimental context.

More broadly, we observed a trend toward increased inversion frequencies across our full experimental generation. A general benefit for inversions would mirror previous findings in which inversions found at intermediate frequencies in natural *D. melanogaster* populations (Inoue 1979 [DOI](#)), as well as inversions from seaweed flies known to be under balancing selection in the wild (Mérot et al. 2020 [DOI](#)), were both inferred to be under directional selection in the lab. However, one laboratory study did report evidence that a *D. melanogaster* inversion not tested here, *In(3R)P*, had frequency-dependent fitness under crowded conditions (Nassar et al. 1973 [DOI](#)). Furthermore, even an inversion found to be under directional selection in the lab might still be subject to a balanced tradeoff a more challenging or complex natural environment. Other studies in *D. melanogaster*, including those on focused seasonal evolution, have found evidence for important tradeoffs between reproduction and robustness to environmental challenges (e.g. Behrman et al. 2015 [DOI](#)). Hence, it would clearly be desirable to study the effects of inversions on a wider range of potential tradeoffs. For example, Said et al. (2018) [DOI](#) suggested that inversions *In(2L)t* and *In(3R)K* might be maintained by immunity-related tradeoffs, in light of an over-representation of immune-related genes among differentially expressed transcripts between inverted and standard karyotypes.

In addition to a broader array of balanced tradeoffs that could modulate *D. melanogaster* inversion frequencies, drift and other forms of selection may contribute as well. Although we emphasize above the potential interplay between ecological adaptation and balanced tradeoffs, some inversions could be impacted by simpler forms of directional selection. However, we must then invoke secondary explanations to account for their limited maximal frequencies (see Introduction). Other forms of balancing selection may also come into play as well. It is clear that some variation in this species is subject to temporally-varying selection, and some inversions have displayed seasonal frequency differences (Dobzhansky 1943 [DOI](#); Kapun et al. 2016 [DOI](#); Machado et al. 2021 [DOI](#); Lange et al. 2022 [DOI](#)). As with spatial clines, temporal shifts could reflect either simple directional selection or else environmentally-modulated tradeoffs in a model such as ours. Spatial selective pressures may vary both between populations and within population scales.

As indicated in the Introduction, balancing selection on inversions could also be achieved by linkage to recessive deleterious variants, which may be observed as an antagonistic pleiotropic phenotype despite potentially involving strictly deleterious variants in linkage (Pei et al. 2023 [DOI](#)). In extreme form inversions might be under associative overdominance (Zhao and Charlesworth 2016 [DOI](#); Gilbert et al. 2020 [DOI](#)), where two haplotypes are balanced by recessive cost to both homozygotes. Inversions may be particularly susceptible to carrying recessive deleterious alleles if population size and/or inversion frequency are low, since under these circumstances inversion homozygotes are rarely generated and hence the crossover rate among inverted chromosomes is low, plus there are fewer inverted chromosomes to facilitate rare double recombination events with standard chromosomes. However, this dynamic is not expected to establish new inversions easily (Charlesworth 2024 [DOI](#)) and is expected to decay with increased age and population size of the inversion, as mutation, gene conversion, or double crossover resolve the negative linkage between recessive deleterious alleles. Most common inversions in *D. melanogaster* are both relatively old and have large effective population sizes. Inversions including *In(2L)t*, *In(2R)NS*, and *In(3R)K* have breakpoints estimated to be several times older than the expansion of the species out of southern-central Africa (Corbett-Detig and Hartl 2012 [DOI](#); Sprengelmeyer et al. 2020 [DOI](#)). Those three inversions, plus *In(3L)Ok*, for which inversion age has not been estimated, all maintain frequencies greater than 0.15 within very large populations (Kapun and Flatt 2019 [DOI](#);

Sprengelmeyer et al. 2020 [↗](#)), and so it is not clear that the persistence of recessive deleterious or lethal alleles in the inverted regions is expected to be high enough that inversions would be maintained by a shared recessive lethal variant preventing fixation. However, sampling of natural populations has demonstrated a somewhat higher frequency of recessive lethal variants in inverted haplotypes, but these are largely unique and heterokaryotypic individuals remain robust, giving low costs to a heavily outbred population (Mukai and Yamaguchi 1974 [↗](#)). Many wild-derived inbred strains of *D. melanogaster* are homokaryotypic for inversions (Lack et al. 2016a). Furthermore, as noted in the Results, an observed lack of inversion homozygotes could reflect strong sexual antagonism as opposed to recessive deleterious variation. Nevertheless, further simulation and experimental testing regarding the potential influence of associative overdominance and other processes on inversion frequencies in *D. melanogaster* and other species would be desirable.

3.2 Generality of Balanced Pleiotropic Inversions

We have proposed that some polymorphic inversions in *D. melanogaster*, and in other species with analogous mating dynamics, may be maintained in part through balancing selection acting on a frequency-dependent tradeoff. While we explored the tradeoff between male reproduction and viability, we emphasize that any antagonistic pleiotropy involving at least one frequency-dependent trait has the potential to generate similar inversion polymorphism. Such antagonistic pleiotropy may be common in evolution, and might be useful in explaining the many examples of inversions associated with distinct phenotypes.

For the specific model explored here, of antagonistic pleiotropy between survival and display in a mate choice system, the prevalence of the dynamic across species and populations is unclear. It seems plausible that a similar model of antagonistic pleiotropy and balancing selection under mate choice underlies other systems, such as the balanced inversion system in the European ruff *Calidris pugnax* (formerly *Philomachus pugnax*), where three inverted haplotypes on an autosome determine aggressive, cooperative, and female-mimic male lekking morphs, albeit with a dominance relationship and homozygous inviability for an inverted haplotype (Küpper et al. 2016 [↗](#)). A recent study on the female-mimic Faeder morph suggests that it may be maintained in part by a balance between its effects lowering female reproductive output while increasing male mating success, and the system is likely more complex as a whole due to the three interacting morphs (Giraldo-Deck et al. 2022 [↗](#)). More generally, a number of factors contribute to the potential prevalence of this dynamic. Mate competition and choice during reproduction is rather common among animals, and this may mean sexually antagonistic inversions are plausibly also common. The model depends on the distribution of fitness effects of new mutations and the availability of sexually antagonistic variation, but such variation does seem prevalent (Zajitschek and Connallon 2018 [↗](#); Ruzicka et al. 2019 [↗](#)). More likely to limit the prevalence of such inversions are the meiotic costs experienced by inversion heterokaryotypes in some systems (White 1973 [↗](#)), whether sexually antagonistic variants are clustered in a region that would allow inversions to increase linkage, and whether other mechanisms like sex-biased expression (Ingleby et al. 2015 [↗](#)), gene duplication (Connallon and Clark 2011 [↗](#)), and linkage to the sex chromosome or sex-determining locus would resolve the antagonism more frequently (though see following subsection on sex chromosome evolution).

Perhaps more importantly, there are a number of well-studied cases of inversion polymorphism that may follow this dynamic of balancing selection between haplotypes simultaneously demonstrating antagonistic pleiotropy and trait-specific frequency dependence, but which do not involve sexual antagonism. *Papilio* butterfly populations harbor well-studied mimicry polymorphisms that segregate between inverted segments and may often experience frequency-dependent selection (Clarke and Sheppard 1960 [↗](#); Kunte et al. 2014 [↗](#); Poul et al. 2014 [↗](#)). The snail *Cepaea nemoralis* harbors diverse and distinct shell patterning haplotypes associated with inversions that remain balanced within single sampling locations (Cook 2013 [↗](#); Surmacki et al.

2013 [↗](#)). Several inversions in different ant species segregate a single-queen colony strategy from a cooperative multi-queen strategy within the same range (Wang et al. 2013 [↗](#); Libbrecht and Kronauer 2014 [↗](#)).

Essentially, any population with antagonistic pleiotropy involving some form of negative frequency-dependent balancing selection on one of the phenotypes might exhibit the same dynamic. The population would experience similar divergent selection towards alternate phenotypic extremes, with the frequency-dependence maintaining the presence of two genotypes. Importantly, this model involves a different mechanism for the maintenance of antagonistically pleiotropic variation than models based on reversals in dominance, where dominance effects generate a net heterozygote advantage at a locus (Connallon and Chenoweth 2019 [↗](#)). The model proposed here integrates well with existing thought on the evolution of divergent selection on balanced systems. For example, Kopp and Hermisson (2006) [↗](#) have proposed a specific form of frequency dependent divergent selection (FDDS), involving competitive exclusion between similar genotypes for a trait otherwise under stabilizing selection, and this model was expected to result in single or few loci of large effect. Their model did not consider pleiotropy or linkage modifiers, but still had comparable dynamics in that selection in the FDDS regime favored small numbers of large effect loci, while the inversion haplotypes in the model presented here simplify the recombination architecture to approximate a single locus. This simplification of the genetic architecture follows the reduction principle in recombination modifier theory (Feldman 1972 [↗](#); Feldman and Balkau 1973 [↗](#); Feldman et al. 1980 [↗](#); Altenberg and Feldman 1987 [↗](#); Altenberg et al. 2017 [↗](#)). When a population experiences selection towards two (or more) divergent optima, under a selection regime that provides a balancing mechanism to maintain both, indirect selection will favor a reduction of the genetic architectures to prevent the generation of intermediate phenotypes. The likelihood of either model likely rests on the inherent polygenicity of the trait under selection, with simpler traits being more likely to resolve by single locus FDDS dynamics than by inversion association (e.g. Yassin et al. 2016 [↗](#)).

Notably, our simulations tended to result in strong sexual antagonism, such that it leads to a substantial fraction of the population failing to survive to reproductive age. It is unclear how common such extreme antagonism may be in natural populations. If it is uncommon in nature, even in taxa that experience high levels of mate competition as modeled here, then it is unclear what prevents our simulation predictions from being realized. It may be that genetic constraints prevent the occurrence of highly antagonistic genotypes, or that the relationship between survival-altering mutations and fitness is different than we have modeled. Or, populations may steadily evolve to mitigate antagonistic effects that arise and persist through processes such as those we modeled. Alternatively, some highly antagonistic haplotypes may become associated with sex determination, as suggested below.

3.3 Antagonistic autosomal inversions may precede novel sex chromosomes

If balanced, sexually antagonistic, autosomal inversions occur frequently enough, they may also contribute to the evolution of sex chromosomes. Given an established inversion with significant sexual antagonism, the inverted haplotype that favors one sex still experiences an antagonistic cost of transmission to offspring of the opposite sex. Mutations that link the haplotype with the sex it benefits are therefore favored under selection, and this can occur either by fusing to an existing sex chromosome (Charlesworth and Charlesworth 1980 [↗](#)), though this only fully resolves the antagonism for linkage to Y and W chromosomes, or by linkage to a new sex determining mutation (van Doorn and Kirkpatrick 2007 [↗](#)).

Here, we discovered that strong inversion-linked antagonism could yield a scenario where essentially all highly-successful males were homokaryotypic for a display-favoring karyotype, whereas reproducing females were overwhelmingly heterokaryotypic (Figure 5 [↗](#)). This situation

is similar to the ZW sex chromosome system, suggesting that this particular pleiotropic trade-off may favor resolution by linkage to a female sex determination factor. Further, simulations in which only females experienced antagonistic survival cost usually generated populations with a chromosome essentially lethal in females, thus exclusively transmitted in males similar to a Y chromosome (**Figure 7**). These simulations sometimes had three states that appeared stable at $20N$ generations, including Y chromosome-like, W chromosome-like, and shared intermediate arrangements. This three-arrangement stable polymorphism seems unusual for a natural population, but might be interesting to explore in populations with polymorphic sex determination.

There are often deleterious effects of chromosome fusion or novel sex determination (Pennell et al. 2015; Saunders et al. 2019), which may represent a general obstacle to sex chromosome evolution. However, balanced sexually antagonistic inversions may concentrate enough antagonism that the built-in benefit of resolving it may sometimes overcome the fitness costs of evolving new sex chromosomes (Blackmon and Brandvain 2017). Hence, antagonistic inversions might be more likely than other loci to be involved in the evolution of new sex chromosomes.

A basic prediction of the above hypothesis is that at least one inversion difference between sex chromosomes should already exist at the time of new sex chromosome formation, and genomic differentiation therefore begins between inversion karyotypes on an autosome even before sex-specific transmission of the region. This scenario contrasts with current models of sex chromosome evolution in which inversion accumulation and genomic differentiation accrue only after sex chromosome formation, due to advantages in suppressing recombination involving sexual antagonistic variants already associated with a sex-determining locus (Wright et al. 2016). In this context, we note that the neo-X and neo-Y homologs of Muller element C in *Drosophila miranda*, which became sex-linked about 1.5 million years ago, already appear to have inversion differences between them (Chen et al. 2024 – **Figure 5A**). Further, a recent study in *Littorina saxatilis* snails has discovered an ecotype-specific sex determination system that may be founded on such tradeoffs, though selection across ecotypes involves discrete geographic boundaries and is only partly comparable (Hearn et al. 2022). Still, additional genomic studies of new sex chromosomes are needed to assess the potential role of sexually antagonistic inversions in sex chromosome evolution.

Summary and Future Prospects

We have advanced a model in which inversion polymorphism is maintained by the frequency-dependent fitness effects of a pleiotropic tradeoff associated with inversion-linked genetic variation. We have used sexual antagonism associated with increased male reproduction versus survival as a specific motivating case, and we explore the predictions of this model using a novel simulation algorithm. We confirmed model predictions that some antagonistic variants can be maintained at an equilibrium frequency, and that inversions can allow multiple antagonistic variants to persist and form more strongly antagonistic haplotypes. We complemented our conceptual and computational modeling with laboratory experiments designed to detect fitness effects of four inversions common in an ancestral range population of *D. melanogaster*. We indeed found some evidence of tradeoffs between male reproduction and survival for these inversions, with two out of four inversions showing significant evidence for such a tradeoff, although we also observed a strong dependence of male reproductive success on female genetic background.

In light of the persistent mystery of how *D. melanogaster* inversion frequencies are determined, and how readily we detected tradeoffs involving male reproduction and survival, we suggest that further studies investigating the potential fitness tradeoffs of inversions in this model system are strongly warranted. While some such studies have previously been conducted, the applicability of our amplicon sequencing approach to estimate inversion frequencies in large pools of flies may improve the sensitivity of such experiments. Our understanding of the phenotypic impacts of *D.*

melanogaster inversions would benefit from investigations of a wider range of fitness-related traits, including female fecundity, immunity, and performance in challenging environments such as cool temperatures and high population densities.

Clearly, it will also be important to test the generality of our model by estimating the fitness associations of inversions in a wider range of species. And particularly in the case of species with very recent changes sex chromosomes, it will be of great interest to test whether one or more inversions are already present between these new sex chromosomes, as predicted a model in which sex chromosome evolution involves the resolution of inversion-associated sexual antagonism. Such studies will advance our broader understanding of the roles of balancing selection, pleiotropic tradeoffs, and genome structure in shaping genetic variation and genomic evolution.

4 Materials and Methods

4.1 Simulations

We developed a new individual-level forward simulator ('SAIsim' for Sexually Antagonistic Inversion simulator) written in python 3.7 for the simulation of inversions and sexually antagonistic alleles at infinite loci, with uniform rates of mutation, crossing over, and gene conversion (per bp, per generation) over a number of independent chromosome arms. The dynamic we were interested in modeling requires direct accounting of pleiotropy between different components of fitness, crossing over and gene conversion within inversions (as well as the possibility of inversion fixation), and directly modeling male display and female choice, and no existing simulator provided this functionality. These simulations were used to establish the potential for the model to occur in a population generally, and not to compare against specific empirical observations. We have used literature estimates of mutation rate and conversion rate of *Zambian D. melanogaster* for these simulations (Comeron et al. 2012 [\[1\]](#); Huang et al. 2016 [\[2\]](#)) but have no information on the joint distribution of effects of mutations on our model of survival and display in *D. melanogaster*. So, we sample from relatively broad but arbitrary combinations of mutation effects, only assuming that there is likely some degree of antagonistic pleiotropy naturally present in genomic variation. In SAIsim, a population is instantiated as a python object, and populated with individuals (also python objects) based on provided (or initially empty) genomes. In each generation, as diagrammed in **Figure 5** [\[3\]](#), the simulator calculates a survival probability for each individual as the product of the survival values of each variant they carry, which range [0,1]. Here, multiplicative selection for survival prevents biologically unreasonable negative survival probabilities from occurring. Male and female survival costs can each be included in a sex-specific manner, but only globally in the simulation, not per mutation. For each offspring in the next generation, a mother is selected randomly, weighted by individual survival probabilities, with replacement. To model the encountered males that might become the father of an offspring individual from this female, a set number of males (i.e. the uniformly specified encounter number) is randomly sampled from the full population of males, weighted by their survival probabilities. The males each have a genetic display quality given as the sum of the reproductive values of their alleles. A noise parameter, normally distributed around zero with a standard deviation of one, is added to each male display value during each encounter. The male with the highest noise-adjusted display quality among the encountered males is chosen to be the father. As offspring are generated, crossover locations are sampled, and resampled in cases where the resultant gamete would be aneuploid. Resampling aneuploids removes the fecundity cost that may exist in some taxa, but was chosen to reflect the lack of observed fecundity effects for most paracentric inversions of *D. melanogaster*. Gene conversion is modeled as a constant probability per heterozygous variant, with each event independently affecting just one locus, rather than modeling a conversion tract length (given that gene conversion tracts are generally short and each would typically contain only a single modeled variant in our simulations). This probability reflects

the likelihood of any gene conversion event affecting a single nucleotide site, rather than the per-bp rate of initiation of gene conversion tracts. The dynamics of gene conversion are expected to have separate effects on the assembly and the maintenance of arrangements. The assembly of an arrangement is positively dependent on the population-scaled rate of recombination of alleles as new polymorphisms start at low frequency and often not on their optimal haplotype. However, the indirect fitness effect of the inversion is dependent on its effect on recombinant offspring, and therefore negatively depends on the per-individual rate of conversion. The simulation is intended to represent *D. melanogaster* life history and allows removal of recombination in males, which also removes gene conversion in males by default. Finally, inversions and new mutations are potentially added to the resultant gamete, with inversions resampled when overlapping the coordinates of an inversion already present on the same individual chromosome. SAIsim also allows the specification of mutation rates and the distribution of effect sizes of new mutations.

The simulations presented in **Figures 5**, **6**, and **7** rescale parameter values by a factor of 2,000 to model an estimated ancestral N_e of around 2×10^6 (Sprenkelmeyer et al. 2020) while using a smaller simulated population of 1,000 individuals. The low crossover rate simulations model a 0.0218 cM region and the high crossover rate simulations model 2.18 cM, which scale to 43.6 cM and 43.6 M female meiotic maps in the simulated populations. Based on an average autosomal recombination rate in females of 2.18×10^{-8} for *D. melanogaster* (Comeron et al. 2012), these represent 10kb and 1Mb regions. Gene conversion occurs at each heterozygous variant with a probability of 0.1295, based on scaling $\gamma = 1.25 \times 10^{-7}$ events/bp/female meiosis with a conversion track length of 518 bp (Comeron et al. 2012) to a per-single-nucleotide-variant estimate of 6.475×10^{-5} conversions/bp/female meiosis. Sexually antagonistic mutations were modeled at a rate corresponding to one per thousand mutations, and so based on an autosomal estimate of 5.21×10^{-9} mutations/bp/gen on autosomes (Huang et al. 2016), and modeling 10kb and 1Mb regions, these were scaled to 1.042×10^{-4} and 1.042×10^{-2} events per gamete respectively. Inversions were approximated as arising once per hundred sexually antagonistic mutations, for 1.042×10^{-6} and 1.042×10^{-4} events per gamete. Both of these rates are empirically underdetermined. The simulations use mutations with uniformly distributed quality values in the range [0,1]; although distributions skewed toward small effects may be more realistic, our uniform approach avoids the need to invoke a specific but empirically unknown distribution. SAIsim's object-oriented design allows specification of a population model by a new user familiar with the basics of python scripting. A population object can be parameterized and populated in a few lines of python script, and then parameters can be changed between periods of simulation, allowing bottlenecks and demographic changes as well as changes in mutational parameters between generations. Replicate SAIsim simulations were performed via the UW-Madison Center For High Throughput Computing (CHTC), which manages cycle servers using the HTCondor distributed computing project (Thain et al. 2005; Erlandson and Theisen 2018).

4.2 Experimental Populations and Sequencing

To act as a high-inversion source population for the potential fathers in our experiment, we combined 15 males and 15 females from each of 10 inbred strains derived from wild flies collected in Zambia. These 10 strains were chosen to collectively carry the following common inversions: *In(2L)t*, *In(2R)NS*, *In(3L)Ok*, *In(3R)K*. We then collected at least 500 F2 males from this high-inversion population to use in each of our independent maternal line experiments, and these males were expected to carry a wide range of inversion genotypes. Each batch of high-inversion males was crossed to an equal number of virgin females from a single inbred Zambian strain (**Figure 6**). Four such experiments were performed, each using one inversion-free maternal strain. Since no inversions could be inherited through the mothers, inversion frequencies of successfully reproducing F2 males could be inferred from their pooled offspring. Unless otherwise noted, all flies were kept in a lab space of 23°C with around a degree of temperature fluctuation and without a controlled day/night cycle. All fly food used was prepared in batches of 4.5 L water, 500 mL cornmeal, 500 mL molasses, 200 mL yeast, 54 g agar, 20 mL propionic acid, and 45 mL tegosept 10% (in 95% ethanol). The four parental populations were each allowed to mate freely in

clear plexiglass cages of dimension 19.3 by 19.3 by 30.0 cm, with a loose mesh covering on one end for access by hand, twisted and clipped shut with a binder clip. The parents were given either fly food bottles or grape agar in 7.9 cm petri dishes as food or laying substrate on alternating days over two weeks, after which all surviving male parents were counted and frozen. The grape agar was prepared in batches of 500 mL from *FlyStuff.com* [FlyStuff.com](#) grape agar powder, and wet Red Star yeast was brushed over the agar plates before use to encourage laying. Embryos were collected from the agar, counted, and frozen for sequencing. Embryos laid in the bottles were allowed to hatch and develop into adults. These adult offspring were separated by sex and by the time after laying at which they eclosed. In our husbandry conditions, the earliest eclosion was 11 days after laying, but this was rare and a first eclosion at 12 days was more common among the bottles. Most flies took several days longer to develop, and we considered flies collected in the first 6 days of a 10 day collection window to be the early-eclosing pool, in order to split the pool sample sizes closer to evenly. The collected adult flies were then aged by sequentially transferring them between bottles weekly for approximately 2.5 months, and then the adult offspring were sexed, counted, and collected. A 10-week aging window was chosen based on the time taken for a control set of four bottles of flies from the paternal outbred population to reach 80% mortality. All adult and embryo samples were frozen at -80°C before DNA extraction. This sampling scheme was chosen to identify inversion frequencies in the set of potential parents, in embryonic offspring, and in older adults, and so to assess potentially selected changes in inversion frequency across male reproduction or in egg to adult (encompassing viability plus longevity). We chose to age the populations to include longevity as long term survival is relevant both in the estimated reproductive age of the species (Pool 2015 [Pool 2015](#)), and the need to survive during periods of unsuitable environments, often while under reproductive diapause (Saunders et al. 1990 [Saunders et al. 1990](#); Ragland et al. 2019 [Ragland et al. 2019](#)). In seasonally varying climates, either temperate or with long dry season, survival through challenging conditions is expected to require several months. In many such cases females are in reproductive diapause, and so longevity is the main selective pressure.

DNA for each replicate age group was extracted by a phenol-chloroform extraction protocol used previously in preparing DNA for the *Drosophila* Genome Nexus and *Drosophila* Population Genomics Project. The protocol is similar to protocol ‘BPC’ used in DPGP phase 1 (Langley et al. 2012 [Langley et al. 2012](#)), except with the addition of proteinase-K during homogenization of adult flies, and generally based on the common quick preparation protocol presented in *Drosophila Protocols* (Huang et al. 2009 [Huang et al. 2009](#)). Inversion genotyping was performed by amplicon sequencing, where amplicons associated with each inversion were designed to target at least one single nucleotide polymorphism in the inversion interior that is fixed for alternate alleles between the inversion karyotypes, while still sharing conserved primer regions across karyotypes to prevent primer bias. Genetic variation in these inversions was assessed using previously sequenced and inversion-called haploid genomes from the Zambian *D. melanogaster* population, available from the *Drosophila* Genome Nexus (Lack et al. 2016a). The amplicon region and primer selection was performed with a custom script. The script calls Primer3 version 2.4.0 for primer candidate generation (Koressaar and Remm 2007 [Koressaar and Remm 2007](#); Untergasser et al. 2012 [Untergasser et al. 2012](#); Koressaar et al. 2018 [Koressaar et al. 2018](#)). Primers thus selected were ordered from IDT as oligos ligated to Illumina stubby adaptors [Table S1]. Illumina sequencing libraries were then prepared for each sample following the Illumina metagenomic amplicon sequencing protocol (Illumina 2013 [Illumina 2013](#)), with 50 ng/uL diluted DNA samples. and sequenced 2×150 on an Illumina NovaSeq to at least a depth of 20000 reads. Depth was selected to ensure high likelihood of detecting a 0.025 change in inversion frequency from an initial frequency of 0.10, based on simulated resampling of the collected flies, DNA extraction, and sequencing.

4.3 Sequence Analysis and Statistics

Sequence preparation and analysis were performed using custom pipeline scripts. Read trimming, alignment, and QC were performed using bwa version 0.7.17-r1188 (Li 2013 [Li 2013](#)), GATK version 3.4-46 (Van der Auwera and O’Connor 2020 [Van der Auwera and O’Connor 2020](#)), and Picard version 1.79 (Broad Institute 2019 [Broad Institute 2019](#)). Reads were trimmed at the ends to remove segments with base quality less than 20, and read pairs were

filtered to retain only those pairs where both reads aligned with mapping quality >20. SNP identity at the sites of interest was called from those remaining reads with base quality >20 at the SNP of interest. Due to the presence of chimeric reads in the amplicon library sequencing, we chose to use a single discriminant SNP to determine each inversion's frequency in the read pool, instead of incorporating haplotype or additional SNP information.

To assess the significance of the difference in inversion frequency between just two sampled life stages in the same experiment, we first calculated a chi-square test statistic from the two-by-two table of inverted and non-inverted allele counts at two life stages. Assessment of the significance of frequency change reversals across three cohorts was done with concordant methods discussed subsequently. Allele counts were taken by multiplying the inversion-specific SNP frequency in the sequencing reads by the allele sample size. We then generated an approximate distribution of the chi-square statistic when selection is absent (our null hypothesis), using a resampling approach (Figure 6). Comparing the observed statistic to our modeled neutral distribution of test-statistic values allowed us to identify significance thresholds. Note that we are not conducting chi-square tests, but simply using the associated statistic as a convenient metric for comparing empirical and simulated data.

To generate the neutral distribution of the chi-square statistic, we took the observed paternal inversion frequencies and then sampled from them to obtain new paternal and offspring read counts to represent a single neutral data set as follows. First, we resampled the collected allelic counts (given in Tables S6 and S7) to account for sampling variance between replicates. For the paternal samples, we needed to account for pre-sequencing mortality as up to 15% of individuals died before collection. So, in generating a new paternal read set, we first modeled down-sampling of the paternal allele sample from the initial census allelic count to the collected allelic count as a hypergeometric sampling. For all other cohorts we used a binomial resampling from the collected allelic count. Second, we used a binomial sampling step to account for the experimental variance introduced by the extraction and library preparation. The extraction and library preparation appeared to introduce variation independent of allele sample size, possibly due to variation introduced during PCR or sample homogenization. This inference was based on analysis of data from control library preparations (Table S5) and from sequencing done in parallel with our experiment from sample fly pools with known inversion frequencies. These control samples included an average of 6 technical replicates (separate library prep and sequencing replicates from the same flies) from four different expected frequencies for *In(2L)t*, as well as one expected frequency for each of *In(2R)NS* and *In(3R)K*. Due to genotyping and husbandry difficulties we did not have a control set for *In(3L)Ok*. Based on the inversion frequency differences observed among these technical replicates, we estimated that taking a binomial sampling of size 316 from the sample allele frequency reflected a maximum likelihood model for the sample-size-independent variation introduced by the library prep and sequencing process. Third, from the preliminary resampled paternal inversion frequency, we then took a binomial sample of size equal to the read count to account for the random sample of sequencing reads, thus yielding resampled paternal read counts (given in Tables S6 and S7) that accounted for the number of sampled and non-sampled individuals, experimental variance, and depth of coverage. Fourth, to generate resampled offspring read sets, we instead used a binomial sampling of half the offspring alleles from the fully resampled paternal allele frequency for each separate offspring cohort, since the homozygous mothers all contributed non-inverted alleles. Finally, each offspring allele sample was resampled using the same size 316 binomial sampling to account for the variation introduced during this library's preparation, and binomial sampled again with the read count size to account for sampling reads from the library, thus yielding resampled offspring read counts that accounted for the same three factors as for paternal counts.

For a pairwise comparison of inversion frequency between two samples, we calculated the same allele count chi-square statistic as described above for the empirical data for each resampled set of allele counts, generating the full distribution of neutral chi-square test statistics from which we

could identify extreme value cutoffs. We then obtained a p-value from the proportion of resampled replicates in which the chi-square statistic was greater than or equal to the empirical value. If the test was directional, we took the proportion of resampled replicates with both a chi-square statistic greater than or equal to the empirical value and a frequency difference in the expected direction.

In order to identify a frequency change reversal between life history stages that could reflect a fitness tradeoff, we tested for the joint significance of two opposing frequency changes using three cohorts. A tradeoff would require a frequency change in one direction between paternal and embryo samples, followed by a frequency change in the opposite direction between embryo and adult offspring samples. These changes both depend on the shared embryonic cohort in the test, and so are not independent. Therefore, we evaluated the two inversion frequency changes jointly from a set of resampled neutral data sets in which the three cohorts arose from the same resampling process. Here we considered how likely it was that a sample of the three cohorts from the associated estimated neutral model had chi-square values for both the comparison of the fathers to embryos and the embryos to combined adults that were equivalent or more extreme in the test direction than the observed data. For example, when testing the significance of an observed paternal-embryo increase and then embryo-adult decrease in frequency, our p-value would be defined by the proportion of resampled replicates in which the frequency changes were in the tested directions of increase then decrease, and their chi-square values were both more extreme than the empirical values.

To assess the significance of inversion frequency changes considered across all four maternal line crosses at once, we used Fisher's combined probability test. We applied this to both particular three-cohort frequency change reversals and to the set of two-cohort frequency changes. The inversion trajectory is independent between maternal crosses, but the differences between each cross in maternal line and initial paternal genetic variation make the replicates difficult to compare in a more structured way, so we felt this was an appropriate test for combined significance. In assessing our tradeoff hypothesis, we applied this test only for inversions where it was well-motivated; *i.e.* where the averaged frequency change across maternal crosses was in opposite directions between paternal/embryo and embryo/adult comparisons.

To account for multiple hypothesis testing in our cohort comparisons, we used a Benjamini-Yekutieli false discovery rate correction for significance. We applied this correction across the set of Fisher's method combined p-values we generated to assess the relevant frequency change reversal hypotheses. So, this corrected for the multiple hypotheses of different tradeoff directions and different inversions involved, from p-values that already combined evidence from different maternal line crosses. Benjamini-Yekutieli was chosen due to its robustness to the unclear dependence relationships between some of the tests involved. For example, multiple tests of frequency changes from a single cohort (*e.g.* an embryo sample compared against both paternal and aged adult samples) are likely to be positively correlated.

Data Availability

Sequencing reads have been uploaded to the NIH Short Read Archive under project XXXXXX. The simulation program and other analysis scripts can be found at <https://github.com/csmcal/>.

Acknowledgements

Particular thanks to Shimeng Gao, who assisted with *Drosophila* husbandry and collection of adult flies during the within-generation experiment. We also thank members of the Pool lab and three anonymous reviewers for helpful comments on earlier versions of this manuscript.

This research was supported by National Science Foundation Graduate Research Fellowship (DGE-1747503 to CSM), and by the National Institutes of Health (grants R35 GM136306 to JEP, T32 GM007133, and T32 HG002760).

This research was performed using the compute resources and assistance of the UW-Madison Center for High Throughput Computing (CHTC) in the Department of Computer Sciences. The CHTC is supported by UW-Madison, the Advanced Computing Initiative, the Wisconsin Alumni Research Foundation, the Wisconsin Institutes for Discovery, and the National Science Foundation, and is an active member of the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

References

- Altenberg L, Feldman MW (1987) **Selection, generalized transmission and the evolution of modifier genes. I. The reduction principle** *Genetics* **117**:559–572
- Altenberg L, Liberman U, Feldman MW (2017) **Unified reduction principle for the evolution of mutation, migration, and recombination** *Proc. Natl. Acad. Sci* **114**:E2392–E2400
- Audet T, Krol J, Pelletier K, Stewart AD, Dworkin I (2024) **Sexually discordant selection is associated with trait specific morphological changes and a complex genomic response** *Evolution:qpae* **71**
- Aulard S, David JR, Lemeunier F (2002) **Chromosomal inversion polymorphism in Afrotropical populations of *Drosophila melanogaster*** *Genet. Res* **79**:49–63
- Babcock CS, Anderson WW (1996) **Molecular evolution of the Sex-Ratio inversion complex in *Drosophila pseudoobscura*: analysis of the Esterase-5 gene region** *Mol. Biol. Evol* **13**:297–308
- Bastide H, Ogereau D, Montchamp-Moreau C, Gérard PR (2022) **The fate of a suppressed X-linked meiotic driver: experimental evolution in *Drosophila simulans*** *Chromosome Res* **30**:141–150
- Bastide H, Yassin A, Johanning EJ, Pool JE (2014) **Pigmentation in *Drosophila melanogaster* reaches its maximum in Ethiopia and correlates most strongly with ultra-violet radiation in sub-Saharan Africa** *BMC Evol. Biol* **14**
- Bateman AJ (1948) **Intra-sexual selection in *Drosophila*** *Heredity* **2**:349–368
- Behrman EL, Watson SS, O'Brien KR, Heschel MS, Schmidt PS (2015) **Seasonal variation in life history traits in two *Drosophila* species** *J. Evol. Biol* **28**:1691–1704
- Berdan EL *et al.* (2023) **How chromosomal inversions reorient the evolutionary process** *J. Evol. Biol* **36**:1761–1782
- Betrán E, Santos M, Ruiz A (1998) **Antagonistic Pleiotropic Effect of Second-Chromosome Inversions on Body Size and Early Life-History Traits in *Drosophila buzzatii*** *Evolution* **52**:144–154
- Blackmon H, Brandvain Y (2017) **Long-Term Fragility of Y Chromosomes Is Dominated by Short-Term Resolution of Sexual Antagonism** *Genetics* **207**:1621–1629
- Blows MW, Brooks R (2003) **Measuring Nonlinear Selection** *Am. Nat* **162**:815–820
- Broad Institute (2019) **Picard Toolkit Available from**
- Charlesworth B. (1974) **Inversion polymorphism in a two-locus genetic system** *Genet. Res* **23**:259–280

- Charlesworth B (1997) **The evolution of lethals in the t-haplotype system of the mouse** *Proc. R. Soc. B Biol. Sci* **258**:101–107
- Charlesworth B (2024) **The fitness consequences of genetic divergence between polymorphic gene arrangements** *Genetics* **226**
- Charlesworth B, Barton NH (2018) **The Spread of an Inversion with Migration and Selection** *Genetics* **208**:377–382
- Charlesworth B, Charlesworth D (1973) **Selection of new inversions in multi-locus genetic systems** *Genet. Res* **21**:167–183
- Charlesworth B, Charlesworth D (2010) **Elements of Evolutionary Genetics** *W. H. Freeman*
- Charlesworth B, Hughes KA, Singh RS, Krimbas CB (2000) **The maintenance of genetic variation in life-history traits** *Evolutionary Genetics: from Molecules to Morphology* :369–392
- Charlesworth D, Charlesworth B (1980) **Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes** *Genet. Res* **35**:205–214
- Chen KH-C, Chatla K, Bachtrög D (2024) **Single-cell RNA-seq of Drosophila miranda testis reveals the evolution and trajectory of germline sex chromosome regulation** *PLOS Genet* **22**
- Cheng C, Kirkpatrick M (2016) **Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies** *PLOS Genet* **12**
- Clarke CA, Sheppard PM (1960) **Super-genes and mimicry** *Heredity* **14**:175–185
- Comeron JM, Ratnappan R, Bailin S (2012) **The Many Landscapes of Recombination in Drosophila melanogaster** *PLOS Genet* **8**
- Connallon T, Chenoweth SF (2019) **Dominance reversals and the maintenance of genetic variation for fitness** *PLOS Biol* **17**
- Connallon T, Clark AG (2011) **The resolution of sexual antagonism by gene duplication** *Genetics* **187**:919–937
- Cook LM. (2013) **Selection and disequilibrium in Cepaea nemoralis** *Biol. J. Linn. Soc* **108**:484–493
- Corbett-Detig RB (2016) **Selection on Inversion Breakpoints Favors Proximity to Pairing Sensitive Sites in Drosophila melanogaster** *Genetics* **204**:259–265
- Corbett-Detig RB, Hartl DL (2012) **Population Genomics of Inversion Polymorphisms in Drosophila melanogaster** *PLOS Genet* **8**
- Courret C, Chang C-H, Wei KH-C, Montchamp-Moreau C, Larracuent AM (2019) **Meiotic drive mechanisms: lessons from Drosophila** *Proc. R. Soc. B Biol. Sci* **286**
- Coyne JA, Meyers W, Crittenden AP, Sniegowski P (1993) **The Fertility Effects of Pericentric Inversions in Drosophila Melanogaster** *Genetics* **134**:487–496

- Cox RM, Calsbeek R (2009) **Sexually Antagonistic Selection, Sexual Dimorphism, and the Resolution of Intralocus Sexual Conflict** *Am. Nat* **173**:176–187
- da Silva Ribeiro T, Galván JA, Pool JE (2022) **Maximum SNP FST Outperforms Full-Window Statistics for Detecting Soft Sweeps in Local Adaptation** *Genome Biol. Evol* **14**
- Dickerson GE (1955) **Genetic slippage in response to selection for multiple objectives** *Cold Spring Harbor Symposia on Quantitative Biology* **20**:213–224
- Dobzhansky TG (1943) **Genetics of Natural Populations IX. Temporal Changes in the Composition of Populations of *Drosophila pseudoobscura*** *Genetics* **28**:162–186
- Dobzhansky TG (1949) **Observations and Experiments on Natural Selection in *Drosophila*** *Hereditas* **35**:210–224
- Dobzhansky TG (1950) **Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*** *Genetics* **35**:288–302
- Dobzhansky TG (1970) **Genetics of the Evolutionary Process**
- van Doorn GS, Kirkpatrick M. (2007) **Turnover of sex chromosomes induced by sexual conflict** *Nature* **449**:909–912
- Erlandson E, Theisen T. (2018) **Erlandson E, Theisen T. 2018. htcondor: HTCondor source repository, formerly the Condor Project. Available from: <https://github.com/htcondor/htcondor>**
- Faria R, Johannesson K, Butlin RK, Westram AM (2019) **Evolving Inversions** *Trends Ecol. Evol* **34**:239–248
- Feldman MW. (1972) **Selection for Linkage Modification: I. Random Mating Populations** *Theor. Pop. Biol* **3**:324–346
- Feldman MW, Balkau B (1973) **Selection for Linkage Modification: II. A Recombination Balance for Neutral Modifiers** *Genetics* **74**:713–726
- Feldman MW, Christiansen FB, Brooks LD (1980) **Evolution of recombination in a constant environment** *Proc. Natl. Acad. Sci* **77**:4838–4841
- Frydenberg O (1963) **Population Studies of a Lethal Mutant in *Drosophila melanogaster*** *Hereditas* **50**:89–116
- Gilbert KJ, Pouyet F, Excoffier L, Peischl S (2020) **Transition from Background Selection to Associative Overdominance Promotes Diversity in Regions of Low Recombination** *Curr. Biol* **30**:101–107
- Gillespie JH (1998) **Population Genetics: A Concise Guide**
- Giraldo-Deck LM, Loveland JL, Goymann W, Tschirren B, Burke T, Kempnaers B, Lank DB, Küpper C (2022) **Intralocus conflicts associated with a supergene** *Nat. Commun* **13**
- Glaser-Schmitt A, Ramnarine TJS, Parsch J (2023) **Rapid evolutionary change, constraints and the maintenance of polymorphism in natural populations of *Drosophila melanogaster*** *Mol. Ecol*

Gong WJ, McKim KS, Hawley RS (2005) **All Paired Up with No Place to Go: Pairing, Synapsis, and DSB Formation in a Balancer Heterozygote** *PLOS Genet* **1**

Gowaty PA, Kim Y-K, Anderson WW (2012) **No evidence of sexual selection in a repetition of Bateman's classic study of *Drosophila melanogaster*** *Proc. Natl. Acad. Sci. U. S. A* **109**:11740–11745

Grieshop K, Arnqvist G (2018) **Sex-specific dominance reversal of genetic variation for fitness** *PLOS Biol* **16**

Hazel LN (1943) **The Genetic Basis for Constructing Selection Indexes** *Genetics* **28**:476–490

Hearn KE, Koch EL, Stankowski S, Butlin RK, Faria R, Johannesson K, Westram AM (2022) **Differing associations between sex determination and sex-linked inversions in two ecotypes of *Littorina saxatilis*** *Evolution Letters* **6**:358–374

Hoffmann AA, Rieseberg LH (2008) **Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation?** *Annu. Rev. Ecol. Evol. Syst* **39**:21–42

Hoquet T, Bridges WC, Gowaty PA (2020) **Bateman's Data: Inconsistent with "Bateman's Principles."** *Ecol. Evol* **10**:10325–10342

Huang AM, Rehm EJ, Rubin GM (2009) **Quick Preparation of Genomic DNA from *Drosophila*** *Cold Spring Harb. Protoc* **2009**

Huang W, Lyman RF, Lyman RA, Carbone MA, Harbison ST, Magwire MM, Mackay TF (2016) **Spontaneous mutations and the origin and maintenance of quantitative genetic variation** *eLife* **5**

Illumina (2013) **Illumina. 2013. 16S metagenomic sequencing library preparation guide. Available from:**
https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf

Ingleby FC, Flis I, Morrow EH (2015) **Sex-Biased Gene Expression and Sexual Conflict throughout Development** *Cold Spring Harb. Perspect. Biol* **7**

Innocenti P, Morrow EH (2010) **The Sexually Antagonistic Genes of *Drosophila melanogaster*** *PLOS Biol* **8**

Inoue Y (1979) **THE FATE OF POLYMORPHIC INVERSIONS OF *DROSOPHILA MELANOGASTER* TRANSFERRED TO LABORATORY CONDITIONS** *Jpn. J. Genet* **54**:83–96

John A, Vinayan K, Varghese J (2016) **Achiasmy: Male Fruit Flies Are Not Ready to Mix** *Front. Cell Dev. Biol* **4**

Jones AG, Arguello JR, Arnold SJ (2002) **Validation of Bateman's principles: a genetic study of sexual selection and mating patterns in the rough-skinned newt** *Proc. R. Soc. B Biol. Sci* **269**:2533–2539

Kapun M, Fabian DK, Goudet J, Flatt T (2016) **Genomic Evidence for Adaptive Inversion Clines in *Drosophila melanogaster*** *Mol. Biol. Evol* **33**:1317–1336

- Kapun M, Flatt T (2019) **The adaptive significance of chromosomal inversion polymorphisms in *Drosophila melanogaster*** *Mol. Ecol* **28**:1263–1282
- Kapun M, Mitchell ED, Kawecki TJ, Schmidt P, Flatt T (2023) **An Ancestral Balanced Inversion Polymorphism Confers Global Adaptation** *Mol. Biol. Evol* **40**
- Kapun M, van Schalkwyk H, McAllister B, Flatt T, Schlötterer C (2014) **Inference of chromosomal inversion dynamics from Pool-Seq data in natural and laboratory populations of *Drosophila melanogaster*** *Mol. Ecol* **23**:1813–1827
- Kirkpatrick M, Barton N (2006) **Chromosome Inversions, Local Adaptation and Speciation** *Genetics* **173**:419–434
- Kopp M, Hermisson J (2006) **The evolution of genetic architecture under frequency-dependent disruptive selection** *Evolution* **60**:1537–1550
- Koury SA (2023) **Predicting recombination suppression outside chromosomal inversions in *Drosophila melanogaster* using crossover interference theory** *Heredity* **130**:196–208
- Koressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M (2018) **Primer3_masker: integrating masking of template sequence with primer design software** *Bioinforma. Oxf. Engl* **34**:1937–1938
- Koressaar T, Remm M (2007) **Enhancements and modifications of primer design program Primer3** *Bioinforma. Oxf. Engl* **23**:1289–1291
- Krimbas CB, Powell JR, Krimbas CB, Powell JR (1992) **Introduction** *Drosophila Inversion Polymorphism* :1–52
- Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR (2014) **. doublesex is a mimicry supergene** *Nature* **507**:229–232
- Küpper C *et al.* (2016) **A supergene determines highly divergent male reproductive morphs in the ruff** *Nat. Genet* **48**:79–83
- Lack JB, Lange JD, Tang AD, Corbett-Detig RB, Pool JE (2016) **A Thousand Fly Genomes: An Expanded *Drosophila* Genome Nexus** *Mol. Biol. Evol* **33**:3308–3313
- Lack JB, Yassin A, Sprengelmeyer QD, Johanning EJ, David JR, Pool JE (2016) **Life history evolution and cellular mechanisms associated with increased size in high-altitude *Drosophila*** *Ecol. Evol* **6**:5893–5906
- Lange JD, Bastide H, Lack JB, Pool JE (2022) **A Population Genomic Assessment of Three Decades of Evolution in a Natural *Drosophila* Population** *Mol. Biol. Evol* **39**
- Langley CH *et al.* (2012) **Genomic Variation in Natural Populations of *Drosophila melanogaster*** *Genetics* **192**:533–598
- Lemeunier F, Aulard S, Krimbas CB, Powell JR (1992) **Inversion polymorphism in *Drosophila melanogaster*** *Drosophila Inversion Polymorphism* :339–405
- Li H (2013) **Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM** *Available from*

- Libbrecht R, Kronauer DJC (2014) **Convergent Evolution: The Genetics of Queen Number in Ants** *Curr. Biol* **24**:R1083–R1085
- Lupiáñez DG *et al.* (2015) **Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions** *Cell* **161**:1012–1025
- Machado HE *et al.* (2021) **Broad geographic sampling reveals the shared basis and environmental correlates of seasonal adaptation in *Drosophila*** *eLife* **10**
- McBroome J, Liang D, Corbett-Detig RB (2020) **Fine-Scale Position Effects Shape the Distribution of Inversion Breakpoints in *Drosophila melanogaster*** *Genome Biol. Evol* **12**:1378–1391
- Mérot C, Llaurens V, Normandeau E, Bernatchez L, Wellenreuther M (2020) **Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly** *Nat. Commun* **11**
- Mettler E, Voelker RA, Mukai T (1977) **Inversion Clines in Populations of *Drosophila melanogaster*** *Genetics* **87**
- Morgan TH (1910) **Sex Limited Inheritance in *Drosophila*** *Science* **32**:120–122
- Mroczek RJ, Melo JR, Luce AC, Hiatt EN, Dawe RK (2006) **The Maize Ab10 Meiotic Drive System Maps to Supernumerary Sequences in a Large Complex Haplotype** *Genetics* **174**:145–154
- Mukai T, Yamaguchi O (1974) **The genetic structure of natural populations of *Drosophila melanogaster* XI. Genetic variability in a local population.** *Genetics* **76**:339–366
- Munasinghe M, Brandvain Y (2023) **Together Inbreeding and Reproductive Compensation Favor Lethal t-Haplotypes**
- Nassar R, Muhs HJ, Cook RD (1973) **Frequency-dependent selection at the Payne inversion in *Drosophila melanogaster*** *Evolution* **27**:558–564
- O'Donald P, Majerus MEN, Clarke BC, Partridge L, Robertson A, Clarke BC, Partridge L (1997) **Frequency-dependent sexual selection** *Philos. Trans. R. Soc. Lond. B Biol. Sci* **319**:571–586
- Ohta T, Kimura M (1970) **Development of associative overdominance through linkage disequilibrium in finite populations** *Genet. Res* **16**:165–177
- Otto SP, Lenormand T (2002) **Resolving the paradox of sex and recombination: Evolution of sex** *Nat. Rev. Genet* **3**:252–261
- Otwinowski J, McCandlish DM, Plotkin JB (2018) **Inferring the shape of global epistasis** *Proc. Natl. Acad. Sci* **115**:E7550–E7558
- Pearse DE *et al.* (2019) **Sex-dependent dominance maintains migration supergene in rainbow trout** *N. Ecol. Evol* **3**:1731–1742
- Pei Y, Forstmeier W, Knief U, Kempenaers B (2023) **Weak antagonistic fitness effects can maintain an inversion polymorphism** *Mol. Ecol* **32**:3575–3585
- Pennell MW, Kirkpatrick M, Otto SP, Vamosi JC, Peichel CL, Valenzuela N, Kitano J (2015) **Y Fuse? Sex Chromosome Fusions in Fishes and Reptiles** *PLOS Genet* **11**

- Pool JE (2015) **The Mosaic Ancestry of the *Drosophila* Genetic Reference Panel and the *D. melanogaster* Reference Genome Reveals a Network of Epistatic Fitness Interactions** *Mol Biol Evol* **32**:3236–3251
- Pool JE, Braun DT, Lack JB (2017) **Parallel Evolution of Cold Tolerance within *Drosophila melanogaster*** *Mol. Biol. Evol* **34**:349–360
- Poul YL, Whibley A, Chouteau M, Prunier F, Llaurens V, Joron M (2014) **Evolution of dominance mechanisms at a butterfly mimicry supergene** *Nat. Commun* **5**
- Ragland GJ, Armbruster PA, Meuti ME (2019) **Evolutionary and functional genetics of insect diapause: a call for greater integration** *Curr Opin Insect Sci* **36**:74–81
- Reinhardt JA, Kolaczowski B, Jones CD, Begun DJ, Kern AD (2014) **Parallel Geographic Variation in *Drosophila melanogaster*** *Genetics* **197**:361–373
- Reinhart M, Carney T, Clark AG, Fiumera AC (2015) **Characterizing Male–Female Interactions Using Natural Genetic Variation in *Drosophila melanogaster*** *J. Hered* **106**:67–79
- Rest JS, Morales CM, Waldron JB, Opulente DA, Fisher J, Moon S, Bullaughey K, Carey LB, Dedousis D (2013) **Nonlinear Fitness Consequences of Variation in Expression Level of a Eukaryotic Gene** *Mol. Biol. Evol* **30**:448–456
- Rice WR (1992) **Sexually antagonistic genes: experimental evidence** *Science* **256**:1436–1439
- Rice WR (1996) **Sexually antagonistic male adaptation triggered by experimental arrest of female evolution** *Nat. Lond* **381**:232–234
- Rose MR (1982) **Antagonistic pleiotropy, dominance, and genetic variation** *Heredity* **48**:63–78
- Rowe L, Chenoweth SF, Agrawal AF (2018) **The Genomics of Sexual Conflict** *Am. Nat* **192**:274–286
- Ruzicka F, Hill MS, Pennell TM, Flis I, Ingleby FC, Mott R, Fowler K, Morrow EH, Reuter M (2019) **Genome-wide sexually antagonistic variants reveal long-standing constraints on sexual dimorphism in fruit flies** *PLOS Biol* **17**
- Said I, Byrne A, Serrano V, Cardeno C, Vollmers C, Corbett-Detig RB (2018) **Linked genetic variation and not genome structure causes widespread differential expression associated with chromosomal inversions** *Proc. Natl. Acad. Sci* **115**:5492–5497
- Saunders DS, Richard DS, Applebaum SW, Ma M, Gilbert LI (1990) **Photoperiodic diapause in *Drosophila melanogaster* involves a block to the juvenile hormone regulation of ovarian maturation** *Gen Comp Endocrinol* **79**:174–184
- Saunders PA, Neuenschwander S, Perrin N (2019) **Impact of deleterious mutations, sexually antagonistic selection, and mode of recombination suppression on transitions between male and female heterogamety** *Heredity* **123**:419–428
- Schaeffer SW, Goetting-Minesky MP, Kovacevic M, Peoples JR, Graybill JL, Miller JM, Kim K, Nelson JG, Anderson WW (2003) **Evolutionary genomics of inversions in *Drosophila pseudoobscura*: Evidence for epistasis** *Proc. Natl. Acad. Sci* **100**:8319–8324

- Sprengelmeyer QD, Mansourian S, Lange JD, Matute DR, Cooper BS, Jirle EV, Stensmyr MC, Pool JE (2020) **Recurrent Collection of *Drosophila melanogaster* from Wild African Environments and Genomic Insights into Species History** *Mol. Biol. Evol* **37**:627–638
- Stearns FW (2010) **One Hundred Years of Pleiotropy: A Retrospective** *Genetics* **186**:767–773
- Sturtevant AH (1921) **A Case of Rearrangement of Genes in *Drosophila*** *Proc. Natl. Acad. Sci* **7**:235–237
- Sturtevant AH, Beadle GW (1936) **The Relations of Inversions in the X Chromosome of *Drosophila melanogaster* to Crossing Over and Disjunction** *Genetics* **21**
- Sturtevant AH, Mather K (1938) **The Interrelations of Inversions, Heterosis and Recombination** *Am. Nat* **72**:447–452
- Surmacki A, Ożarowska-Nowicka A, Rosin ZM (2013) **Color polymorphism in a land snail *Cepaea nemoralis* (Pulmonata: Helicidae) as viewed by potential avian predators** *Naturwissenschaften* **100**:533–540
- Sved JA (1968) **The Stability of Linked Systems of Loci with a Small Population Size** *Genetics* **59**:543–563
- Tatarenkov A, Healey CIM, Grether GF, Avise JC (2008) **Pronounced reproductive skew in a natural population of green swordtails, *Xiphophorus helleri*** *Mol. Ecol* **17**:4522–4534
- Thain D, Tannenbaum T, Livny M (2005) **Distributed computing in practice: the Condor experience** *Concurr. Comput. Pract. Exp* **17**:323–356
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) **Primer3—new capabilities and interfaces** *Nucleic Acids Res* **40**
- Van der Auwera GA, O'Connor BD (2020) **Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly Media Incorporated**
- Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang Y-C, Shoemaker D, Keller L (2013) **A Y-like social chromosome causes alternative colony organization in fire ants** *Nature* **493**:664–668
- Wellenreuther M, Bernatchez L (2018) **Eco-Evolutionary Genomics of Chromosomal Inversions** *Trends Ecol. Evol* **33**:427–440
- White MJD (1973) **Animal Cytology and Evolution**
- Wittmann MJ, Bergland AO, Feldman MW, Schmidt PS, Petrov DA (2017) **Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift** *Proc. Natl. Acad. Sci* **114**:E9932–E9941
- Wright AE, Dean R, Zimmer F, Mank JE (2016) **How to make a sex chromosome** *Nat. Commun* **7**:1–8
- Yamamoto A, Huang W, Carbone MA, Anholt RRH, Mackay TFC (2024) **The Genetic Basis of Incipient Sexual Isolation in *Drosophila melanogaster*** *bioRxiv*

Yassin A, Bastide H, Chung H, Veuille M, David JR, Pool JE (2016) **Ancient balancing selection at tan underlies female colour dimorphism in *Drosophila erecta*** *Nat. Commun* **7**

Zajitschek F, Connallon T (2018) **Antagonistic pleiotropy in species with separate sexes, and the maintenance of genetic variation in life-history traits and fitness** *Evolution* **72**:1306–1316

Zhao L, Charlesworth B (2016) **Resolving the Conflict Between Associative Overdominance and Background Selection** *Genetics* **203**:1315–1334

Editors

Reviewing Editor

Dieter Ebert

University of Basel, Basel, Switzerland

Senior Editor

Detlef Weigel

Max Planck Institute for Biology Tübingen, Tübingen, Germany

Reviewer #1 (Public Review):

The hypothesis is based on the idea that inversions capture genetic variants that have antagonistic effects on male sexual success (via some display traits) and survival of females (or both sexes) until reproduction. Furthermore, a sufficiently skewed distribution of male sexual success will tend to generate synergistic epistasis for male fitness even if the individual loci contribute to sexually selected traits in an additive way. This should favor inversions that keep these male-beneficial alleles at different loci together at a cis-LD. A series of simulations are presented and show that the scenario works at least under some conditions. While a polymorphism at a single locus with large antagonistic effects can be maintained for a certain range of parameters, a second such variant with somewhat smaller effects tends to be lost unless closely linked. It becomes much more likely for genomically distant variants that add to the antagonism to spread if they get trapped in an inversion; the model predicts this should drive accumulation of sexually antagonistic variants on the inversion versus standard haplotype, leading to the evolution of haplotypes with very strong cumulative antagonistic pleiotropic effects. This idea has some analogies with one of predominant hypotheses for the evolution of sex chromosomes, and the authors discuss these similarities. The model is quite specific, but the basic idea is intuitive and thus should be robust to the details of the model assumption. It makes perfect sense in the context of the geographic pattern of inversion frequencies.

To provide empirical support for this idea, the authors study the dynamics of inversions in population cages over one generation, tracking their frequencies through amplicon sequencing at three time points: (young adults), embryos and very old adult offspring of either sex (>2 months from adult emergence). Out of four inversions included in the experiment, two show patterns consistent with antagonistic effects on male sexual success (competitive paternity) and the survival of offspring, especially females, until an old age, which the authors interpret as consistent with their theory.

There are several reasons why the support from these data for the proposed theory is not waterproof.

(1) As I have already pointed out in my previous review, survival until 2 months (in fact, it is 10 weeks and so 2.3 months) of age is of little direct relevance to fitness, whether under natural conditions or under typical lab conditions.

The authors argue this objection away with two arguments

First, citing Pool (2015) they claim that the average generation time (i.e. the average age at which flies reproduce) in nature is 24 days. That paper made an estimate of 14.7 generations per year under the North Carolina climate. As also stated in Pool (2015), the conditions in that locality for *Drosophila* reproduction and development are not suitable during three months of the year. This yields an average generation length of about 19.5 days during the 9 months during which the flies can reproduce. On the highly nutritional food used in the lab and at the optimal temperature of 25 C, *Drosophila* need about 11-12 days to develop from egg to adult. Even assuming these perfect conditions, the average age (counted from adult eclosion) would be about 8 days. In practice, larval development in nature is likely longer for nutritional and temperature reasons, and thus the genomic data analyzed by Pool imply that the average adult age of reproducing flies in nature would be about 5 days, and not 24 days, and even less 10 weeks. This corresponds neatly to the 2-6 days median life expectancy of *Drosophila* adults in the field based on capture-recapture (e.g., Rosewell and Shorrocks 1987). Second, the authors also claim that survival over a period of 2 month is highly relevant because flies have to survive long periods where reproduction is not possible. However, to survive the winter flies enter a reproductive diapause, which involves profound physiological changes that indeed allow them to survive for months, remaining mostly inactive, stress resistant and hidden from predators. Flies in the authors' experiment were not diapausing, given that they were given plentiful food and kept warm. It is still possible that survival to the ripe old age of 10 weeks under these conditions still correlates well with surviving diapause under harsh conditions, but if so, the authors should cite relevant data. Even then, I do not think this allows the authors to conclude that longevity is "the main selective pressure" on *Drosophila* (l. 936).

(2) It appears that the "parental" (in fact, paternal) inversion frequency was estimated by sequencing sires that survived until the end of the two-week mating period. No information is provided on male mortality during the mating period, but substantial mortality is likely given constant courtship and mating opportunities. If so, the difference between the parental and embryo inversion frequency could reflect the differential survival of males until the point of sampling rather than / in addition to sexual selection.

(3) Finally, irrespective of the above caveats, the experimental data only address one of the elements of the theoretical hypothesis, namely antagonistic effects of inversions on reproduction and survival, notably that of females. It does not test for two other key elements of the proposed theory: the assumption of frequency-dependence of selection on male sexual success, and the prediction of synergistic epistasis for male fitness among genetic variants in the inversion. To be fair, particularly testing the latter prediction would be exceedingly difficult. Nonetheless, these limitations of the experiment mean that the paper is much stronger theoretical than empirical contribution.

<https://doi.org/10.7554/eLife.93338.2.sa3>

Reviewer #2 (Public Review):

Summary:

In their manuscript the authors address the question of whether the inversion polymorphism in *D. melanogaster* can be explained by sexually antagonistic selection. They designed a new simulation tool to perform computer simulations, which confirmed their hypothesis. They

also show a tradeoff between male reproduction and survival. Furthermore, some inversions display sex-specific survival.

Strengths:

It is an interesting idea on how chromosomal inversions may be maintained

Weaknesses:

General points:

The manuscript lacks clarity of writing. It is impossible to fully grasp what the authors did in this study and how they reached their conclusions. Therefore, I cannot guarantee that I spotted all the problems in the study and in my review I will only highlight some cases that I found problematic.

Although this is an interesting idea, but it clearly cannot explain the apparent influence of seasonal and clinal variation on inversion frequencies.

Comments on the latest version:

I would like to give an example of the confusing terminology of the authors:

"Additionally, fitness conveyed by an allele favoring display quality is also frequency-dependent: since mating success depends on the display qualities of other males, the relative advantage of a display trait will be diminished as more males carry it..."

I do not understand the difference to an advantageous allele, as it increases in frequency the frequency increase of this allele decreases, but this has nothing to do with frequency dependent selection. In my opinion, the authors re-define frequency dependent selection, as for frequency dependent selection needs to change with frequency, but from their verbal description this is not clear.

One example of how challenging the style of the manuscript is comes from their description of the DNA extraction procedure. In principle a straightforward method, but even here the authors provide a convoluted uninformative description of the procedure.

It is not apparent to the reviewer why the authors have not invested more effort to make their manuscript digestible.

<https://doi.org/10.7554/eLife.93338.2.sa2>

Reviewer #3 (Public Review):

Summary:

In this study, McAllester and Pool develop a new simulation model to explain the maintenance of balanced inversion polymorphism, based on (sexually) antagonistic alleles and a trade-off between male reproduction and survival (in females or both sexes). In support of the plausibility of this model, the authors use laboratory experiments on four naturally occurring inversion polymorphisms in *Drosophila melanogaster*, finding evidence for the existence of the above-mentioned trade-off in two out of the four cases.

Strengths:

(1) The study develops and analyzes a new (*Drosophila melanogaster*-inspired) model for the maintenance of balanced inversion polymorphism, combining elements of (sexually) antagonistically (pleiotropic) alleles, negative frequency-dependent selection and synergistic epistasis. To this end, the authors develop and use a new simulator.

(2) The above-mentioned model assumes, as a specific example, a trade-off between male reproductive display and survival; in the second part of their study, the authors perform laboratory experiments on four common *D. melanogaster* inversions to study whether these

polymorphisms may be subject to such a trade-off. The authors find that two of the four inversions show suggestive evidence that is consistent with a trade-off between male reproduction and survival. The new amplicon sequencing approach to track inversion frequencies used by the authors seems promising in terms of studying fitness effects / trade-offs associated with polymorphic inversions and how such effects play out dynamically.

Weaknesses:

A gap in the current modeling is that, while a diploid situation is being studied, the model does not investigate the effects of varying degrees of dominance. It would be important and interesting to fill this gap in future work.

Comments on the latest version:

Most of the comments which I have made in my public review have been adequately addressed.

Some of the writing still seems somewhat verbose and perhaps not yet maximally succinct; some additional line-by-line polishing might still be helpful at this stage in terms of further improving clarity and flow (for the authors to consider and decide).

<https://doi.org/10.7554/eLife.93338.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Reviewer #1 (Public Review):

[...]

(1) The authors claim that the negative frequency dependence that maintains polymorphism in their model results from a non-linear relationship between the display trait and sexual success [...] Maybe I missed something, but the authors do not provide support for their claim about the negative frequency-dependence of sexual selection in their simulations. To do so they could (1) extract the relationship between the relative mating success of the two male types from the simulations and (2) demonstrate that polymorphism is not maintained if the relationship between male display trait and mating success is linear.

We believe that there is a confusion of terminology here. We agree that for the two alleles at a locus impacting male display in our model, the allele conferring inferior display quality will have a fitness that increases as its frequency increases, so this allele displays positive frequency dependent fitness. And, the alternate, display-favoring allele at the locus does display negative frequency dependence. Our use of the terminology 'negative frequency dependence' was meant to refer to the negative dependence of the fitness of the display-favoring allele with respect to its own frequency. However, a significant body of literature instead discusses models in which both an allele and its alternate(s) are beneficial when at low frequency and deleterious when at high frequency under the same selective challenge, entailing negative frequency dependence of fitness for all alleles involved. This benefit-when-rare model of a single trait is often described simply as negative frequency dependence, and generates balancing selection at the locus, but is not the model we are presenting here, and does not encompass all models involving negative frequency dependent fitness. This lexical expectation may make the interpretation of our work more difficult, and we have amended the manuscript to make our model clearer (lines 227-231). In this model, we have a negative frequency dependence for the fitness of the display-favoring allele in mate competition, but

the net selective disadvantage of this allele at high frequency is due to a cost in another, pleiotropic, fitness challenge: the constant survival effect. So, the alleles are under balancing selection where alternate alleles are favored by selection when rare, but not due solely to selection during mate competition. Instead, our model relies on pleiotropy for an emergent form of frequency-dependent balancing selection (in the sense that each allele is predicted to be beneficial on balance when rare).

In the reviewer's model of the success of two alleles at one locus, the ratio of success is vaguely linear with allele frequency for $n=3$, though it starts quite convex and has an inflection point between convex and concave segments (for the disfavored allele) at $p=0.532$. This is visualized easily by plotting the function and its derivatives in Wolfram-Alpha. For $n \geq 4$, the fitness function with respect to the display-favoring/disfavoring allele becomes increasingly concave/convex respectively, and this specific nonlinearity is needed to act along with the antagonistic pleiotropy to maintain balancing selection, rather than being maintained by a model that favors any rare allele on the basis of its rarity in some manner. In an attempt to make the importance of the encounter number parameter clearer, we've generated new panels for Figure S1 which simulate encounter numbers 2, 3, and 4, and we have updated corresponding text and figure references in lines 335-338.

For (1-2), it is not clear how to modify the simulation such that the relationship between the trait value and mating success can be perfectly linear - either linear with respect to allele frequency in a one locus model or linear with respect to trait value at a specific population composition, without removing the simulation of mate competition altogether. While it may be of interest to explore a more comprehensive range of biological trade-offs in future studies, we are not able to meaningfully do so within the context of the present manuscript.

(2) The authors only explore versions of the model where the survival costs are paid by females or by both sexes. We do not know if polymorphism would be maintained or not if the survival cost only affected males, and thus if sexual antagonism is crucial.

We now present simulations with male costs only as added panels to Figure S1 and mention these results in the main text (lines 334-335). Maintenance of the polymorphism is significantly reduced or completely absent in such simulations.

(3) The authors assume no cost to aneuploidy, with no justification. Biologically, investment in aneuploid eggs would not be recoverable by Drosophila females and thus would potentially act against inversions when they are rare.

We did offer some discussion and justification of our decision to model no inherent fitness of the inversion mutation itself, specifically aneuploidy, in lines 36-39 and 78-80 of the original reviewed preprint. Previous research suggests that *D. melanogaster* females may not actually invest in aneuploid eggs generated from crossover within paracentric inversions. While surprising, and potentially limited to a subset of clades, many 'r-selected' taxa or those in which maternal investment is spread out over time may have some degree of reproductive compensation for non-viable offspring, which can reduce the costs of generating aneuploids significantly (for example, t-haplotypes in mice). We have added this example and citation to lines 34ff in the current draft.

(4) The authors appear to define balanced polymorphism as a situation in which the average allele frequency from multiple simulation runs is intermediate between zero and one (e.g., Figure 3). However, a situation where 50% of simulation runs end up with the fixation of allele A and the rest with the fixation of allele B (average frequency of 0.5) is not a balanced polymorphism. The conditions for balanced polymorphism require that selection favors either variant when it is rare.

We originally chose mean final frequency for presenting the single locus simulations based on the ease of generating a visual plot that included information on fixation vs loss and equilibrium frequency. Figure 3 and related supplemental images have been changed to now also represent the proportion of simulations retaining polymorphism at the locus in the final generation.

(5) Possibly the most striking result of the experiment is the fact that for 14 out of 16 combinations of inversion x maternal background, the changes in allele frequencies between embryo and adult appear greater in magnitude in females than in males irrespective of the direction of change, being the same in the remaining two combinations. The authors interpret this as consistent with sexually antagonistic pleiotropy in the case of In(3L)Ok and In(3R)K. The frequencies of adult inversion frequencies were, however, measured at the age of 2 months, at which point 80% of flies had died. For all we know, this may have been 90% of females and 70% of males that died at this point. If so, it might well be that the effects of inversion on longevity do not systematically differ between the ages and the difference in Figure 9B results from the fact that the sample includes 30% longest-lived males and 10% longest-lived females.

This critique deserves some consideration. The aging adults were separated by sex during aging, but while we recorded the number of survivors, we did not record the numbers of eclosed adults and their sexes initially collected out of an interest in maintaining high throughput collection. We therefore cannot directly calculate the associated survival proportions, but we can estimate them. We collected 1960 females and 3156 males, and we can very roughly estimate survival if we assume that equal numbers of each sex eclosed, and that the survivors represent 20% of the original population. That gives 12790 individuals per sex, or 84.7% female mortality and 75.3% male mortality.

So, we have added a qualification discussing the possibility of stronger selection on females and its influence on observed sex-specific frequency changes, on lines 602-605.

(6) Irrespective of the above problem, survival until the age of 2 months is arguably irrelevant from the viewpoint of fitness consequences and thus maintenance of inversion polymorphism in nature. It would seem that trade-offs in egg-to-adult survival (as assumed in the model), female fecundity, and possibly traits such as females resistance to male harm would be much more relevant to the maintenance of inversion polymorphisms.

Adult *Drosophila* will continue to reproduce in good conditions until mortality, and the estimated age of a mean reproductive event for a *Drosophila melanogaster* individual is 24 days (Pool 2015), and likewise for *D. simulans* (Turelli and Hoffman 1995). Given that reproduction is centered around 24 days, we expect sampling at 2 months of age to still be relevant to fitness. In seasonally varying climates, either temperate or with long dry season, survival through challenging conditions is expected to require several months. In many such cases, females are in reproductive diapause, and so longevity is the main selective pressure. See lines 931-936 in the revised manuscript.

As we agreed above, it would of interest to investigate a wider range of trade-offs in future studies. We focused here on the balanced between survival and male reproductive success because the latter trait generates negative frequency dependence for display-favoring alleles and a disproportionate skew towards higher quality competitors, whereas many other fitness-relevant traits lack that property.

(7) The experiment is rather minimalistic in size, with four cages in total; given that each cage contains a different female strain, it essentially means N=1. The lack of replication

makes statements like "In(2L)t and In(2R)NS each showed elevated survival with all maternal strains except ZI418N" (l. 493) unsubstantiated because the claimed special effect of ZI418N is based on a single cage subject to genetic drift and sampling error. The same applies to statements on inversion x female background interaction (e.g., l. 550), as this is inseparable from residual variation. It is fortunate that the most interesting effects appear largely consistent across the cages/female backgrounds. Still, I am wondering why more replicates had not been included.

Our experimental approach might be described as “diversity replication”. Essentially, the four maternal genetic backgrounds are serving dual purposes – both to assess experimental consistency and to ensure that our conclusions are not solely driven by a single non-representative genotype (which in so many published studies, can not be ruled out). It would indeed be interesting if we could have quadrupled the size of our experiment by having four replicates per maternal background. However, we suspect the reviewer may not recognize the substantial effort involved in our four existing experiments. Each of these involved collecting 500+ virgin females, hand-picking thousands of embryos during the duration of egg-laying, and repeatedly transferring offspring to maintain conditions during aging, such that cages had to be staggered by more than a month. These four cages took a year of benchwork just to collect frozen samples, before any preparation and quality control of the associated amplicon libraries for sequencing. Adding a further multiplier would take it well beyond the scope of a single PhD thesis. Fortunately, we were able to obtain the key results of interest without that additional effort, even if clearer insights into the role of maternal background would also be of strong interest.

We do agree that no firm conclusions about maternal background can be reached without further replication, and so we have qualified or removed relevant statements accordingly (lines 568ff, 620-622).

Reviewer #1 (Recommendations For The Authors):

The description of the model is confusing and incomplete, e.g., the values of several parameters used to obtain the numerical results are not given. It is first stated (l. 223) that the model is haploid, but text elsewhere talks about homozygotes and heterozygotes. If the model is diploid (this in itself is not clear), what is assumed about dominance?

We are not presenting results for a mathematical model estimated numerically. We have now clarified our transition from a conceptual depiction of our model, in which we use haploid representations for simplified presentation, to our forward population genetic simulations, which are entirely diploid. More broadly, we have improved our communication of the assumptions and parameters used in our simulations. The scenarios we investigate involve purely additive trait effects within and between loci (except that survival probabilities are multiplicative to avoid negative values). We think that considering other dominance scenarios would be a worthy subject for a follow-up study, whereas the present manuscript is already covering a great deal of ground.

Similarly, it is hard to understand the design (l.442ff). I was confused as to whether a population was set up for each inversion or for all of them and what the unit or replication was. I found the description in Methods (l. 763-771) much clearer and only slightly longer; I suggest the authors transfer it to the Results. Also, Figure 8 should contain the entire crossing scheme; the current version is misleading in that it implies males with only two genotypes.

All four tested inversions were segregating within the same karyotypically diverse population of males, and were assayed from the same experiments. We have attempted to

improve the relevant description. For Figure 8, we had trouble conceiving a graphic update that contained a more complete cross scheme without seeming much more confused and cluttered. We have tried to clarify in the relevant text and the figure caption instead.

There are a number of small issues that should be addressed:

- No epistasis for viability assumed - what would be the consequence?

We explored a model in which we intentionally included no terms for epistatic effects on phenotype. All epistasis with regard to fitness is emergent from competition between individuals with phenotypes composed of non-epistatic, non-dominant genetic effects. So, the simplest model of antagonism would have no epistasis for viability whatsoever. One could explore a model that has emergent viability epistasis in a similar way, by implementing stabilizing selection on a quantitative trait with a gaussian or similar non-linear phenotype-to-fitness map, but that might be better served as a topic for a future study. We have, however, tried to make this intent clearer in the text.

I. 750 implies that aneuploidy generated by the inversion has no cost (aneuploid games are resampled)

Yes, as addressed in public review item (3). Alternately see lines 34ff, 293, 369, 392 for in-text edits.

I. 24-25: unclear; is this to mean that there is haplotype x sex interaction for survival?

I. 25: success in what? (I assume this will be explained in the paper, but the abstract should stand on its own).

I. 193-4: "producing among most competitive males": something missing or a word too much?? Figure 1B,C: a tiny detail, but the plots would be more intuitive if the blue (average) bars were aged (i.e., to the right) of the male and female ones, given that the average is derived from the two sex-specific values.

Each of the above have been edited or implemented as suggested

I. 205. It is convex function, but I do not understand what the authors mean by "convex distribution".

Hopefully the updated text is clearer: "yielding a distribution of male reproductive output that follows a relatively convex trend".

I. 223ff: some references to Fig 1 panels in this paragraph seem off by one letter (i.e., A should be B, etc.).

I. 231 "fitness...are equally fit": rephrase

I. 260: maybe "thrown out" is not the most fortunate term, maybe "eliminated" would be better?

Each of the above have been edited or implemented as suggested

Figure 3: I do not understand the meaning of "additive" and "multiplicative" in the case of a single locus haploid model

All presented simulations are diploid, and these refer to the interactions between the two alleles at the locus. Hopefully the language is overall clearer in this draft.

l. 274: "Mutation of new nucleotide" meaning what? Or is it mutation_to_ a new nucleotide?

Hopefully the revised text is clearer.

Figure 5. The right panel of figure 5A implies that, with the inversion, the population evolves to an extreme display trait that is so costly that it fills 95% of all individuals (or of all females?

What is assumed about this here?). Apart from the biological realism of this result, what does it say about the accumulation of polymorphism and maintenance of the inversion? The graphs in fig 5B do plot a divergence between haplotypes, but it is not clear how they relate to those in panel A - the parameter values used to generate these plots are again not listed. Furthermore, from the viewpoint of the polymorphism, it would be good to report the frequencies at the steady-state.

We have now clarified the figure description, including the parameter values used. The distribution of frequencies at the end of the simulation is represented in figure 6. Given that we set up the simulation with assumptions that are otherwise common to population models, what biological process would prevent this extreme? Why isn't this extreme observed in natural populations? One possible explanation is that they become sex chromosomes, with increasing likelihood as the cost increases. Or other compensatory changes may occur that we don't simulate, like regulatory evolution giving a complementary phenotype. Maybe genetic constraints in natural populations prevent the mutation of the kind of pleiotropic mutations that drive this dynamic. The populations still survive, though they are parameterized by relative fitness. What would an absolute fitness population function be? Would it go extinct or not? It would be of interest to explore a wider range of models, but it is the purpose of this paper to establish that this is a viable model for the maintenance of sexually antagonistic polymorphism and association with inversions. We have added a paragraph motivated by this comment to the Discussion starting on line 765.

l. 401-2: Z-like, W-like : please specify you are talking about patterns resembling sex chromosomes.

l. 738: "population calculates"?

l. 743-4 and 746-7: is this the same thing said twice, or are there two components of noise? l. 357: there is no figure 5C.

Each of the above have been addressed with text edits.

L. 473-5: Yes, the offspring did not contain inversion homozygotes, but the sire pool did, didn't it? So homozygous inversions may have affected male reproductive success. Anyway, most of this paragraph (from line 473) seems to belong in Discussion rather than Results.

We have revised this sentence to focus on offspring survival.

We can understand the reviewer's suggestion about Results vs. Discussion text. While this can often be a challenging balance, we find that papers are often clearer if some initial interpretation is offered within the Results text. However, we moved the portion of this paragraph relating our findings to the published literature to the Discussion.

I. 516: "In(3L)Ok favored male survival": this is misleading/confusing given the data, "In(3L)Ok reduced female survival more strongly than male survival..."

Hopefully the phrasing is clearer now.

I. 663ff: I did not have an impression that this section added anything new and could safely be cut.

We have done some editing to make this more concise and emphasize what we think is essential, but we believe that the model of an autosomal, sexually antagonistic inversion differentiating before contributing to the origin of a sex chromosome is novel and interesting. And, that this additional emphasis is worthwhile to encourage thought and consideration of this idea in future research and among interested researchers.

I. 751: "flat probability per locus": do the authors mean a constant probability?

Edited.

Reviewer #2 (Public Review):

The manuscript lacks clarity of writing. It is impossible to fully grasp what the authors did in this study and how they reached their conclusions. Therefore, I will highlight some cases that I found problematic.

Hopefully the revised manuscript improves writing clarity.

Although this is an interesting idea, it clearly cannot explain the apparent influence of seasonal and clinal variation on inversion frequencies.

We do not believe that our model predicts a non-existence of temporal and spatial dependence of the fitness of inverted haplotypes, nor do we seek to identify the manner in which seasonal and clinal differences affect fitness of inverted haplotypes. Rather, we argued that the influence of seasonal and clinal selection on inversions does not on its own predict the observed maintenance of inversions at low to intermediate frequencies across such a diverse geographic range, along with the higher frequencies of many derived inversions in more ancestral environments.

We might imagine that trade-offs between life history traits such as mate competition and survival should be universal across the range of an organism. But in practice, the fitness benefits and costs of a pleiotropic variant (or haplotype) may be heavily dependent on the environment. A harsh environment such as a temperate winter may both reduce the number of females that a male encounters (decreasing the benefit of display-enhancing variants) and also increase the likelihood that survival-costly variants lead to mortality (thus increasing their survival penalty). In light of such dynamics, our model would predict that equilibrium inversion frequencies should be spatially and temporally variable, in agreement with a number of empirical observations regarding *D. melanogaster* inversions.

We have edited the introduction to emphasize that inversion frequencies vary temporally as well as seasonally, on lines 144ff. We also note relevant discussion of the potential interplay between the environment and trade-offs such as those we investigate, on lines 153-155.

The simulations are highly specific and make very strong assumptions, which are not well-justified.

We respond to all specific concerns expressed in the Recommendations For The Authors section below. We also note that we have made further clarifications throughout the text regarding the assumptions made in our analysis and their justification.

Reviewer #2 (Recommendations For The Authors):

I think that the manuscript would greatly benefit from a major rewrite and probably also a reanalysis of the empirical data.

In particular, a genome-wide analysis of differences in SNP frequencies between sexes and developmental stages would help the reader to appreciate that inversions are special.

[moved up within this section for clarity] We are lacking a genomic null model-how often do the authors see similar allele frequency differences when looking at the entire genome? This could be easily done with whole genome Pool-Seq and would tell us whether inversions are really different from the genomic background. I think that this information would be essential given the many uncertainties about the statistical tests performed.

We expect that autosome-wide SNP frequencies will be heavily influenced by the frequencies of inversions, which occur on all four major autosomal chromosome arms. These inversions often show moderate disequilibrium with distant variants (e.g. Corbett-Detig & Hartl 2012).

Furthermore, the limited number of haplotypes present, given that the paternal population was founded from 10 inbred lines, would further enhance associations between inversions and distant variants. Therefore, we do not expect that whole-genome Pool-Seq data would provide an appropriate empirical null distribution for frequency changes. Instead, we have generated appropriate null predictions by accounting for both sampling effects and experimental variance, and we have aimed to make this methodology clearer in the current draft.

Some basic questions:

why start at a frequency of 50% (line 287)?

Isn't it obvious that in this scenario strong alleles with sexually antagonistic effects can survive?

The initial goal of the associated Figure 4 was not to show that a strongly antagonistic variant could persist. Instead, we wanted to test the linkage conditions in which a second, relatively weaker antagonistic variant survived – which did not occur in the absence of strong linkage.

We have now added simulations with relatively lower initial frequencies, in which the weaker variant and the inversion both start at 0.05 frequency, while the stronger variant is still initialized at 0.5 to reflect the initial presence of one balanced locus with a strongly antagonistic variant. Here, the weaker antagonistic variant is still usually maintained when it is close to the stronger variant, and while the inversion-mediated maintenance of the weaker variant at greater distance from the stronger variant because less frequent than the original investigated case, it still happens often enough to hypothetically allow for such outcomes over evolutionary time-scales.

Still, we should also emphasize that the goals of this proof-of-concept analysis are to establish and convey some basic elements of our model. Subsequently, analyses such as those presented in Figures 5 and 6 provide clearer evidence that the hypothesized dynamics of

inversions facilitating the accumulation of sexual antagonism actually occur in our simulations.

The experiments seem to be conducted in replicate (which is of course essential), but I could not find a clear statement of how many replicates were done for each maternal line cross.

How did the authors arrive at 16 binomial trials (line 473)? 4 inversions, 4 maternal genotypes?

How were replicates dealt with?

In Figure 9, it would be important to visualize the variation among replicates.

Unfortunately, we did not have the bandwidth to perform replicates of each maternal line. Instead, we use four maternal backgrounds to simultaneously establish consistency across independent experiments and genetic backgrounds (see our response to Reviewer 1, point 7). We've edited the draft to make this clearer and more clearly delineate what is supported and not supported by our data. Replicate variation for the control replicates of the extraction and sequencing process, and the exact read counts of the experiment, are available in Supplemental Tables S5, S6, and S7.

The statistical analysis of trade-off is not clear: which null model was tested? No frequency change? In my opinion, two significances are needed: a significant difference between parental and embryo and then embryo and adult offspring. The issue with this is, however, that the embryo data are used twice and an error in estimating the frequency of the embryos could be easily mistaken as antagonistic selection.

Hopefully the description of our null model is clearer in the text, now starting around line 967 in the Methods. We are aware of the positive dependence when performing tests comparing the paternal to embryo and then embryo to offspring frequencies, and this is accounted for by our analysis strategy - see lines 1009-1012.

It was not clear how the authors adjusted their chi-squared test expectations. Were they reinventing the wheel? There is an improved version of the chi-squared test, which accounts for sampling variation.

We did not actually perform chi-square tests. Instead, we used the chi statistic from the chi-squared test as a quantitative summary of the differences in read counts between samples. We compared an observed value of chi to values for this statistic obtained from simulated replicates of the experiment. Sampling from this simulation generated our 'expected' distribution of read counts, sampled to match sources of variance introduced in the experimental procedure, but without any effect of natural selection, per lines 825ff in the original submission. Hence, we are approximating the likelihood of observing an empirical chi statistic by generating random draws from a model of the experiment and comparing values calculated from each draw to the experimental value: a Monte Carlo method of approximating a p-value for our data. We have attempted to make the structure of these simulations and their use as a null-model clearer in this draft.

It is not sufficiently motivated why the authors model differences in the extraction procedure with a binomial distribution.

Adding a source of variance here seemed necessary as running control sequencing replicates revealed that there was residual variance not fully recapitulated by sample-size-dependent resampling. Given that we were still sampling a number of draws from a binomial outcome

(the read being from the inverted or standard arrangement), a binomial distribution seemed a reasonable model, and we fit the level of this additional noise source to an experiment-wide constant, read-count or genome-count independent parameter that best fit the variance observed in the controls (lines 830ff in the original draft). Clarification is made in this manuscript draft, lines 979-989.

How many reads were obtained from each amplicon? It looks like the authors tried to mimic differences between technical replicates by a binomial distribution, which matches the noise for a given sample size, but this depends on the sequence coverage of the technical replicates.

We provide read counts in Supplemental Tables S6 and S7. The relevant paragraph in the methods has been edited for clarity, lines 972ff. Accounting for sampling differences between replicates used a hypergeometric distribution for paternal samples to account for paternal mortality before collection, and the rest were resampled with a binomial distribution. There were two additional binomial samplings, to account for resampling the read counts and to capture further residual variance in the library prep that did not seem to depend on either allele or read counts.

It would be good to see an estimate for the strength of selection: 10% difference in a single generation appears rather high to me.

Estimates of selection strength based on solving for a Wright-Fisher selection coefficient for each tested comparison can now be found in Table S8, mentioned in text on lines 589-590. The mean magnitude of selection coefficients for all paternal to embryo comparisons was 0.322, and for embryo to all adult offspring it was 0.648. For In(3L)Ok the mean selection coefficients were 0.479 and -0.53, and for In(3R)K they were -0.189 and 1.28, respectively. Some are of quite large magnitude, but we emphasize that the coefficients for embryo to adult are based on survival to old age, rather than developmental viability. That factor, in addition to the laboratory environment, makes these estimates distinct from selection coefficients that might be experienced in natural populations.

Reviewer #3 (Public Review):

Strengths:

(1) ...the authors developed and used a new simulator (although it was not 100% clear as to why SLiM could not have been used as SLiM has been used to study inversions).

Before SLiM 3.7 or so (and including when we did the bulk of our simulation work), we do not think it would have been feasible to use SLiM to model the mutation of inversions with random breakpoints and recombination between without altering the SLiM internals. Separately, needing to script custom selection, mutation, and recombination functions in Eidos would have slowed SLiM down significantly. Given our greater familiarity with python and numpy, and the ability to implement a similar efficiency simulator more quickly than through learning C++ and Eidos, we chose to write our own.

It should be a fair bit easier to implement comparable simulations in SLiM now, but it will still require scripting custom mutation, selection, and recombination functions and would still result in a similarly slow runtime. The current script recipe recommended by SLiM for simulating inversions uses constants to specify the breakpoints of a single inversion, without the ability to draw multiple inversions from a mutational distribution, or model recombination between more complicated karyotypes. Hence, our simulator still seems to be a more versatile and functional option for the purposes of this study.

Weaknesses:

[Comments 1 through 4 on Weaknesses included numerous citation suggestions, and some discussion recommendations as well. In our revised manuscript, we have substantially implemented these suggestions. In particular, we have deepened our introduction of mechanisms of balancing selection and prior work on inversion polymorphism, integrating many

suggested references. While especially helpful, these suggestions are too extensive to completely quote and respond to in this already-copious document. Therefore, we focus our response on two select topics from these comments, and then proceed to comment 5 thereafter.]

(2) The general reduction principle and inversion polymorphism. In Section 1.2., the authors state that "there has not been a proposed mechanism whereby alleles at multiple linked loci would directly benefit from linkage and thereby maintain an associated inversion polymorphism under indirect selection." Perhaps I am misunderstanding something, but in my reading, this statement is factually incorrect. In fact, the simplest version of Dobzhansky's epistatic coadaptation model

(see Charlesworth 1974; also see Charlesworth and Charlesworth 1973 and discussion in Charlesworth & Flatt 2021; Berdan et al. 2023) seems to be an example of exactly what the authors seem to have in mind here: two loci experiencing overdominance, with the double heterozygote possessing the highest fitness (i.e., 2 loci under epistatic selection, inducing some degree of LD between these loci), with subsequent capture by an inversion; in such a situation, a new inversion might capture a haplotype that is present in excess of random expectation (and which is thus fitter than average)...

We agree that the quoted statement could be misleading and have rewritten it. We intended to point out that we are presenting a model in which all loci contribute additively (with respect to display) or multiplicatively (with respect to survival probability), without any dominance relationships or genetic interaction terms. And yet, the model generates epistatic balancing selection in a panmictic population under a constant environment. This represents a novel mechanism by which (the life-history characteristics of) a population would generate epistatic balancing selection as an emergent property, instead of assuming a priori that there is some balancing mechanism and representing frequency dependence, dominance effects, or epistatic interactions directly using model parameters. We have therefore refined the scope of the statement in question (lines 155-158).

(4) Hearn et al. 2022 on Littorina saxatilis snails.

A good reference. There is considerable work on ecotype-associated inversions in *L. saxatilis*, but we previously cut some discussion of this and of other populations with high gene flow but identifiable spatial structure for inversion-associated phenotypes (e.g. butterfly mimicry polymorphisms, *Mimulus*, etc.). Due to the spatially discrete environmental preferences and sampled ranges of the inversions in these populations, we considered these examples to be somewhat distinct from explaining inversion polymorphism in a potentially homogenous and panmictic environment.

(4) cont. A very interesting paper that may be worth discussing is Connallon & Chenoweth (2019) about dominance reversals of antagonistically selected alleles (even though C&C do not discuss inversions): AP alleles (with dominance reversals) affecting two or more life-history traits provide one example of such antagonistically selected alleles (also see Rose 1982, 1985; Curtsinger et al. 1994) and sexually antagonistically selected alleles provide another. The two

are of course not necessarily mutually exclusive, thus making a conceptual connection to what the authors model here.

We had removed a previously drafted discussion of dominance reversal for brevity's sake, but this topic is once again represented in the updated draft of the manuscript with a short reference in the introduction, lines 76-80. We also mention 'segregation lift' (Wittmann et al. 2017) involving a similar reversal of dominance for fitness between temporally fluctuating conditions, as opposed to between sexes or life history stages.

(5) The model. In general, the description of the model and of the simulation results was somewhat hard to follow and vague. There are several aspects that could be improved: [5](1) it would help the reader if the terminology and distinction of inverted vs. standard arrangements and of the three karyotypes would be used throughout, wherever appropriate.

We have attempted to do so, using the suggested heterokaryotypic/homokaryotypic terminology.

[5](2) The mention of haploid populations/situations and haploid loci (e.g., legend to Figure 1) is somewhat confusing: the mechanism modelled here, of course, requires suppressed recombination in the inversion/standard heterokaryotype; and thus, while it may make sense to speak of haplotypes, we're dealing with an inherently diploid situation.

While eukaryotes with haploid-dominant life history may still experience similar dynamics, we do expect that most male display competition is in diploid animals, and we are only simulating diploid fitnesses and experimenting with diploid *Drosophila*. We have tried to minimize the discussion of haploids in this draft.

[5](3) The authors have a situation in mind where the 2 karyotypes (INV vs. STD) in the heterokaryotype carry distinct sets of loci in LD with each other, with one karyotype/haplotype carrying antagonistic variants favoring high male display success and with the other karyotype/haplotype carrying non-antagonistic alternative alleles at these loci and which favor survival. Thus, at each of the linked loci, we have antagonistic alleles and non-antagonistic alleles - however, the authors don't mention or discuss the degree of dominance of these alleles. The degree of dominance of the alleles could be an important consideration, and I found it curious that this was not mentioned (or, for that matter, examined).

In this study, our goal was to show that the investigated model could produce balanced and increasing antagonism without the need to invoke dominance. We think there would be a strong case for a follow-up study that more investigates how dominance and other variables impact the parameter space of balanced antagonism, but this goal is beyond our capacity to pursue in this initial study. We've added several lines clarifying the absence of dominance from our investigated models, and pointing out that dominance could modulate the predictions of these models (lines 211-213, 278-282).

[5](4) In many cases, the authors do not provide sufficient detail (in the main text and the main figures) about which parameter values they used for simulations; the same is true for the Materials & Methods section that describes the simulations. Conversely, when the text does mention specific values (e.g., 20N generations, 0.22-0.25M, etc.), little or no clear context or justification is being provided.

We have sought to clarify in this draft that 20N was chosen as an ample time frame to establish equilibrium levels and frequencies of genetic variation under neutrality. We present a time sequence in Figure 5, and these results indicate that that antagonism has stabilized in models without inversions or with higher recombination rates, whereas its rate of increase has slowed in a model with inversions and lower levels of crossing over.

The inversion breakpoints and the position of the locus with stronger antagonistic effects in Figure 4 were chosen arbitrarily for this simple proof of concept demonstration, with the intent that this locus was close to one breakpoint. Hopefully these and other parameters are clearer in the revised manuscript.

5 The authors sometimes refer to "inversion mutation(s)" - the meaning of this terminology is rather ambiguous.

Edited, hopefully the wording is clearer now. The quoted phrase had uniformly referred to the origin of new inversions by a mutagenic process.

(6) Throughout the manuscript, especially in the description and the discussion of the model and simulations, a clearer conceptual distinction between initial "capture" and subsequent accumulation / "gain" of variants by an inversion should be made. This distinction is important in terms of understanding the initial establishment of an inversion polymorphism and its subsequent short- as well as long-term fate. For example, it is clear from the model/simulations that an inversion accumulates (sexually) antagonistic variants over time - but barely anything is said about the initial capture of such loci by a new inversion.

We do not have a good method of assessing a transition between these two phases for the simulations in which both antagonistic alleles and inversions arise stochastically by a mutagenic process. However, we have tried to be clearer on the distinction in this draft: we have included simulations in Figure 4 with variants starting at lower frequencies, and we have tried to better contextualize the temporal trajectories in Figure 5 as (in part) modeling the accumulation of variants after such an origin.

Reviewer #3 (Recommendations For The Authors):

- In general: the whole paper is quite long, and I felt that many parts could be written more clearly and succinctly - the whole manuscript would benefit from shortening, polishing, and making the wording maximally precise. Especially the Introduction (> 8 pages) and Discussion (7.5 pages) sections are quite long, and the description of the model and model results was quite hard to follow.

We have attempted to condense some portions of the manuscript, but inevitably added to others based on important reviewer suggestions. Regarding the length Introduction and Discussion, we are covering a lot of intellectual territory in this study, and we aim to make it accessible to readers with less prior familiarity. At this point, we have well over 100 citations – far more than a typical primary research paper – in part thanks to the relevant sources provided by this reviewer. We are therefore optimistic that our text will provide a valuable reference point for future studies. We have also made significant efforts to clarify the Results and Methods text in this draft without notably expanding these sections.

- In general: the conceptual parts of the paper (introduction, discussion) could be better connected to previous work - this concerns e.g. the theoretical mechanisms of balancing selection that might be involved in maintaining inversions; the general, theoretical role of antagonistic pleiotropy (AP) and trade-offs in maintaining polymorphisms; previously

made empirical connections between inversions and AP/trade-offs; previously made empirical connections between inversions and sexual antagonism.

In the revised manuscript, we have improved the connection of these topics to prior work.

- L3: "accumulate". A clearer distinction could be made, throughout, between initial capture of alleles/haplotypes by an inversion vs. subsequent gain.

Please see point 6 in the response to the Public Review, above.

- L29: I basically agree about the enigma, however, there are quite many empirical examples in *D. melanogaster* / *D. pseudoobscura* and other species where we do know something about the nature of selection involved, e.g., cases of NFDS, spatially and temporally varying selection, fitness trade-offs, etc.

At least for our focal species, we have emphasized that geographic (and now temporal) associations have been found for some inversions. For the sake of length and focus, we probably should not go down the road of documenting each phenotypic association that has been reported for these inversions, or say too much about specific inversions found in other species. As indicated in our response to reviewer 2, some previously documented inversion-associated trade-offs may be compatible with the model presented here. However, we did locate and add to our Discussion one report of frequency-dependent selection on a *D. melanogaster* inversion (Nassar et al. 1973).

- L43: it is actually rather unlikely, though not impossible, that new inversions are ever completely neutral (see the review by Berdan et al. 2023).

This line was intended to convey that, in line with Said et al. 2018's results, the structural alterations involved in common segregating inversions are not expected to contribute significantly to the phenotype and fitness (as indicated by lack of strong regulatory effects), and that their phenotypic consequences are instead due to linked variation. We have rewritten this passage to better communicate this point, now lines 44-52. Interpreting Section 2 and Figure 1 of Berdan et al. 2023, the linked variation may be what is in mind when saying that inversions are almost never neutral. We have also added a line referencing the expected linked variation of a new inversion (lines 49-52).

- L51-73: I felt this overview should be more comprehensive. The model by Kirkpatrick & Barton (2016) is in many ways less generic than the one of Charlesworth (1974) which essentially represents one way of modeling Dobzhansky's epistatic coadaptation. Also, the AOD mechanism is perhaps given too much weight here as this mechanism is very unlikely to be able to explain the establishment of a balanced inversion polymorphism (see Charlesworth 2023 preprint on bioRxiv). NFDS, spatially varying selection and temporally varying selection (for all of which there is quite good empirical evidence) should all be mentioned here, including the classical study of Wright and Dobzhansky (1946) which found evidence for NFDS (also see Chevin et al. 2021 in Evol. Lett.)

On reflection, we agree that we put too much emphasis on AOD and have edited the section to be more representative.

*- L57. Two earlier Dobzhansky references, about epistatic coadaptation, would be: Dobzhansky, T. (1949). Observations and experiments on natural selection in *Drosophila*. *Hereditas*, 35(S1), 210-224. <https://doi.org/10.1111/j.1601-5223.1949.tb03334.x>; Dobzhansky, T. (1950). Genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. *Genetics*, 35, 288-302. <https://doi.org/10.1093/genetics/35.3.288> - In general, in the introduction, the*

classical chapter by Lemeunier and Aulard (1992) should be cited as the primary reference and most comprehensive review of D. melanogaster inversion polymorphisms.

- L101: this is of course true, though there are some exceptions, such as In(3R)Mo.

- L110: the papers by Knibb, the chapter by Lemeunier and Aulard (1992), and the meta-analysis of INV frequencies by Kapun & Flatt (2019) could be cited here as well.

Citation suggestions integrated.

- L123 and elsewhere: the common D. melanogaster inversions are old but perhaps not THAT old - if we take the Corbett-Detig & Hartl (2012) es7mates, then most of them do not really exceed an age of Ne generations, or at least not by much. I mean: yes, they are somewhat old but not super-old (cf. discussion in Andolfatto et al. 2001).

Edited to curb any hyperbole. We agree that there are much more ancient polymorphisms in populations.

- L133-135. This needs to be rewritten: this claim is incorrect, to my mind (Charlesworth 1974; also see Charlesworth and Charlesworth 1973; discussion in Charlesworth & Flatt 2021).

Edited. See public review response (2).

- L154: the example of inversion polymorphism is actually explicitly discussed in Altenberg's and Feldman's (1987) paper on the reduction principle.

Edited to mention this. Inversions are also mentioned in Feldman et al. 1980, Feldman and Balkau 1973, Feldman 1972, and have been in discussion since the origins of the idea.

- L162ff: see Connallon & Chenoweth (2019).

Citation suggestion integrated, along with Cox & Calsbeek 2009 which seems more directly applicable, now line 185ff.

- L169: why? There is much evidence for other important trade-offs in this system.

Reworded.

- L178-179: other studies have found that trade-offs/AP contribute to the maintenance of inversion polymorphisms, e.g. Mérot et al. 2020 and Betrán et al. 1998, etc.

Added Betrán et al. 1998 - a good reference. Moved up mention of Mérot et al. 2020 from later in the text and directed readers to the Discussion, lines 202-205.

- L198. "alternate inversion karyotypes" - you mean INV vs. STD? It would be good to adopt a maximally clear, uniform terminology throughout.

Edited to communicate this better.

- L215-217: this is a theoretically well-known result due to Hazel (1943); Dickerson (1955); Robertson (1955); e.g., see the discussion in the quantitative genetics book by Roff (1997) or in the review of Flatt (2020).

Citations integrated, now lines 232ff.

- L223 and L245: "haploid" - somewhat confusing (see public review).
- L259-260: This may need some explanation.
- L261-262: simply state that there is no recombination in *D. melanogaster* males.

Edited for increased clarity.

- L274 (and elsewhere): the meaning of "mutation...of new..inversion polymorphisms" is ambiguous - do you mean a polymorphic inversion and hence a new inversion polymorphism or do you mean polymorphisms/variants accumulating in an inversion?
- L275: maybe better heterokaryotypic instead of heterozygous? (note that INV homokaryotypes or STD homokaryotypes can be homo- or heterozygous, so when referring to chromosomal heterozygotes instead of heterozygous chromosomes it may be best to refer to heterokaryotypes).

Per [5](1) and 5 in the public review, we have edited our terminology.

- L276: referral to M&M - I found the description of the model/simulation details there to be somewhat vague, e.g. in terms of parameter settings, etc.

Further described.

- L281-282: would SLiM not have worked?

See public review response.

- L286-287: why these parameters?

Further described.

- L296ff: it is not immediately clear that the loci under consideration are polymorphic for antagonistic alleles vs. non-antagonistic alternative alleles - maybe this could be made clear very explicitly.

Edited to be explicit as suggested.

- L341, 343: "inversion mutation" - meaning ambiguous.
- L348, 352: "specified rate" - vague.
- L354-357: initial capture and/or accumulation/gain?
- L401, 402, 404: Z-, W- and Y- are brought up here without sufficient context/explanation.

The above have been addressed by edits in the text.

- L523, 557, 639, 646, and elsewhere: not the first evidence - see the paper by Mérot et al. (2020) (and e.g. also by Yifan Pei et al. (2023)).

Citations integrated in the introduction and discussion. Mérot et al. (2020) was cited (L486 in original) but discussion was curtailed in the previous draft.

- L558-559. I agree but it is clear that there are many mechanisms of balancing selection that can achieve this, at least in principle; for some of them (NFDS, etc.) we have pretty good evidence.
- L576-577. This is correct but for *In(3R)C* that study did find a differential hot vs. cold selection response.

Addressed with text edit.

- L584-L586: cf. Betrán et al. (1998), Mérot et al. (2020), Pei et al. (2023), etc.
- L591. "other forms of balancing selection": yes! This should be stressed throughout. Multiple forms of balancing selection exist and they are not mutually exclusive.
- L593: consider adding Dobzhansky (1943), Machado et al. (2021)
- L596-597: this is rather unlikely, at least in terms of inversion establishment (see Charlesworth 2023; <https://www.biorxiv.org/content/10.1101/2023.10.16.562579v1>).
- L608: consider adding Kapun & Flal (2019).
- L611-612: see studies by Mukai & Yamaguchi, 1974; and Watanabe et al., 1976.
- L639, 646: AP - see general literature on AP as a factor in maintaining polymorphism (Rose 1982, 1985; Curtsinger et al. 1994; Charlesworth & Hughes 2000 chapter in Lewontin Festschrift; Conallon & Chenoweth 2019 - this latter paper is particularly relevant in terms of AP effects in the context of sexual antagonism)

Citation suggestions integrated.

- L657: inversion polymorphism is explicitly discussed in Altenberg's and Feldman's (1987) paper on the reduction principle.

Hopefully this is better communicated.

- L724-755: I felt that this section generally lacks sufficient details, especially in terms of parameter choices and settings for the simulations.
- L732L: why not state these rates?

Parameter values are now given a fuller description in figure legends and in the methods.

- L746: but we know that mutational effect sizes are not uniformly distributed (?).

We made this choice for simplicity and to avoid invoking seemingly arbitrary distribution, but one could instead simulate trait effects with some gamma distribution. Display values would still have variable fitness effects that fluctuate with population composition, but we agree that distribution shifted toward small effects would be more realistic.

- L765: *In(3R)P* is not mentioned elsewhere - is this really correct?

That was incorrect, fixed.

<https://doi.org/10.7554/eLife.93338.2.sa0>