**Research Article**

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

Christopher S. McAllester[1,2], John E. Pool[1,2]

[1]*Laboratory of Genetics, University of Wisconsin – Madison, USA*

[2]Corresponding Authors: cmcallester@gmail.com (CSM), jpool@wisc.edu (JEP)

# 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.

# 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 that another 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.

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

113 Yet, it is unclear whether ecological adaptation is a sufficient explanation for *D. melanogaster* inversion frequency
114 patterns, for two reasons. First, most of the common inversions become less frequent as one samples populations
115 farther from the sub-Saharan ancestral range (or similar warm environments) into potentially more challenging,
116 colder high latitude and altitude environments (Aulard et al. 2002; Kapun et al. 2016; Lack et al. 2016a; Pool et al.
117 2017; Kapun and Flatt 2019) with *In(3R)Mo* representing an exception to this pattern (Kapun et al. 2014; Kapun et
118 al. 2016). As these inversions are all derived (being absent from closely related species), local adaptation theory
119 would predict that they arose to protect local adaptation against ongoing migration, more likely in the novel
120 environment if it is experiencing high levels of immigration (Kirkpatrick and Barton 2006). It is unclear why
121 inversions would be needed to adapt to the species' ancestral environments, or why ancestral "standard"
122 arrangements would have an ecological advantage in derived environments that are very different from the
123 ancestral range.

124 Second, even in the locations where inversions would hypothetically be ecologically advantageous, they rarely
125 surpass intermediate frequencies (Lemeunier and Aulard 1992; Pool et al. 2017; Kapun and Flatt 2019;
126 Sprengelmeyer et al. 2020). Under a local adaptation model, that observation would require a secondary

127 explanation that resists fixation, such as very high levels of long distance gene flow or the fixation of recessive
128 deleterious variants on inverted haplotypes (Ohta and Kimura 1970; Zhao and Charlesworth 2016; Gilbert et al.
129 2020). While migration may well limit inversion frequencies in some populations, we note that some putatively
130 adaptive Single Nucleotide Polymorphism (SNP) variants do achieve much greater frequency differentials between
131 populations (da Silva Ribeiro et al. 2022), and thus we would have to suppose that inversions are under relatively
132 weaker selection to explain their lesser differentiation. Regarding the hypothesis of high recessive load, we note
133 that *D. melanogaster* populations do generally contain substantial recessive genetic load (Greenberg and Crow
134 1960), and it has been speculated that associative overdominance between sampled backgrounds might boost
135 inversion frequencies in *D. melanogaster* populations during founder events (Pool et al. 2012) or in isolated
136 wilderness environments (Sprengelmeyer et al. 2020). However, it seems likely that in large human-commensal
137 populations, *D. melanogaster* inversions would escape from linked deleterious variants through gene conversion or
138 double crossover events with standard chromosomes, in light of both the age of most of these inversions (Corbett-
139 Detig and Hartl 2012) and the very high population recombination rates of *D. melanogaster*. Homozygous lines for
140 each common inversion also appear at appreciable frequency in inbred lab strains, demonstrating a lack of perfect
141 linkage of the inversions with recessive lethal or infertile variants (*e.g.* Lack et al. 2016a).

142 While local adaptation, gene flow, and associative overdominance may all contribute to geographic variation in
143 inversion frequency among *D. melanogaster* populations, it is not clear that these processes are collectively
144 sufficient to explain observed patterns. Alternatively, the allele frequencies and diversity patterns of at least some
145 of these inversions might be primarily shaped by some form of balancing selection acting on inversion-linked
146 variation, in which the balanced frequency is dependent on the environment (Kapun et al. 2023). Balancing
147 selection on inversions has received increasing attention in recent literature (Wellenreuther and Bernatchez 2018;
148 Faria et al. 2019). In *Drosophila*, inversions have been associated with balancing selection due to seasonal,
149 temporally fluctuating selection (Machado et al. 2021). However, the potential for temporally varying selection to
150 maintain stable balanced polymorphisms on its own is somewhat limited (Gillespie 1988). Further, most common
151 inversions maintain relatively high polymorphic frequencies across broad geographic ranges without obvious
152 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 benefit an individual in competitive mate choice systems often decrease in fitness as the beneficial allele increases in frequency, 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. Albeit, the direction and magnitude of the frequency dependent effect can depend on the way mate choice is parameterized even among "best of n" models, as well as the strength of selective versus stochastic effects (O'Donald 1973; O'Donald et al. 1997). Combinations of display-enhancing alleles at multiple loci 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). In a mate choice system such as a best of n model in which a female chooses the highest quality male encountered, a marginally higher quality male similarly outcompetes either a marginally lower or an exceptionally worse male. So, the last few alleles to push display into the top quantile should 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

178 sexually antagonistic variants that have the greatest fitness when carried together. This tendency of selection to
179 favor the reduction of recombination when the haplotypes present are at a fitness optimum is well studied for
180 recombination modifiers including inversions, and has been termed the 'reduction principle' (Feldman 1972;
181 Feldman and Balkau 1973; Feldman et al. 1980; Altenberg and Feldman 1987; Altenberg et al. 2017; see Altenberg
182 and Feldman 1987 for a direct treatment of inversions).

183 While the emergent epistasis generated by a mate choice system may favor linkage of variants, alleles that
184 contribute only to a sexually selected display are not expected to remain as balanced polymorphisms, rather they
185 are expected to fix, albeit perhaps slowly as their fitness benefits diminish at high frequencies (Hazel 1943;
186 Dickerson 1955; Charlesworth and Hughes 2000). However, if these alleles simultaneously contribute deleteriously
187 to other components of fitness, they are pleiotropic (Stearns 2010), and given the sex-limited benefits of the
188 display they would likely also be sexually antagonistic (Rowe et al. 2018), whereby traits increase the fitness of one
189 sex but decrease fitness if found in the other. Pleiotropy can result from traits that are inherently physiologically
190 coupled, from resource allocation restraints, or from the specific role or behavior of genes. Pleiotropy can further
191 be sexually antagonistic when the traits in the tradeoff are differentially important to different sexes, which may
192 be common given the differing fitness requirements of the sexes in a population (Cox and Calsbeek 2009;
193 Connallon and Clark 2014; Connallon and Chenoweth 2019). It should be noted that a model of two pleiotropically
194 determined traits each with potentially sex-specific fitness can be treated separately from a model where one trait
195 has sexually antagonistic fitness effects.

196 *D. melanogaster* mating involves a male display performance and female acceptance or rejection. A number of
197 experimental evolution studies (Rice 1992; Rice 1996; Audet et al. 2024) and genetic studies (Innocenti and
198 Morrow 2010; Cheng and Kirkpatrick 2016; Glaser-Schmitt et al. 2023) have confirmed the prevalence of sexually
199 antagonistic loci in *D. melanogaster*. Pleiotropic alleles that trade off between survival and male display quality
200 represent one clear potential case of antagonism to consider for a model of balancing selection, and one highly
201 relevant to the biology of *D. melanogaster*. Many such antagonistically pleiotropic scenarios are possible under
202 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.

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 that is subject to balancing selection and for which intermediate states are inherently disfavored. 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.

# 2 Results

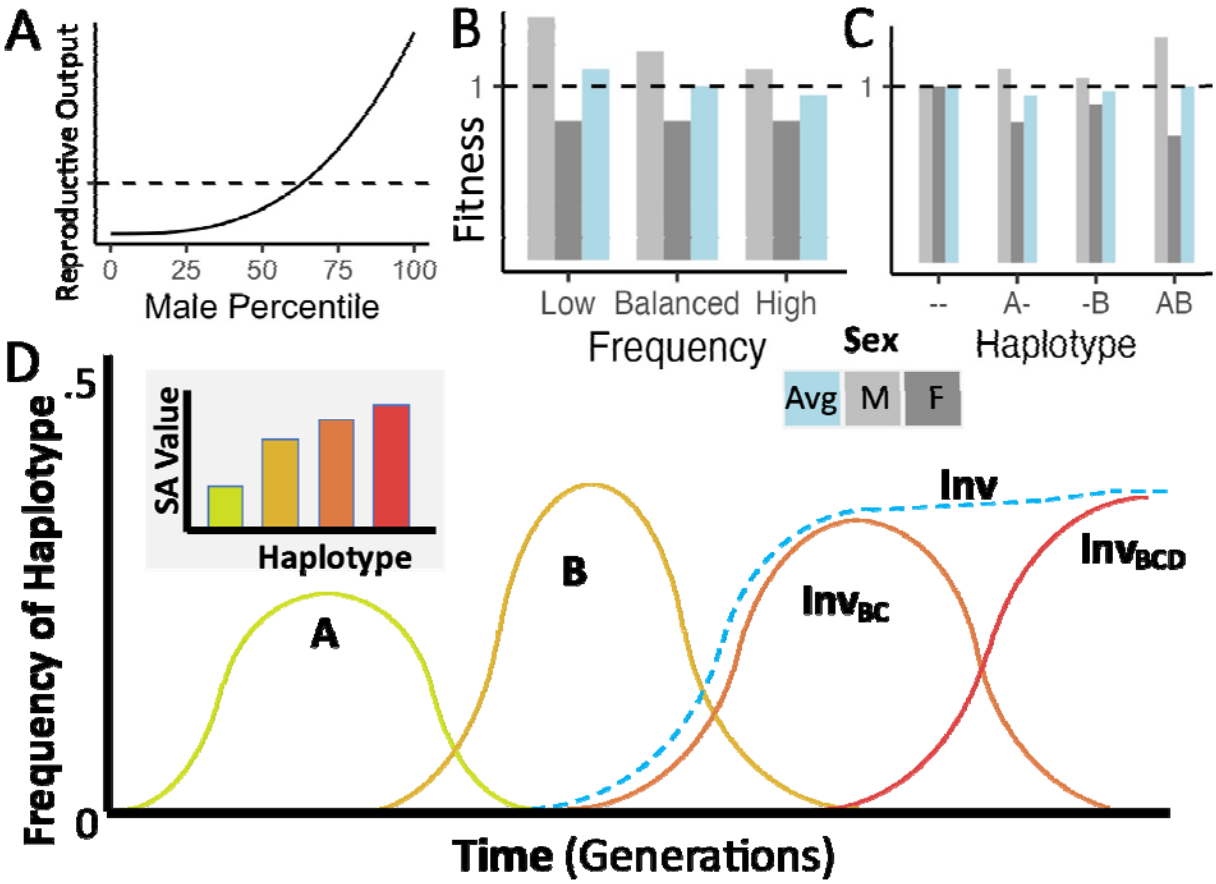
## 2.1 A Model of Sexually Antagonistic Inversions

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 either 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 from substituting alleles at two loci, or from substituting two homologous alleles at a single locus in a diploid or polyploid organism, 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 frequency of comparable display qualities of other males in a best of n model, the relative advantage of a display trait can diminish as more males carry it, for both haploid and diploid populations (Figure 1B, male fitnesses; O'Donald 1973; O'Donald et al. 1997). 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. Whereas, if the variant is antagonistically pleiotropic for a trait 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 note that this dynamic can be described as a "negative frequency dependence" of the fitness component for mate choice alone, as the selection coefficient of the favored allele may decrease towards zero while remaining positive, though some authors instead use negative frequency dependence to specifically describe fitness functions in which all alternate alleles are of selective benefit when rare (*e.g.* Charlesworth and Charlesworth 2010). In addition to the relationship between sex averaged fitness and frequency, the contrasting sex-specific fitnesses of antagonistic alleles under this model could further enhance the potential for balanced polymorphism.

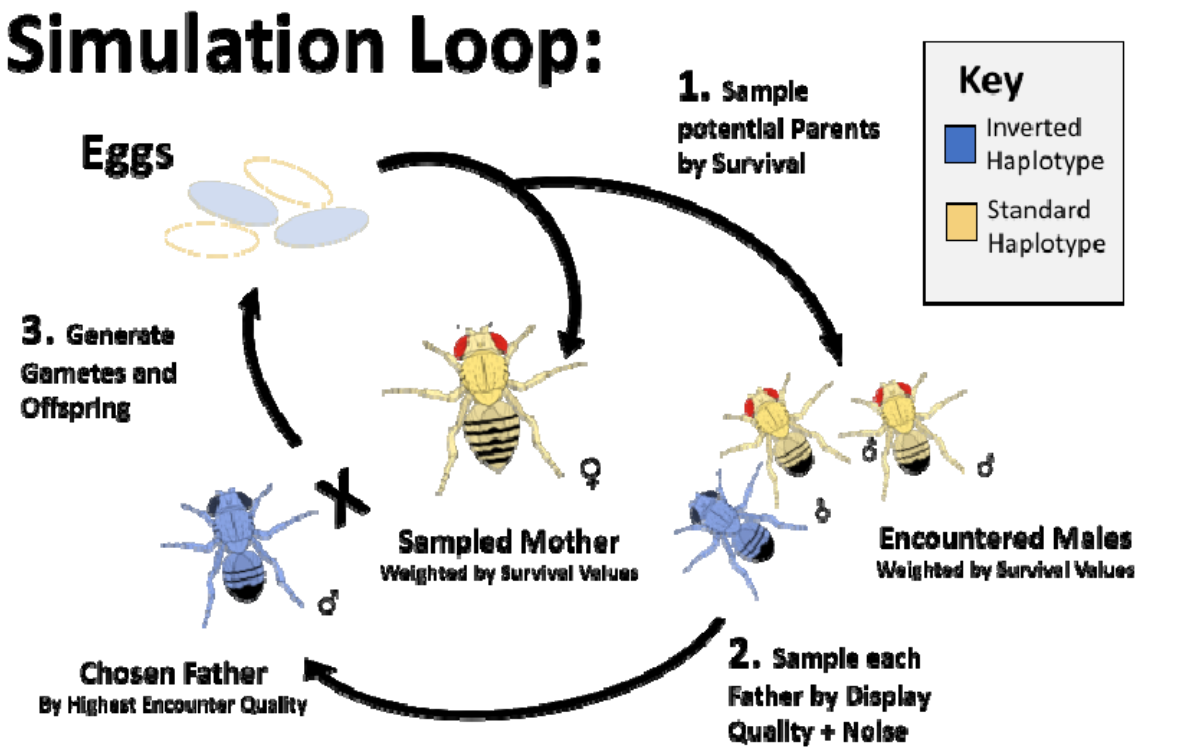
We further predict that multiple such balanced pleiotropic loci should generate interesting emergent behavior. Because of the emergent synergistic epistasis from mate competition (see discussion of Figure 1A above), the haplotype with the greatest male display benefit should increase in frequency relative 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.




[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, such as expected under a "best of n" mate choice 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 furthers the accumulation of antagonistic variants, reaching an equilibrium frequency while displacing less extreme display-favoring haplotypes.]

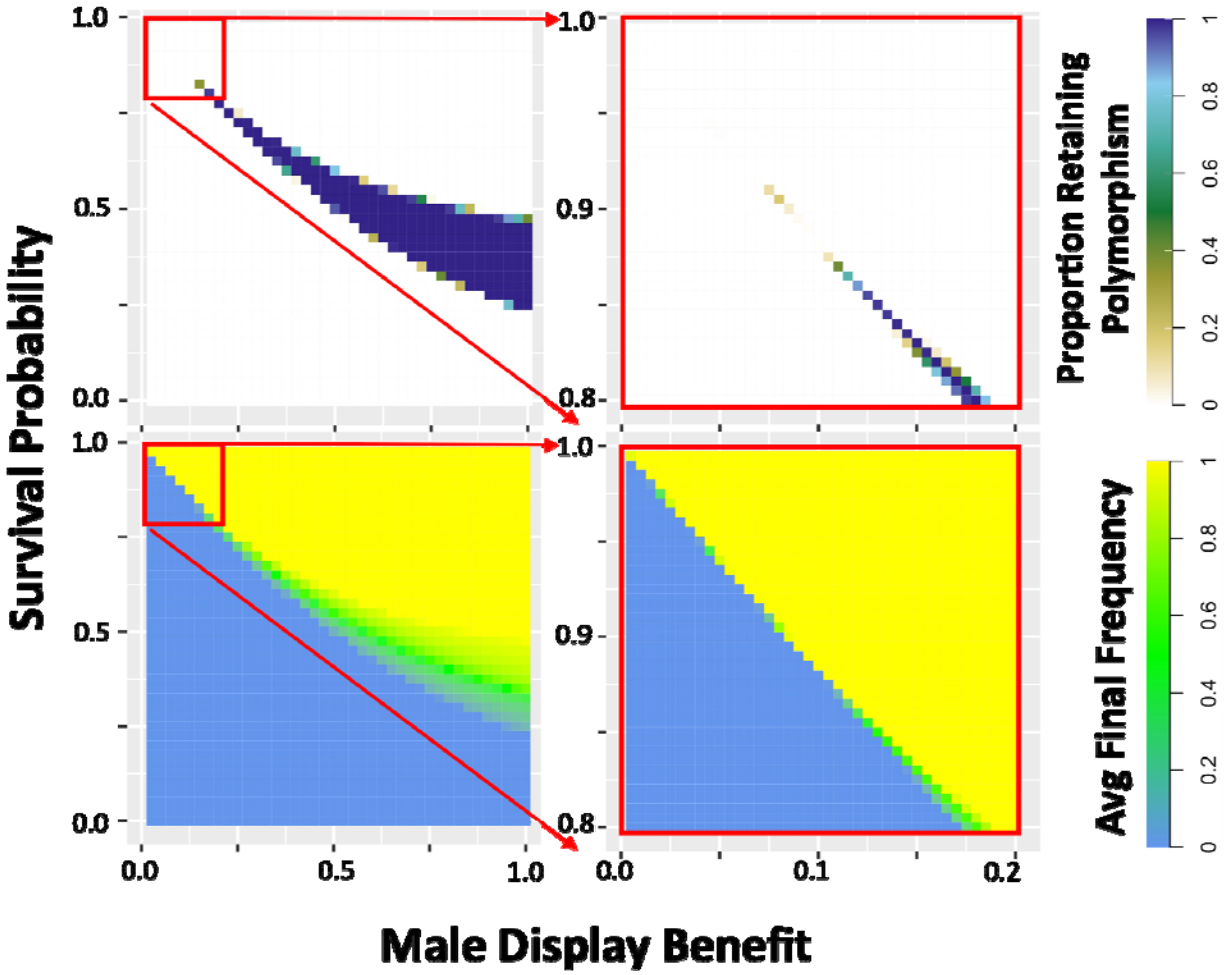
While Figure 1 illustrates a haploid scenario for the sake of simplicity, this model extends naturally to a diploid case. The main difference is that heterokaryotypic diploids – those heterozygous for the arrangement – would have intermediate trait values and be less fit than either homokaryotype, for 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




[**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 for the best of n mate choice - without replacement for each pool and with replacement between pools. Unless otherwise noted, we defaulted to 100 encountered males in simulations presented here. 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 there is one antagonistic allele with the indicated effects reducing survival proportion (of both sexes) and increasing male reproductive quality score, initiated at equal frequency with an alternative allele

313 that has no such effects. Larger values of both effect sizes are plotted in the left panels, with values incrementing
314 by 0.025, and smaller effect sizes (a subset of the larger) are plotted in greater detail in the right panels, with
315 values incrementing by 0.005. Reproductive quality contributes to male success in a best of 100 mate competition
316 with an added normally distributed noise effect of standard deviation 1 (Figure 2; Materials and Methods). See
317 Figure S1 for plots with sex-specific costs and differing encounter numbers in the best of n mate competition. Most
318 parameters result in either all simulations retaining polymorphism (purple), or else all simulations losing
319 polymorphism (white). However, simulations retaining polymorphism have a broad range of equilibrium
320 frequencies (shades of green). For each parameter combination, 500 replicates were simulated for a diploid
321 population of 1,000 individuals evolving for 20,000 generations, with further details as indicated in the Materials
322 and Methods.]

323

324 We investigated the predictions of the above model of inversion-associated balanced sexual antagonism using a
325 purpose-built forward simulation program, "Sexually Antagonistic Inversion simulator" (SAIsim; see Materials and
326 Methods). This simulation program incorporates the generation of new nucleotide and inversion polymorphisms
327 by mutation, where each nucleotide mutation independently draws a positive or negative effect on survival and on
328 male display. It also simulates crossing over and gene conversion, and inherently the effects of drift (Materials and
329 Methods). SAIsim includes gene conversion between all homologous sites regardless of karyotype, and models
330 crossing over in heterokaryotypic individuals from even numbers of crossover events within the inverted region. It
331 explicitly models mate selection, in that for each female selected to reproduce, a set number of 'encountered'
332 males are randomly selected, and the observed display value of each male for this competitive event is assigned
333 based on the sum of the male's genetic effects and a noise effect resampled at every encounter. The male with the
334 highest observed display value then succeeds in this mating (Figure 2).

335 With this simulator, we first confirmed that some individual variants with an antagonistic pleiotropy between
336 survival (initially of both sexes) and male reproduction can generate stable selectively balanced polymorphisms *in*
337 *silico*. We simulated a single locus in diploid populations of $N$ = 1000, polymorphic for an allele of a specified

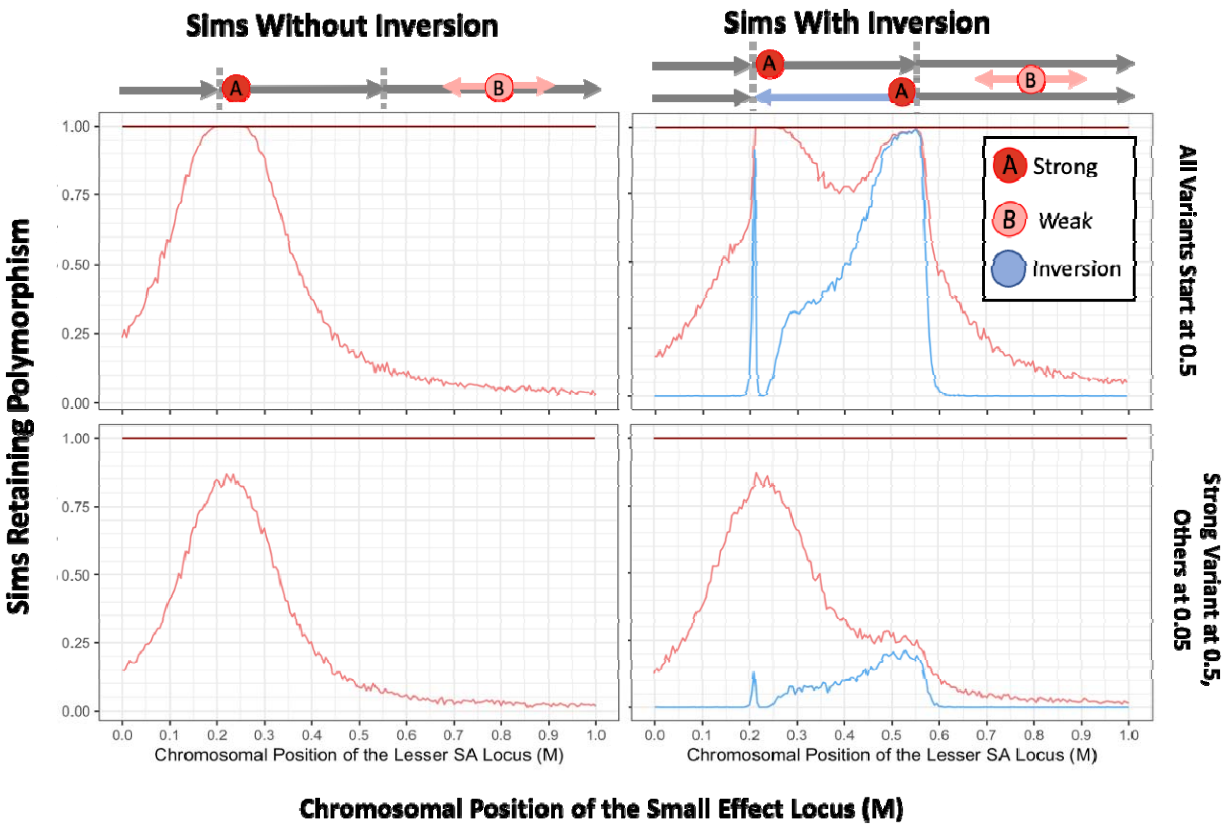
338 survival cost and display benefit at an initial frequency of 0.5, with the alternate allele having no trait effect. For
339 each female's mating competition, 100 males were sampled, though see Figure S1 for plots with varying encounter
340 number. After 20*N* generations, we calculated the proportion of simulations retaining polymorphism and the
341 average final frequency of the variant across all simulations, revealing that there is a range of paired trait effects
342 for which the mutation consistently remained polymorphic, with intermediate mean frequency, after 20*N*
343 generations (Figure 3) – implying the action of balancing selection. Alleles with small effect sizes had only a narrow
344 band of selective maintenance, outside of which they were either fixed or lost, while large-effect alleles remained
345 balanced across a larger area of simulated parameter space. While the overall balanced range is somewhat
346 narrow, such variants may persist under balancing selection indefinitely in the absence of other forces, and many
347 small-effect alleles in linkage blocks may confer larger effects in synergy. In a scenario where only females pay the
348 survival cost, the balanced parameter space of antagonistic trait effects is broader and shifted somewhat (Figure
349 S1). In a scenario where only males pay the survival cost, and variants are neutral in females, essentially no
350 simulations retain polymorphism (Figure S1). The number of males sampled to assess mate competition for each
351 offspring also has a significant effect, with larger values giving broader polymorphic ranges, though simulations
352 with costs only in females show robust polymorphism even when only sampling two males for each competition
353 (Figure S1).

354 We then tested the behavior of two linked polymorphic loci with antagonistic alleles. We selected antagonistic
355 effects for these alleles from variants previously found to undergo balancing selection separately, and performed
356 simulations with varying genetic distances between them (Figure 4). Crossing over and gene conversion were
357 simulated, with gene conversion occurring for each heterozygous variant at a probability of $10^{-2}$ per female meiosis
358 per site, with random direction. This behavior was intended to be a conservative approximation of the number of
359 gene conversion events affecting a given site in *D. melanogaster*, scaled to simulations of *N*=1000 (Comeron et al.
360 2012). The stronger antagonistic variant (modelled at fixed position 0.225M) persisted in all simulations, in line
361 with expectations, and so we focus on whether the weaker variant (at varying positions) also persisted with it. As
362 seen in the upper panel of Figure 4, the weaker variant (pale red line) was retained at close linkage, as the
363 haplotype containing both antagonistic variants acts as a single larger effect supergene. The haplotypes with

364 intermediate trait values were less fit, leaving less offspring per haplotype across simulations (Figure S2A), whereas
365 the extreme haplotypes were maintained at much higher frequency (Figure S2B), generating linkage disequilibrium
366 between display-favoring variants ($r^2$=0.2139 across simulations in which the smaller effect variant was positioned
367 at 0.250M). This demonstrates the synergistic epistasis proposed in our verbal model. The maintenance of
368 haplotypes AB and ab under selection while Ab and aB are selectively removed relies on the lesser relative fitness
369 of Ab and aB. And since survival values are multiplicative, this additional contribution must come from the mate
370 success of AB being disproportionately larger than Ab or aB. As the weaker variant was then moved farther from
371 the stronger variant, recombination increasingly generated unfit intermediate genotypes compared to the
372 extremes of the reproduction-survival tradeoff, and the stronger variant outcompeted the weaker variant among
373 these haplotypes despite the epistatic benefit from associating the two variants together, leading to the loss of the
374 weaker variant. In the scenario examined by Figure S2, the differences among haplotypes in sex-averaged fitness
375 are modest, whereas the sex-specific fitness effects are more dramatic. These results could indicate that in
376 addition to the frequency-dependent dynamics proposed above, the contrasting fitness effects in females and
377 males may also play an important role in maintaining balanced polymorphism, particularly when near frequency-
378 dependent balance (Connallon and Clark 2014; Kaufmann et al. 2023).

379 We then examined the same spacings between antagonistic variants across a region containing a polymorphic
380 inversion, with its left breakpoint close to the stronger variant, and the right breakpoint within the range of the
381 positions in which the weaker variant was simulated (Figure 4, lower panel). When the inversion was added, the
382 weaker variant was then retained when the two antagonistic variants were relatively more distant, due to
383 association of the weaker variant with the right-hand inversion breakpoint (pale red line), at which point this
384 association also preserved the inversion itself in a polymorphic state (light blue line). While less predictable, a
385 weaker antagonistic variant near the middle of the inversion also appeared to maintain polymorphism more
386 frequently in systems starting with an inversion than without, i.e. across approximate positions 0.30-0.45M in
387 Figure 4, in spite of its weaker linkage with the inversion itself due to double cross-over events. The inversion was
388 not maintained when the weaker antagonistic variant was within the inversion but tightly linked to the stronger
389 variant (see positions 0.22-0.25M); here there is little crossing-over for an inversion to suppress, and so the effect

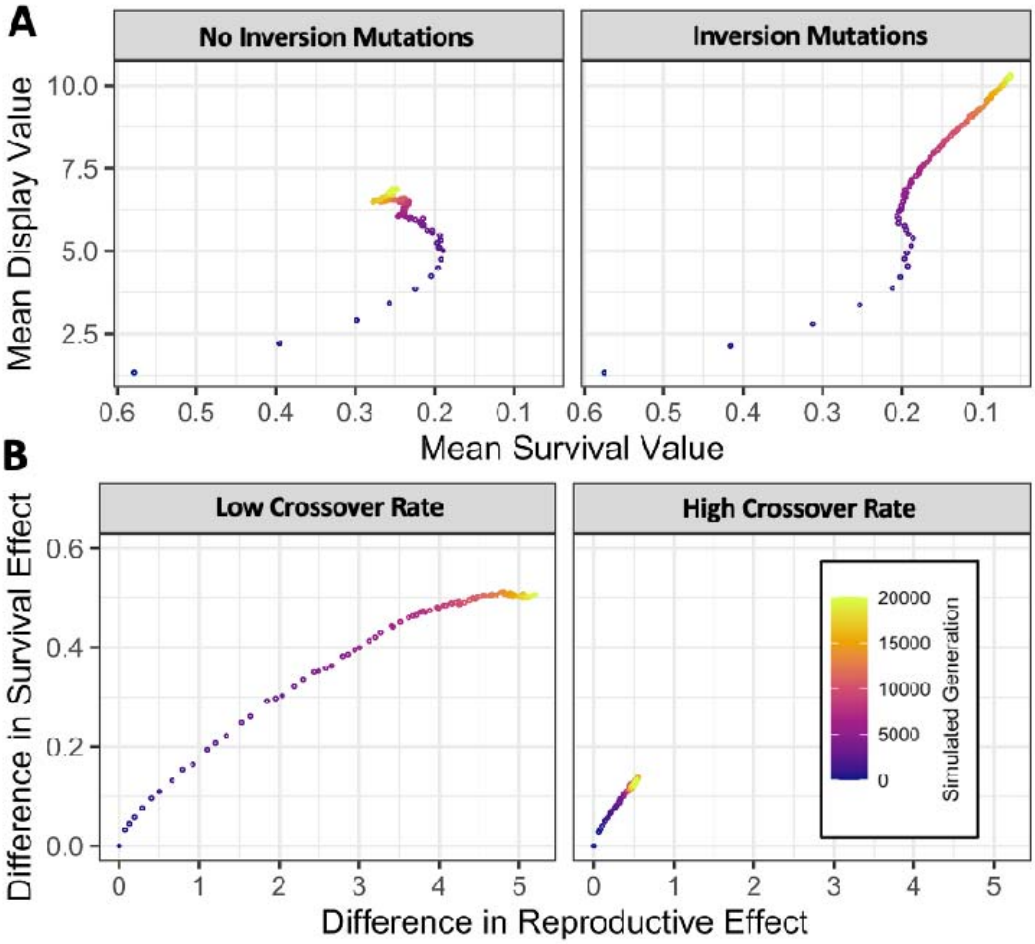
of the inversion approaches neutrality. Interestingly, there was a small window of weaker variant position in which the inversion remained polymorphic despite the already tight linkage between SA variants, where the inversion seemingly does not modify recombination much (see positions 0.2-0.22M). Here, 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.




403

404 [Figure 4: Simulated inversions may persist as polymorphisms when linked to sexually antagonistic, pleiotropic
405 variants. In turn, the presence of inversions facilitates the accumulation of antagonistic variation, demonstrating
406 synergistic epistasis. Rates of long-term persistence are shown for simulations with two defined antagonistic
407 variants alone (left panels) and with the additional presence of a defined inversion (right panels). To focus the
408 potential addition of a second linked variant to an antagonistic haplotype, we begin with the stronger variant
409 already at an intermediate frequency in the population (0.5). The weaker antagonistic variant and the inversion
410 either both start at 0.5 frequency as well (upper row) or both start at a lower frequency of 0.05 (lower row), with
411 all variants initially in linkage equilibrium. The 1 Morgan chromosomal segment is diagrammed above the plots,
412 with the inversion breakpoints marked (0.21M, 0.55M), chosen to include space across and beyond the inverted
413 region. The proportion of simulations in which the second, variably positioned weaker antagonistic variant
414 (survival and reproductive values 0.86, 0.12) maintains polymorphism is shown (light red line), dependent on the
415 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 20*N* 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.]



425 [**Figure 5**: Simulated populations with stochastic mutation of new inversions and sexually antagonistic variants
426 generate persistent inversion polymorphisms and accumulate linked sexually antagonistic variants. (A) In simulated
427 populations with stochastic mutation of new variants having pleiotropic effects on survival and male display, levels
428 of antagonism increased with time before stabilizing at an intermediate value (left panel). Whereas, simulations
429 which also allowed the mutation of new inversions continued accumulating stronger antagonism throughout the
430 simulated time interval (right panel). (B) In simulations which allowed the mutation of new inversions, a low
431 crossover rate (left panel, 43.6 cM in females after scaling) allowed populations to accumulate larger differences
432 between the most common karyotype and all others than when the rate of crossover was particularly high (right
433 panel, 43.6 M in females after scaling). Gene conversion occurs at each heterozygous variant with a scaled
434 probability of 0.1295 per heterozygous variant in both scenarios. Values given are the mean of 1000 replicate
435 simulations of 1000 diploid individuals run for 20*N* generations.]

436

437 In further simulations, we began with a population initially free of variation, in which sexually antagonistic
438 mutations arise stochastically. To assess differences between longer or shorter antagonistic regions, or higher or
439 lower crossover rates, we modeled populations with parameters scaled to either 1 Mb or 10 kb, using estimates
440 from Comeron et al. (2012) and Huang et al. (2016). The mutations had pleiotropic effects increasing reproductive
441 quality score and decreasing survival proportion, each drawn sampled independently (Materials and Methods).
442 Depending on the effect scores drawn, a mutation could be broadly advantageous (potentially leading to a
443 selective sweep), broadly deleterious, or potentially subject to balancing selection (Figure 3, Figure S1). These
444 mutations arose at a rate equal to one thousandth of the scaled single nucleotide mutation rate given in Huang et
445 al. (2016). Inversions (of random lengths and positions) also arose stochastically, at a hundredth of the rate of
446 antagonistic variants. The frequency trajectories of antagonistic variants and inversions were tracked, and the
447 simulation was run for 20*N* generations to allow near-equilibrium patterns to be established. These simulations
448 demonstrated the potential of a population without initial genetic variation to establish alternate karyotypes
449 which link sexually antagonistic, pleiotropic alleles and an inversion, to maintain polymorphism in a balanced

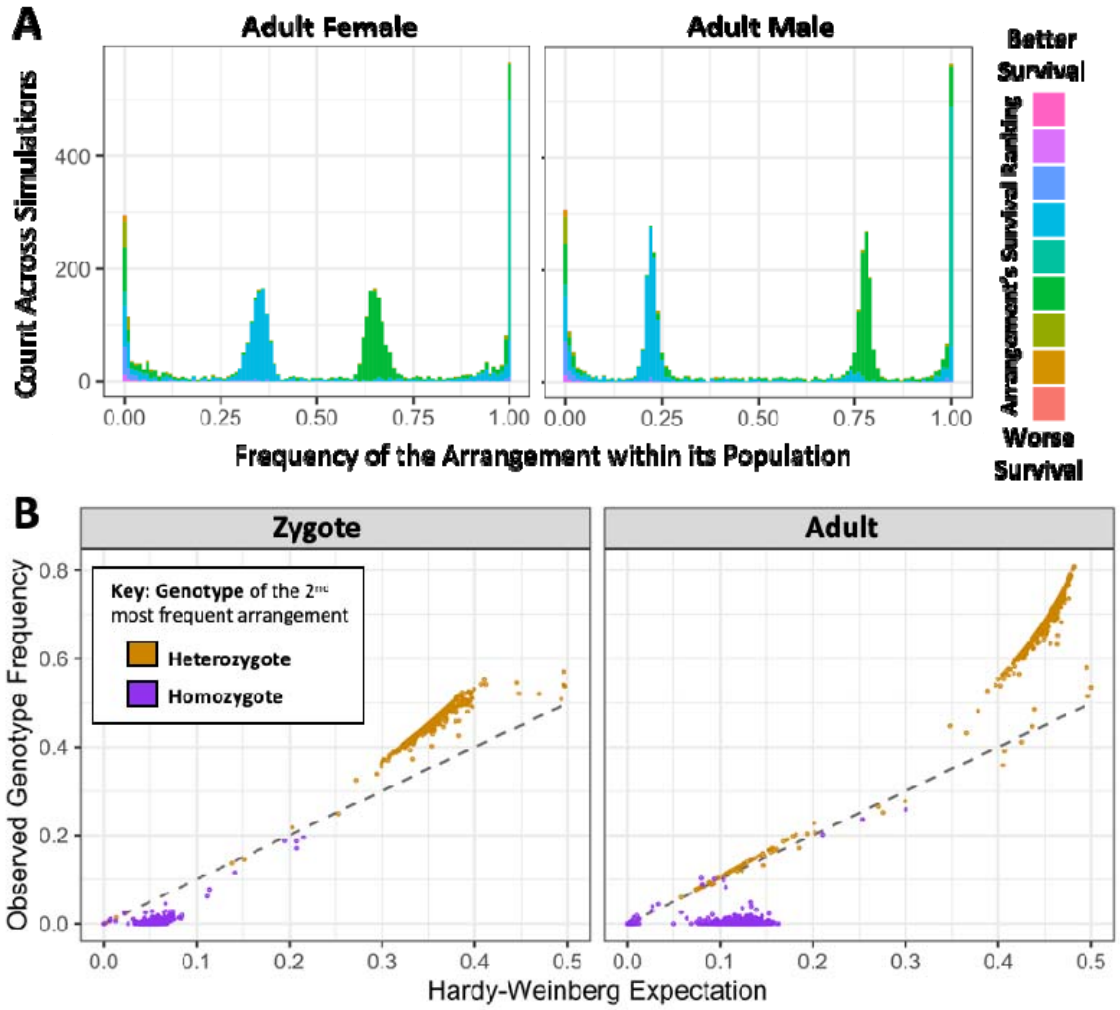
450 manner, and then to subsequently accumulate significant antagonistic character (Figure 5A), although the degree
451 of antagonism accumulated depends strongly on the rate of crossing over (relative to the occurrence of
452 antagonistic mutations; Figure 5B).

453 The simulations that successfully maintained balanced antagonistic chromosomes also had an interesting
454 distribution of allele and genotype frequencies of the balanced haplotypes at equilibrium. Without any selection,
455 the inversion frequencies followed a pattern characteristic of the expected neutral site frequency spectrum (Figure
456 S3). But strikingly, in parameter spaces where simulations with shared survival costs between sexes maintained
457 polymorphism under balancing selection, inversion frequency distributions were tightly clustered around 0.25 and
458 0.75 among the zygote population (Figure S3, 6A). We found that this outcome was related to the genotype
459 frequencies of reproducing adults. These simulations evolve toward strong antagonism, and it is predominantly
460 males homozygous for the display-favoring haplotype (whether that is inverted or standard) who succeed in
461 reproducing, and these males therefore contribute one copy of the display-favoring haplotype to each of their
462 male and female offspring. However, this haplotype's significant cost to viability and longevity would select for the
463 heterokaryotypic females over the display-favoring homokaryotypes. Given that successful parents would
464 predominantly be heterokaryotypic females and display-favoring homokaryotypic males, around three quarters of
465 transmitted chromosomes have the display-favoring arrangement and one quarter does not, yielding the 0.25 and
466 0.75 zygote inversion frequencies seen in simulations. Although these are autosomal simulations, selection yields
467 an inheritance pattern similar to that of a ZW sex chromosome system, where the Z chromosome is
468 homokaryotypic in the fathers, and heterokaryotypic with the W chromosome in the mothers.

469 The sex-specific fitness dynamics present at equilibrium under this model also resulted in population-wide
470 genotype frequencies that depart from standard Hardy-Weinberg expectations. If the preponderance of matings
471 are between heterokaryotypic females and homokaryotypic display-favoring males, then homokaryotypes for the
472 survival-focused haplotype will rarely be produced (as confirmed in Figure 6B). We note that in cases where the
473 standard arrangement favors male display, then especially if antagonism is strong, this model could explain an
474 observed deficiency or absence of inversion homokaryotypes, even if the inverted arrangement was not associated

475 with any recessive deleterious load. Further, heterokaryotypes may become disproportionately frequent after
476 viability selection (Figure 6B). And still, this enrichment for heterokaryotypes reflects neither the existence of
477 recessive deleterious load associated with each arrangement, nor overdominance with respect to viability. Hence,
478 our model may offer an alternative explanation for deviations from Hardy-Weinberg genotype frequencies
479 observed for polymorphic inversions in natural populations.

480




481

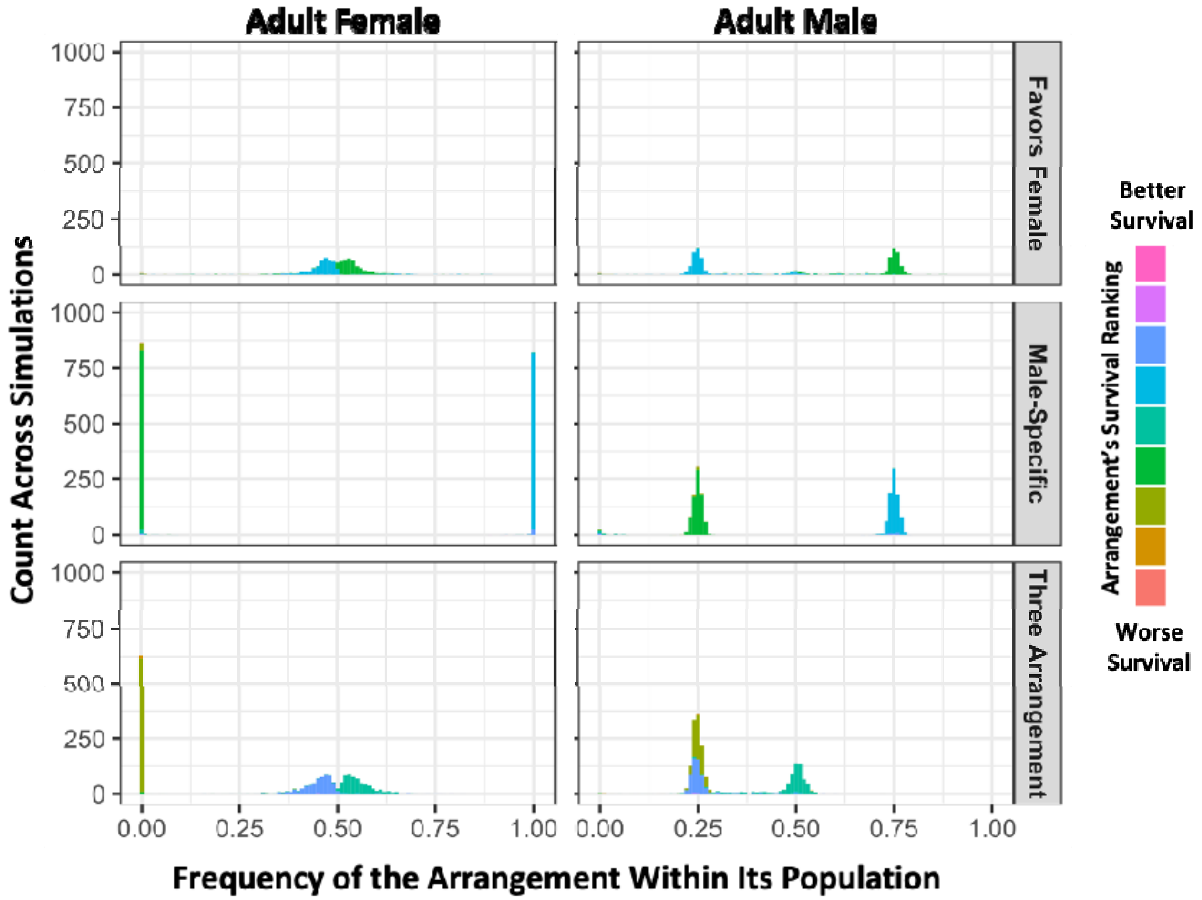
482 [**Figure 6:** Antagonism-associated karyotypes reach predictable equilibrium frequencies in simulations with
483 antagonistic and inversion mutations, when survival costs are shared between sexes. (A) Histograms of the number

of karyotypic 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 between zygote (left panel) and adult (right panel) populations as the genotype affects survival.]

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). In 20.0% of simulations, we observed 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. For 37.0% of simulations, there were again 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. In 40.6% of simulations, we instead detected 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 (Figure 7).

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

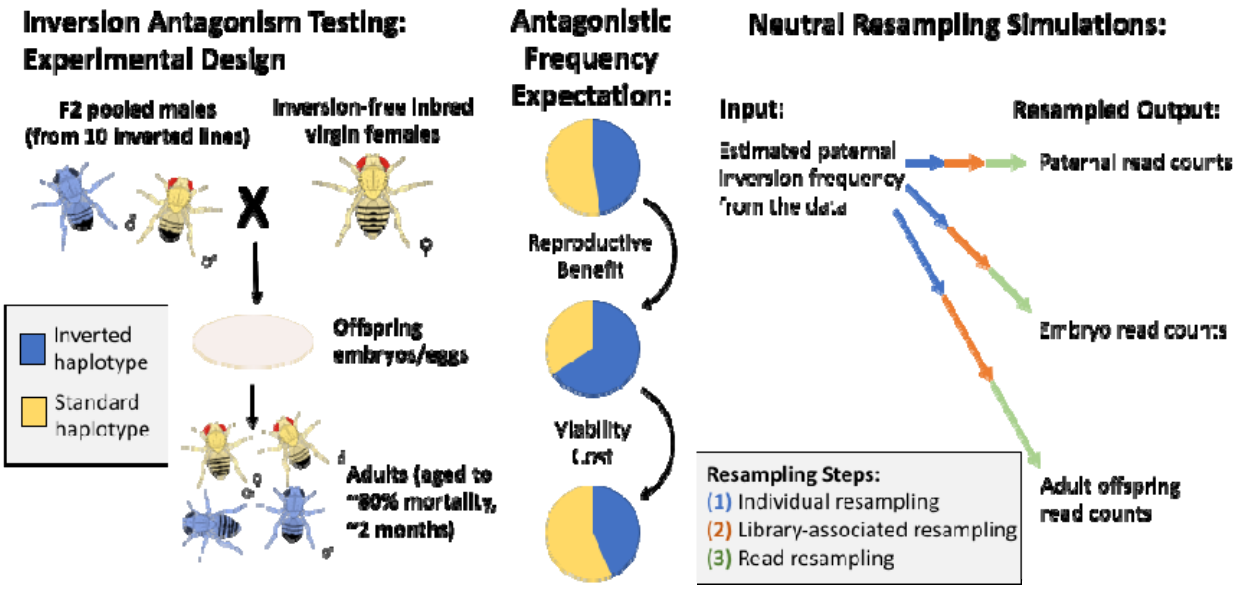
or not, though two chromosomes sampled from the population may have distinct inversions with overlapping breakpoints. This means that if two arrangements, each with an inversion, become the prevalent haplotypes and thereby remove all standard arrangement chromosomes from the population, there is no mutational 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 which 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 in this set of simulations.



[**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).




542

543 [**Figure 8:** The layout and potential expectations of the laboratory evolution experiment. We cross outbred males
544 from a high-inversion population to inbred, non-inverted females from a specific inbred line, and collect DNA
545 samples from the fathers, embryo offspring, and aged adult offspring for inversion frequency estimation via
546 sequencing (left). We hypothesize from our antagonistic pleiotropy model that one or more of the four inversions
547 present on separate chromosome arms in our experimental populations may experience opposing selection
548 between fathers and embryos versus between embryos and aged adults (middle). To test the significance of
549 observed frequency changes, we designed a resampling model to represent the sources of neutral variation
550 expected in the experiment (right).]

551

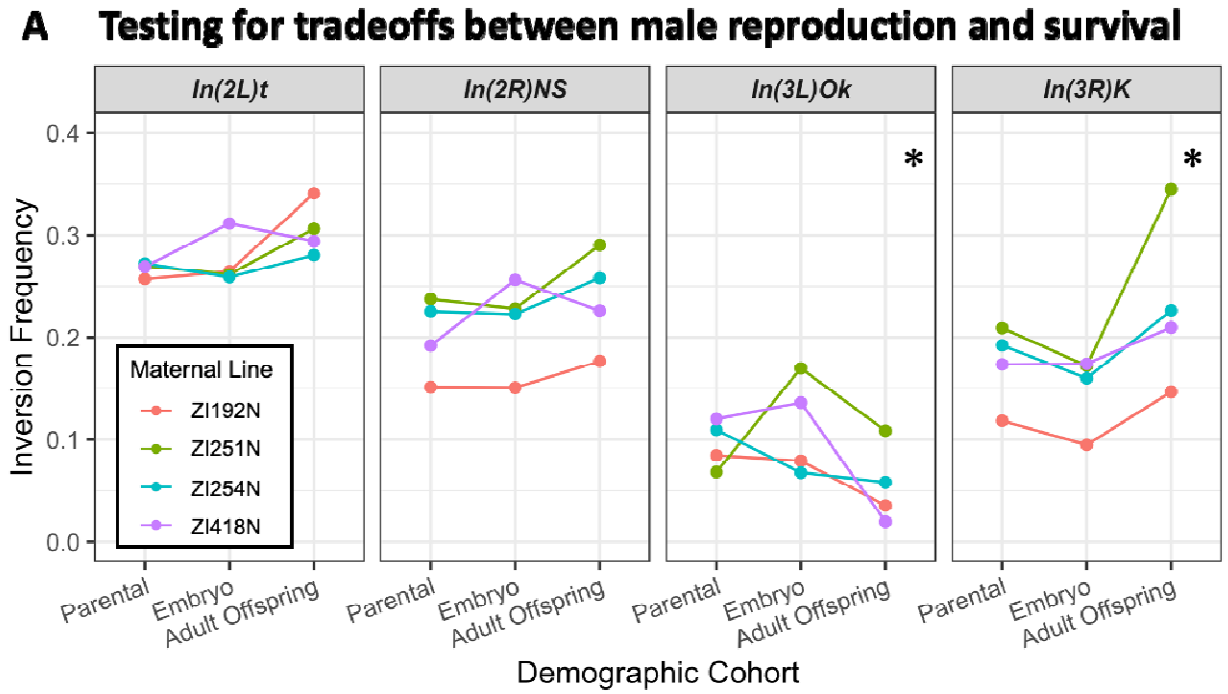
552 We assembled an outbred population with moderate frequencies of each inversion to obtain karyotypically diverse
553 males. We set up four crosses using these males along with females from one of four different inversion-free
554 inbred Zambia lines (Figure 8). Each of these parental crosses occurred in a single large population cage, with
555 populations of at least 600 and equal numbers of each sex (Table S1). We developed a PCR and next-generation
556 sequencing-based assay to estimate inversion-associated SNP frequencies in each pool of males and in each
557 corresponding pool of embryos. These SNPs (given in Table S2) were based on our own analysis of variants that

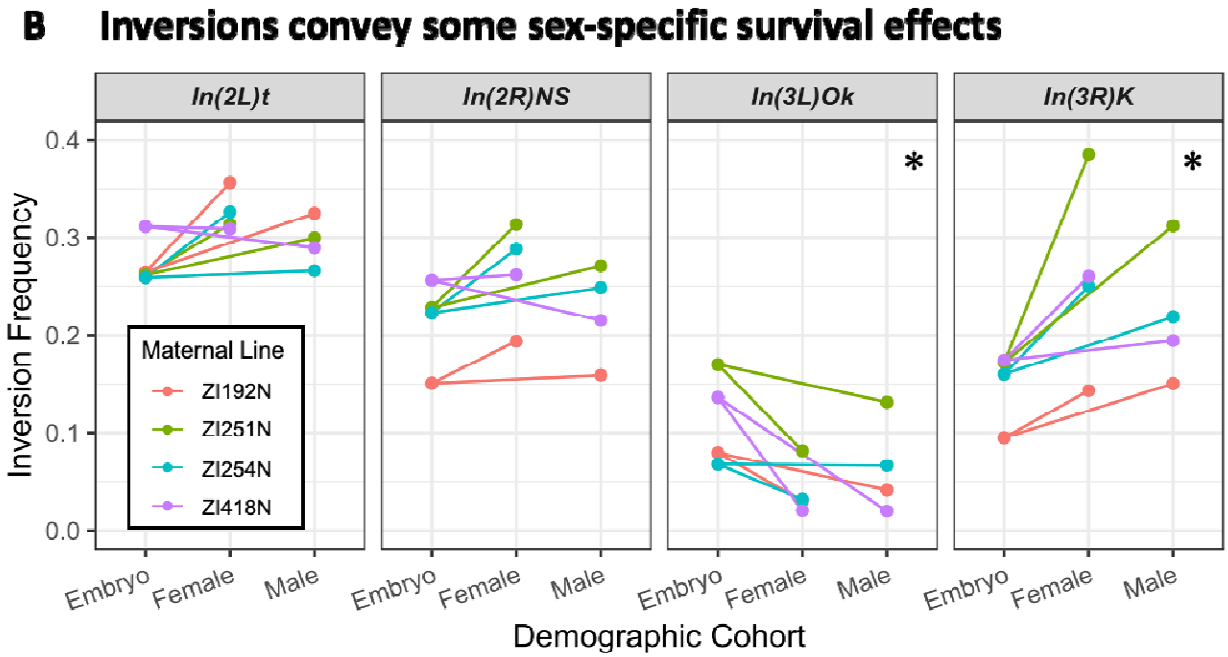
558 showed perfect association with inversions among 197 Zambia genomes (Table S3; see Materials and Methods).
559 We then implemented a resampling method to test for significant inversion frequency changes between parents
560 and offspring (while accounting for the variance observed in technical replicates; see Materials and Methods), in
561 order to determine whether males carrying each inversion were more or less successful in reproducing than the
562 average male. We also compared frequencies between embryos and aged adults collected 10 weeks after eclosion,
563 to investigate the relationship between inversions and viability/longevity. These aged adults eclosed either in the
564 first or second half of the collection period from the same batch of eggs, thus allowing us to test for inversion
565 associations with development time. We calculated p-values from the resampling tests to assess the significance of
566 each individual change in inversion frequency between relevant samples (Table S4). However, in accord with our
567 primary question regarding fitness tradeoffs, we primarily focus here on joint tests of two opposing frequency
568 changes, as further described below. We could therefore test the hypothesis that a given inversion had opposing
569 influences on survival and male reproduction, which would provide evidence consistent with our model of
570 balanced sexual antagonism, while also examining whether any such tradeoffs depend on maternal genetic
571 background.

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

584 multiple inversions harbor alleles that benefit survival or longevity under our experimental conditions (see

585 Discussion).




586

587 [**Figure 9**: Some *Drosophila* inversions show evidence for antagonistic tradeoffs and sex-specific survival. (A)
588 Testing for antagonistic fitness effects on reproduction and survival (encompassing viability and longevity). Each
589 inversion's frequency in parents after collection (2 to 3 weeks from eclosion), embryo offspring, and aged adult
590 offspring (10 weeks from eclosion) are shown for each maternal inbred line. Parental frequency is represented
591 instead of paternal frequency for easier visual comparison, and is taken as the paternal frequency divided by two
592 to account for the inversion-free inbred mothers. Adult offspring frequencies are taken from the sum of estimated
593 allele counts across all adult offspring cohorts. * indicates p-value < 0.05 for evidence of significant reversed
594 frequency change from parental to embryo and from embryo to offspring, combined across all four maternal line
595 crosses and corrected for multiple tests, when compared to neutral simulated experiments (details in Materials
596 and Methods). (B) Testing for evidence of sex-specific survival effects. The changes in inversion frequencies
597 between embryos and the aged adult offspring are again illustrated, now separated by their sex. * indicates p-
598 value < 0.05 for significant differences in female versus male inversion frequency among adult offspring, combined
599 across all four maternal crosses and corrected for multiple tests.]

600

601 When we separated the effects of male reproduction (based on father to embryo changes) and survival (embryos
602 versus adults), we found that the effects of inversions on survival were generally consistent across maternal
603 genotypes. Three inversions were associated with increased survival to 10 weeks: *In(3R)K* was in all four crosses,
604 while *In(2L)t* and *In(2R)NS* each showed elevated survival with all maternal strains except ZI418N (Figure 9A). In
605 contrast, *In(3L)Ok* was associated with lower survival in each cross (Figure 9A). With regard to male reproduction,
606 the trajectory of inversion frequencies from fathers to embryos to offspring was less consistent between different
607 crosses (Figure 9A), implying that differences in male reproductive success may be dependent on the maternal
608 genotype, which is consistent with past findings in *D. melanogaster* (Reinhart et al. 2015). Despite such variability,
609 each of the four inversions had multiple frequency changes that departed strongly from null expectations (Figure
610 9A; Table S4), as might be expected if selection was present but varied in strength and direction between crosses
611 and life stages.

612 To test our central hypothesis of a tradeoff between male reproductive success and cumulative survival during
613 development and adult aging, we used neutral resampling to test the likelihood of observing the magnitudes of
614 opposing inversion frequency changes (between fathers and embryos, and between embryos and aged adult
615 offspring) under the null hypothesis, integrating across the four maternal crosses and accounting for multiple test
616 correction (Materials and Methods). For *In(3R)K*, which had particularly consistent changes in frequency across
617 replicate lines (decreasing between parents and embryos, and then increasing between embryos and aged
618 offspring; Figure 9A), this frequency reversal was significantly non-random ($p = 2.70 \times 10^{-3}$; Table S5). In contrast to
619 *In(3R)K*, *In(3L)Ok* was associated with (variably) higher male reproductive success and (consistently) lower survival;
620 this reversal was also statistically significant ($p = 2.97 \times 10^{-7}$).

621 We also tested for frequency differences between aged adult offspring of different sexes and from early-eclosing
622 versus late-eclosing cohorts. All four inversions displayed differences in frequency between female versus male
623 offspring (Figure 9B), consistent with differential effects on sex-specific survival. Strikingly, *In(3L)Ok* and *In(3R)K*
624 each displayed sex survival differences in alignment with their antagonistic effects implied above: *In(3L)Ok* (which
625 favored male reproduction) reduced female survival more strongly than male survival ($p = 0.0450$), whereas
626 *In(3R)K* (associated with lower male reproduction) favored the survival of females more than it did males ($p = 4.61$
627 $\times 10^{-3}$; Table S5; see Materials and Methods). The other two inversions, *In(2L)t* and *In(2R)NS*, each consistently
628 favored female over male survival in all four maternal crosses, although their raw unidirectional cross-combined $p$
629 values (0.0260 and 0.0126) were not significant after multiple test correction ($p = 0.282$ and 0.137). In contrast, we
630 found no consistent differences in inversion frequency between early- and late-eclosing offspring (Figure S4).

631 Estimated selection coefficients for each tested change based on a Wright-Fisher population approximation are
632 available in Table S6; however these may overestimate inversion fitness effects since the survival terms reflect the
633 probability of reaching advanced age. Further, there may be an influence of differential sex-specific mortality
634 during aging leading to greater frequency shifts in females. We collected a total of 1960 females and 3156 males,
635 and if the survivors represent 20% of an original population with an equal sex ratio, then there was 84.7% female
636 mortality and 75.3% male mortality. Additionally, there was some mortality among the parents over the laying

period, so some of the frequency change observed in embryos may be attributable to paternal mortality. The mean paternal mortality was 9.83% among the crosses, with a range of 5.11 to 15.77% mortality (Table S1), and the fathers were at most three weeks old, suggesting mild effects on inversion frequency estimation. 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)

659 over survival. These results provide significant but measured evidence for the presence of sexually antagonistic
660 autosomal inversions in this model species, while emphasizing the importance of genetic background.

661 As indicated in the Introduction, models like ours in which inversion frequencies are influenced by balancing
662 selection on pleiotropic tradeoffs may help explain multiple aspects of observed *D. melanogaster* inversion
663 frequencies. This class of models inherently accounts for the intermediate maximal frequency of most *D.*
664 *melanogaster* inversions in natural populations. To the extent that inversions favor male reproduction, the sexually
665 antagonistic model we propose here could also help account for the predominantly autosomal locations of
666 common inversions in this species, given that male-specific fitness effects have a weaker influence on the
667 frequencies of X-linked variants since the X chromosome spends only one third of its time in males.

668 This model could also help explain the observed geographic variability of inversion frequencies, as a change in
669 environment may change the fitness value of either pleiotropic effect of the alleles. If a population experiences a
670 novel harsh environment with low survivability, then the population density may decrease, and the number of
671 encountered competing males may decrease as well, reducing the relative fitness advantage of enhanced male
672 display. If there is a genetic tradeoff between mating display and viability, the equilibrium frequency of an
673 antagonistic inversion that favors the genetic extreme of display should be lower. Additionally, in hostile
674 environments, the survival cost of display-favoring variants could be increased, which could further shift their
675 equilibrium frequencies downward or render them not worth carrying at any frequency. These predictions could
676 contribute to the observed negative relationships between some inversion frequencies and latitude or altitude
677 (Kapun et al. 2016; Pool et al. 2017). As a specific example, it is notable that in a high altitude Ethiopian
678 environment with year-round cool weather, which hosts some of this species' most pronounced phenotypic
679 evolution (Bastide et al. 2014; Lack et al. 2016b), inversions are virtually absent - occurring at the lowest
680 frequencies of any analyzed population (Sprengelmeyer et al. 2020). Concordantly, geographically marginal
681 populations of other *Drosophila* species are often found to host less inversion polymorphism than more central
682 populations (Krimbas and Powell 1992). Also consistent with tradeoffs involving inversion-associated survival costs,
683 a disadvantage of some *D. melanogaster* inversions in challenging thermal environments (both cold and hot) was

684 supported by a previous experimental evolution analysis that included *In(2L)t*, *In(2R)NS*, and *In(3R)K* (Kapun et al.
685 2014), though they did find that *In(3R)C* consistently increased in frequency in hot conditions and *In(3R)Mo* in cold.

686 In contrast, our experiments yielded evidence that *In(3R)K* may have a tradeoff opposite from our prediction - with
687 the inverted arrangement associated with greater survival but lower male reproductive success. In line with our
688 findings, Said et al. (2018) found that among Zambia strains, *In(3R)K* was associated with strongly reduced
689 expression of *desat2*, involved in pheromone production. Whereas, higher *desat2* expression – as reported by Said
690 et al. (2018) for standard Zambia karyotypes – has been linked to the "Z-type" males preferred by females of some
691 strains from this region (*e.g.* Fang et al. 2002; but see Grillet et al. 2012). Conversely, a recent genome-wide
692 association study found *In(3R)K* heterokaryotypes to be associated with elevated male reproductive success
693 compared to standard homokaryotypes, when males from various North Carolina strains were paired with a Z-type
694 Zimbabwe strain known to exhibit preference for southern African male phenotypes (Yamamoto et al. 2024),
695 underscoring that inversion-related tradeoffs may depend on genetic background and/or experimental context.

696 Given that our model relies on the ability of either an inverted or a standard karyotype to place males carrying it
697 among the top reproductive competitors, it may seem surprising that we found two different inversions associated
698 with a tradeoff between survival and male reproduction – in that an individual's karyotype at one inversion could
699 undercut the advantage of their karyotype the other. In our case, it might seem unfavorable for a male carrying
700 In(3L)Ok (associated with increased male reproductive performance but lower survival) to also carry In(3R)K (which
701 showed the opposite patterns). Strikingly then, among 197 sequenced haploid embryo genomes from this Zambia
702 population (each reflecting one female gamete from independent isofemale lines; Lack et al. 2016a), not a single
703 genome carried both of these inversions (even though *In(3L)Ok* and *In(3R)K* were carried individually by 36 and 35
704 of these genomes respectively), representing a significant lack of this hypothetically disadvantageous doubly-
705 inverted karyotype ($p = 0.0031$, two-tailed Fisher's Exact Test). These inversions are 20.4 cM apart in terms of
706 meiosis in standard karyotype females (Comeron et al. 2012), but the contributions of natural selection versus
707 potentially altered meiotic dynamics in individuals heterozygous for two inversions in generating this pattern will
708 require further investigation.

709 More broadly, we observed a trend toward increased inversion frequencies across our full experimental
710 generation. A general benefit for inversions would mirror previous findings in which inversions found at
711 intermediate frequencies in natural *D. melanogaster* populations (Inoue 1979), as well as inversions from seaweed
712 flies known to be under balancing selection in the wild (Mérot et al. 2020), were both inferred to be under
713 directional selection in the lab. However, one laboratory study did report evidence that a *D. melanogaster*
714 inversion not tested here, *In(3R)P*, had frequency-dependent fitness under crowded conditions (Nassar et al.
715 1973). Furthermore, even an inversion found to be under directional selection in the lab might still be subject to a
716 balanced tradeoff under a more challenging or complex natural environment. Other studies in *D. melanogaster*,
717 including those on focused seasonal evolution, have found evidence for important tradeoffs between reproduction
718 and robustness to environmental challenges (*e.g.* Behrman et al. 2015). Hence, it would clearly be desirable to
719 study the effects of inversions on a wider range of potential tradeoffs. For example, Said et al. (2018) suggested
720 that inversions *In(2L)t* and *In(3R)K* might be maintained by immunity-related tradeoffs, in light of an over-
721 representation of immune-related genes among differentially expressed transcripts between inverted and
722 standard karyotypes.

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

733 As indicated in the Introduction, balancing selection on inversions could also be achieved by linkage to recessive
734 deleterious variants, which may be observed as an antagonistic pleiotropic phenotype despite potentially involving
735 strictly deleterious variants in linkage (Pei et al. 2023). In extreme form inversions might be under associative
736 overdominance (Zhao and Charlesworth 2016; Gilbert et al. 2020), where two haplotypes are balanced by
737 recessive cost to both homozygotes. Inversions may be particularly susceptible to carrying recessive deleterious
738 alleles if population size and/or inversion frequency are low, since under these circumstances inversion
739 homozygotes are rarely generated and hence the crossover rate among inverted chromosomes is low, plus there
740 are fewer inverted chromosomes to facilitate rare double recombination events with standard chromosomes.
741 However, this dynamic is not expected to establish new inversions easily (Charlesworth 2024) and is expected to
742 decay with increased age and population size of the inversion, as mutation, gene conversion, or double crossover
743 resolve the negative linkage between recessive deleterious alleles. Most common inversions in *D. melanogaster*
744 are both relatively old and have large effective population sizes. Inversions including *In(2L)t*, *In(2R)NS*, and *In(3R)K*
745 have breakpoints estimated to be several times older than the expansion of the species out of southern-central
746 Africa (Corbett-Detig and Hartl 2012; Sprengelmeyer et al. 2020). Those three inversions, plus *In(3L)Ok*, for which
747 inversion age has not been estimated, all maintain frequencies greater than 0.15 within very large populations
748 (Kapun and Flatt 2019; Sprengelmeyer et al. 2020), and so it is not clear that the persistence of recessive
749 deleterious or lethal alleles in the inverted regions is expected to be high enough that inversions would be
750 maintained by a shared recessive lethal variant preventing fixation. However, sampling of natural populations has
751 demonstrated a somewhat higher frequency of recessive lethal variants in inverted haplotypes, but these are
752 largely unique and heterokaryotypic individuals remain robust, giving low costs to a heavily outbred population
753 (Mukai and Yamaguchi 1974). Many wild-derived inbred strains of *D. melanogaster* are homokaryotypic for
754 inversions (Lack et al. 2016a). Furthermore, as noted in the Results, an observed lack of inversion homozygotes
755 could reflect strong sexual antagonism as opposed to recessive deleterious variation. Nevertheless, further
756 simulation and experimental testing regarding the potential influence of associative overdominance and other
757 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

782 locus would resolve the antagonism more frequently (though see following subsection on sex chromosome
783 evolution).

784 Perhaps more importantly, there are a number of well-studied cases of inversion polymorphism that might
785 potentially follow this dynamic of balancing selection between haplotypes simultaneously demonstrating
786 antagonistic pleiotropy and trait-specific frequency dependence, but which do not involve sexual antagonism. An
787 inversion polymorphism in the seaweed fly *Coelopa frigida* has been found to trade off between larval survival and
788 adult reproduction (Mérot et al. 2020). *Papilio* butterfly populations harbor well-studied mimicry polymorphisms
789 that segregate between inverted segments and may often experience frequency-dependent selection (Clarke and
790 Sheppard 1960; Kunte et al. 2014; Poul et al. 2014). The snail *Cepaea nemoralis* harbors diverse and distinct shell
791 patterning haplotypes associated with inversions that remain balanced within single sampling locations (Cook
792 2013; Surmacki et al. 2013). Several inversions in different ant species segregate a single-queen colony strategy
793 from a cooperative multi-queen strategy within the same range (Wang et al. 2013; Libbrecht and Kronauer 2014).
794 Some vertebrates found to harbor antagonistically pleiotropic inversions include the rainbow trout (Pearse et al.
795 2019) and the zebra finch (Pei et al. 2023). These balanced polymorphisms are variously hypothesized to involve
796 maintenance by overdominance (Mérot et al. 2020), linked deleterious recessives (Pei et al. 2023), or reversals of
797 dominance (Pearse et al. 2019), but these mechanisms are not mutually exclusive with each other, or with the
798 model presented here.

799 Essentially, any population with antagonistic pleiotropy involving some form of negative frequency-dependent
800 balancing selection on one of the phenotypes might exhibit the same dynamic. The population would experience
801 similar divergent selection towards alternate phenotypic extremes, with the frequency-dependence maintaining
802 the presence of two genotypes. Importantly, this model involves a different mechanism for the maintenance of
803 antagonistically pleiotropic variation than models based on reversals in dominance, where dominance effects
804 generate a net heterozygote advantage at a locus (Connallon and Chenoweth 2019). The model proposed here
805 integrates well with existing thought on the evolution of divergent selection on balanced systems. For example,
806 Kopp and Hermisson (2006) have proposed a specific form of frequency dependent divergent selection (FDDS),

807 involving competitive exclusion between similar genotypes for a trait otherwise under stabilizing selection, and this
808 model was expected to result in single or few loci of large effect. Their model did not consider pleiotropy or linkage
809 modifiers, but still had comparable dynamics in that selection in the FDDS regime favored small numbers of large
810 effect loci, while the inversion haplotypes in the model presented here simplify the recombination architecture to
811 approximate a single locus. This simplification of the genetic architecture follows the reduction principle in
812 recombination modifier theory (Feldman 1972; Feldman and Balkau 1973; Feldman et al. 1980; Altenberg and
813 Feldman 1987; Altenberg et al. 2017). When a population experiences selection towards two (or more) divergent
814 optima, under a selection regime that provides a balancing mechanism to maintain both, indirect selection will
815 favor a reduction of the genetic architectures to prevent the generation of intermediate phenotypes. The
816 likelihood of either model likely rests in part upon the inherent polygenicity of the trait under selection, with
817 simpler traits being more likely to resolve by single locus FDDS dynamics than by inversion association (*e.g.* Yassin
818 et al. 2016).

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

## 827 3.3 Antagonistic autosomal inversions may precede novel sex chromosomes

828 If balanced, sexually antagonistic, autosomal inversions occur frequently enough, they may also contribute to the
829 evolution of sex chromosomes. Given an established inversion with significant sexual antagonism, the inverted
830 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 20*N* 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 only 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 saxatalis* 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. In our simulations, antagonistic variants could be maintained at an equilibrium frequency, and inversions enabled 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

880 phenotypic impacts of *D. melanogaster* inversions would benefit from investigations of a wider range of fitness-
881 related traits that might be involved in balanced tradeoffs, including female fecundity and immunity. It will also be
882 important to verify the potential frequency dependence and phenotypic consequences of *D. melanogaster*
883 inversions under a wider range of conditions, including challenging environments such as cool temperatures and
884 high population densities. While inherently challenging, a further goal would be to identify loci within an inversion
885 such as *In(3L)Ok* or *In(3R)K* that contribute to observed tradeoffs, and in the case of an arrangement favoring male
886 reproduction, to test for the predicted synergistic epistasis (for male reproductive fitness) among linked variants.

887 Clearly, it will also be important to test the generality of our model by estimating the fitness associations of
888 inversions in a wider range of species. And particularly in the case of species with very recent changes sex
889 chromosomes, it will be of great interest to test whether one or more inversions are already present between
890 these new sex chromosomes, as predicted a model in which sex chromosome evolution involves the resolution of
891 inversion-associated sexual antagonism. Further analytic and simulation work will also be important in clarifying
892 the range of potential scenarios in which balancing selection and antagonistic pleiotropy may lead to the
893 maintenance of polymorphic haplotypes involving inversions (or parts of the genome that rarely recombine),
894 including studies that consider the effects of dominance and that consider tradeoffs other than the one examined
895 here. Collectively, such studies will advance our broader understanding of the roles of balancing selection,
896 pleiotropic tradeoffs, and genome structure in shaping genetic variation and genomic evolution.

897

# 898 4 Materials and Methods

## 899 4.1 Simulations

900 We developed a new individual-level forward simulator ('SAIsim' for Sexually Antagonistic Inversion simulator)
901 written in python 3.7 for the simulation of inversions and sexually antagonistic alleles at infinite loci, with uniform
902 rates of mutation, crossing over, and gene conversion (per bp, per generation) over a number of independent

903 chromosome arms. The dynamic we were interested in modeling requires direct accounting of pleiotropy between
904 different components of fitness, crossing over and gene conversion within inversions (as well as the possibility of
905 inversion fixation), and directly modeling male display and female choice, and no existing simulator provided this
906 functionality. These simulations were used to establish the potential for the model to occur in a population
907 generally, and not to compare against specific empirical observations. We have used literature estimates of
908 mutation rate and conversion rate of Zambian *D. melanogaster* for these simulations (Comeron et al. 2012; Huang
909 et al. 2016) but have no information on the joint distribution of effects of mutations on our model of survival and
910 display in *D. melanogaster*. So, we sampled from relatively broad but arbitrary combinations of mutation effects,
911 only assuming that there is likely some degree of antagonistic pleiotropy naturally present in genomic variation.

912 In SAIsim, a population is instantiated as a python object, and populated with individuals which are also
913 represented by python objects. These individuals may be instantiated using genomes specified by the user, or by
914 default carry no genomic variation. In each generation, as diagrammed in Figure 2, the simulator calculates a
915 survival probability for each individual as the product of the survival values of each variant they carry, which range
916 [0,1]. Here, multiplicative selection for survival prevents biologically unreasonable negative survival probabilities
917 from occurring. Male and female survival costs can each be included in a sex-specific manner, but only globally in
918 the simulation, not per mutation. For each offspring in the next generation, a mother is selected randomly,
919 weighted by individual survival probabilities, with replacement. To model the encountered males that might
920 become the father of an offspring individual from this female, a set number of males (*i.e.* the uniformly specified
921 encounter number, here 100 males unless otherwise specified) is randomly sampled from the full population of
922 males, weighted by their survival probabilities. The males each have a genetic display quality given as the sum of
923 the reproductive values of their alleles, each of which range [0,1]. A noise parameter, normally distributed around
924 zero with a standard deviation of one, is added to each male display value during each encounter. The male with
925 the highest noise-adjusted display quality among the encountered males is chosen to be the father. As offspring
926 are generated, crossover locations are sampled, and resampled in cases where the resultant gamete would be
927 aneuploid. While the simulator can allow recombination in both sexes, all simulations presented only generate
928 crossovers and gene conversion events for female gametes, in accordance with the biology of *D. melanogaster*.

929 Resampling aneuploids removes the fecundity cost that may exist in some taxa, but was chosen to reflect the lack
930 of observed fecundity effects for most paracentric inversions of *D. melanogaster*. Gene conversion is modeled as a
931 constant probability per heterozygous variant, with each event independently affecting just one locus, rather than
932 explicitly modeling a per-bp conversion rate and tract length (given that gene conversion tracts are generally short
933 and each would typically contain only a single modeled variant in our simulations). The dynamics of gene
934 conversion are expected to have separate effects on the assembly and the maintenance of arrangements. The
935 assembly of an arrangement is positively dependent on the population-scaled rate of recombination of alleles as
936 new polymorphisms start at low frequency and often not on their optimal haplotype. However, the indirect fitness
937 effect of the inversion is dependent on its effect on recombinant offspring, and therefore negatively depends on
938 the per-individual rate of conversion. Finally, inversions and new mutations are potentially added to the resultant
939 gamete, with inversions resampled when overlapping the coordinates of an inversion already present on the same
940 individual chromosome. SAIsim also allows the specification of mutation rates and the distribution of effect sizes of
941 new mutations.

942 The simulations presented in Figures 5, 6, and 7 rescale parameter values by a factor of 2,000 to emulate an
943 estimated ancestral $N_e$ of roughly $2 \times 10^6$ (Sprengelmeyer et al. 2020) while using a smaller simulated population of
944 1,000 individuals. The low crossover rate simulations model a 0.0218 cM region and the high crossover rate
945 simulations model 2.18 cM, which scale to 43.6 cM and 43.6 M female meiotic maps in the simulated populations.
946 Based on an average autosomal recombination rate in females of $2.18 \times 10^{-8}$ for *D. melanogaster* (Comeron et al.
947 2012), these represent 10kb and 1Mb regions. Gene conversion occurs at each heterozygous variant with a
948 probability of 0.1295, based on scaling $\gamma=1.25 \times 10^{-7}$ events/bp/female meiosis with a conversion track length of
949 518 bp (Comeron et al. 2012) to a per-single-nucleotide-variant estimate of $6.475 \times 10^{-5}$ conversions/bp/female
950 meiosis.

951 Rates of sexually antagonistic mutations and inversion mutations are empirically undetermined. Here, sexually
952 antagonistic mutations were modeled at a rate corresponding to one per thousand mutations, and so based on an
953 estimate of $5.21 \times 10^{-9}$ mutations/bp/gen on autosomes (Huang et al. 2016), and modeling 10kb and 1Mb regions,

954 these were scaled to $1.042 \times 10^{-4}$ and $1.042 \times 10^{-2}$ events per gamete respectively. Inversions were approximated
955 as arising once per hundred sexually antagonistic mutations, *i.e.* $1.042 \times 10^{-6}$ and $1.042 \times 10^{-4}$ events per gamete.
956 The simulations use mutations with uniformly distributed quality values in the range [0,1]; although distributions
957 skewed toward small effects may be more realistic, our uniform approach avoids the need to invoke a specific but
958 empirically unknown distribution.

959 SAIsim's object-oriented design allows specification of a population model by a new user familiar with the basics of
960 python scripting. A population object can be parameterized and populated in a few lines of python script, and then
961 parameters can be changed between periods of simulation, allowing bottlenecks and demographic changes as well
962 as changes in mutational parameters between generations. Replicate SAIsim simulations were performed via the
963 UW-Madison Center For High Throughput Computing (CHTC), which manages cycle servers using the HTCondor
964 distributed computing project (Thain et al. 2005; Erlandson and Theisen 2018).

## 965 4.2 Experimental Populations and Sequencing

966 To act as a high-inversion source population for the potential fathers in our experiment, we combined 15 males
967 and 15 females from each of 10 inbred strains derived from wild flies collected in Zambia, and allowed the
968 population to breed for two generations. These 10 strains were chosen to collectively carry the following common
969 inversions: *In(2L)t*, *In(2R)NS*, *In(3L)Ok, In(3R)K*. This high-inversion population may not and need not have been at
970 an equilibrium frequency for any particular inversion; instead it was simply designed to transmit each inversion to
971 offspring at meaningful frequency. We collected males from this second generation to use as the paternal pool in
972 each of our independent maternal line experiments, and these males were expected to carry a wide range of
973 inversion genotypes. Each batch of high-inversion males was crossed to an equal number of virgin females from a
974 single inbred Zambian strain (Figure 8; Table S1). Four such experiments were performed, each using one
975 inversion-free maternal strain, in clear plexiglass cages (19.3 by 19.3 by 30.0 cm) with a loose mesh covering on
976 one end for access by hand.

977 Embryos were collected by placing four 7.9 cM petri dishes of grape agar in a population cage for one day, then
978 counting and freezing the collected eggs. To obtain adult offspring, six bottles of standard *Drosophila* medium (see
979 below) were placed in a population cage for one day. Egg-laying for embryo and adult collections occurred on
980 alternating days over two weeks, after which all surviving parents were counted and frozen. From the first day in
981 which flies eclosed in a bottle, adults were collected at the end of each period of a day to two days and were
982 separated by sex. These cohorts were aged by weekly transfer to new bottles. The adult offspring were therefore
983 initially separated by sex and by the time after laying at which they eclosed, as well as the time in which they were
984 laid, with different early- or late-eclosing batches from a given cross later pooled before DNA extraction. In our
985 husbandry conditions, the earliest eclosion was 11 days after laying, but this was rare and a first eclosion at 12
986 days was more common among the bottles. Most flies took several days longer to develop, and we considered flies
987 collected in the first 6 days of a 10 day collection window to be the early-eclosing pool, in order to split the pool
988 sample sizes closer to evenly. The collected adult flies were then aged by sequentially transferring them between
989 bottles weekly for approximately 2.5 months, and then the adult offspring were sexed, counted, and collected. A
990 10-week aging window was chosen based on the time taken for a control set of four bottles of flies from the
991 paternal outbred population to reach 80% mortality, to reflect the survival and longevity stressors of the
992 laboratory bottle environment. All adult and embryo samples were frozen at -80°C before DNA extraction.

993 Since no inversions could be inherited through the mothers, inversion frequencies among successful male gametes
994 could be inferred from their pooled offspring. Therefore, the above sampling scheme allowed us to identify
995 inversion frequencies in the set of potential parents, in embryonic offspring, and in older adults, and so to assess
996 potentially selected changes in inversion frequency across male reproduction or between eggs and aged adults
997 (encompassing viability plus longevity). We chose to age the populations to include longevity as long term survival,
998 which both encompasses fitness challenges that generate mortality at earlier time points, and is relevant to the
999 need to survive during periods of unsuitable environments. Although a natural population from North Carolina was
1000 estimated to have an average generation time of about 24 days (including developmental time; Pool 2015), in
1001 many climates fly populations cycle seasonally between periods of rapid reproduction and survival for months as
1002 adults under reproductive diapause (Saunders et al. 1990; Ragland et al. 2019). While longevity is a key selective

1003 pressure underlying overwintering, the relationship between longevity in permissive lab conditions without
1004 diapause and in natural conditions under diapause is unclear (Schmidt et al. 2005; Flatt 2020), and our experiment
1005 represents just one of many possible ways to examine tradeoffs involving survival.

1006 Unless otherwise noted, all flies in the above experiments were kept in a lab space of 23°C with around a degree of
1007 temperature fluctuation and without an artificially controlled day/night cycle. Light exposure was dependent on
1008 the varying use of the space by laboratory workers but amounted to near constant exposure to at least a minimal
1009 level of lighting, with some variable light due to indirect lighting from adjacent rooms with exterior windows. All fly
1010 food used was prepared in batches of 4.5 L water, 500 mL cornmeal, 500 mL molasses, 200 mL yeast, 54 g agar, 20
1011 mL propionic acid, and 45 mL tegosept 10% (in 95% ethanol). The grape agar was prepared in batches of 500 mL
1012 from FlyStuff.com grape agar powder, and wet Red Star yeast was brushed over the agar plates before use to
1013 encourage laying.

## 1014 4.3 Sequencing and Data Preparation

1015 DNA for each replicate age group was extracted by a phenol-chloroform extraction protocol used previously in
1016 preparing DNA for the Drosophila Genome Nexus and *Drosophila* Population Genomics Project. We followed
1017 protocol 'BPC' used in DPGP phase 1 (presented in Langley et al. 2012), except with the addition of proteinase-K
1018 during homogenization of adult flies. These protocols are ultimately based on the common quick preparation
1019 protocol presented in *Drosophila Protocols* (Huang et al. 2009). Inversion genotyping was performed by amplicon
1020 sequencing, where amplicons associated with each inversion were designed to contain at least one single
1021 nucleotide polymorphism in the inversion interior that is fixed for alternate alleles between the inversion
1022 karyotypes, while still sharing conserved primer regions across karyotypes to prevent primer bias. Genetic
1023 variation in these inversions was assessed using 197 previously sequenced and inversion-called haploid genomes
1024 from the same Zambian *D. melanogaster* population studied here, available from the *Drosophila* Genome Nexus
1025 (Lack et al. 2016a). The amplicon region identification and primer selection was performed with a custom script,
1026 which calls Primer3 version 2.4.0 for primer candidate generation (Koressaar and Remm 2007; Untergasser et al.

1027 2012; Koressaar et al. 2018). Primers thus selected were ordered from IDT as oligos ligated to Illumina stubby
1028 adaptors (Table S2). Illumina sequencing libraries were then prepared for each sample following the Illumina
1029 metagenomic amplicon sequencing protocol (Illumina 2013), with 50 ng/uL diluted DNA samples. and sequenced
1030 using 150bp paired end reads on an Illumina NovaSeq to at least a depth of 20,000 reads. Depth was selected to
1031 ensure high likelihood of detecting a 0.025 change in inversion frequency from an initial frequency of 0.10, based
1032 on simulated resampling of the collected flies, DNA extraction, and sequencing.

1033 Sequence preparation and analysis were performed using custom pipeline scripts. Read trimming, alignment, and
1034 QC were performed using bwa version 0.7.17-r1188 (Li 2013), GATK version 3.4-46 (Van der Auwera and O'Connor
1035 2020), and Picard version 1.79 (Broad Institute 2019). Reads were trimmed at the ends to remove segments with
1036 base quality less than 20, and read pairs were filtered to retain only those pairs where both reads aligned with
1037 mapping quality >20. SNP identity at the sites of interest was called from those remaining reads with base quality
1038 >20 at the SNP of interest. Due to the presence of chimeric reads in the amplicon library sequencing, we chose to
1039 use a single discriminant SNP to determine each inversion's frequency in the read pool, instead of incorporating
1040 haplotype or additional SNP information.

## 1041 4.4 Statistical Analysis of Inversion Frequencies

1042 Before analyzing the significance of frequency change reversals across three cohorts (which is described below),
1043 we assessed the significance of the difference in inversion frequency between just two sampled life stages in the
1044 same experiment. For each such comparison, we first calculated a chi-square test statistic (as a convenient metric
1045 to use when comparing empirical data with simulated experimental replicates, rather than as a formal test) from
1046 the two-by-two table of inverted and non-inverted allele counts at two life stages. Allele counts were estimated by
1047 multiplying the inversion-specific SNP allele frequency in the sequencing reads by the allelic sample size based on
1048 the number of flies present in the sample. We then generated an approximate distribution of the chi-square
1049 statistic when selection is absent (our null hypothesis), using a resampling approach (Figure 8), as described in the

following paragraph. Comparing the observed statistic to our modeled neutral distribution of test-statistic values allowed us to calculate a p-value for this inversion's frequency change between these two life stages.

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 S7 and S8) to account for sampling variance between replicates. For the paternal samples, we needed to account for pre-sequencing mortality as a mean of 9.83% of males died before collection (a range of 5.11 to 15.77%, Table S1). 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 S9) 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 S7 and S8) 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.

1076 Finally, each offspring allele sample was resampled using the same size 316 binomial sampling to account for the
1077 variation introduced during this library's preparation, and binomial sampled again with the read count size to
1078 account for sampling reads from the library, thus yielding resampled offspring read counts that accounted for the
1079 same three factors as for paternal counts.

1080 For a pairwise comparison of inversion frequency between two samples, we calculated the same allele count chi-
1081 square statistic as described above for the empirical data for each resampled set of allele counts, generating the
1082 full distribution of neutral chi-square test statistics from which we could identify extreme value cutoffs. We then
1083 obtained a p-value from the proportion of resampled replicates in which the chi-square statistic was greater than
1084 or equal to the empirical value. If the test was directional, we took the proportion of resampled replicates with
1085 both a chi-square statistic greater than or equal to the empirical value and a frequency difference in the expected
1086 direction.

1087 In order to identify a frequency change reversal between life history stages that could reflect a fitness tradeoff, we
1088 tested for the joint significance of two opposing frequency changes using three cohorts: specifically a frequency
1089 change in one direction between paternal and embryo samples, followed by a frequency change in the opposite
1090 direction between embryo and adult offspring samples. These changes both depend on the shared embryonic
1091 cohort in the test, and so are not independent. Therefore, we evaluated the two inversion frequency changes
1092 jointly from a set of resampled neutral data sets in which the three cohorts arose from the same resampling
1093 process. Here we considered how likely it was that a sample of the three cohorts from the associated estimated
1094 neutral model had chi-square values for both the comparison of the fathers to embryos and the embryos to adult
1095 offspring that were each equivalent or more extreme in the test direction than the observed data. For example,
1096 when testing the significance of an observed paternal-embryo increase and then embryo-adult decrease in
1097 frequency, our p-value would be defined by the proportion of resampled replicates in which the frequency
1098 changes were in the tested directions of increase then decrease, and their chi-square values were both more
1099 extreme than the empirical values. The potential for empirical frequency change reversals to occur in either
1100 direction was accounted for at the multiple test correction step as indicated below.

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 two potential tradeoff directions and the four inversions studied, 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 BioProject ID PRJNA1213821. The simulation program and other analysis scripts can be found at https://github.com/csmcal/.

## Acknowledgements

1123 Particular thanks to Shimeng Gao, who assisted with *Drosophila* husbandry and collection of adult flies during the
1124 within-generation experiment. We also thank members of the Pool lab and three anonymous reviewers for helpful
1125 comments on earlier versions of this manuscript.

1126 This research was supported by National Science Foundation Graduate Research Fellowship (DGE-1747503 to
1127 CSM), and by the National Institutes of Health (grants R35 GM136306 to JEP, T32 GM007133, and T32 HG002760).

1128 This research was performed using the compute resources and assistance of the UW-Madison Center for High
1129 Throughput Computing (CHTC) in the Department of Computer Sciences. The CHTC is supported by UW-Madison,
1130 the Advanced Computing Initiative, the Wisconsin Alumni Research Foundation, the Wisconsin Institutes for
1131 Discovery, and the National Science Foundation, and is an active member of the Open Science Grid, which is
1132 supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

1133

1134 # Works Cited


1135 Altenberg L, Feldman MW. 1987. Selection, generalized transmission and the evolution of modifier genes. I. The
1136 reduction principle. *Genetics* 117:559–572.

1137 Altenberg L, Liberman U, Feldman MW. 2017. Unified reduction principle for the evolution of mutation, migration,
1138 and recombination. *Proc. Natl. Acad. Sci. U.S.A.* 114:E2392–E2400.

1139 Audet T, Krol J, Pelletier K, Stewart AD, Dworkin I. 2024. Sexually discordant selection is associated with trait
1140 specific morphological changes and a complex genomic response. *Evolution* 78:1426–1440.

1141 Aulard S, David JR, Lemeunier F. 2002. Chromosomal inversion polymorphism in Afrotropical populations of
1142 *Drosophila melanogaster*. *Genet. Res.* 79:49–63.

1143 Babcock CS, Anderson WW. 1996. Molecular evolution of the Sex-Ratio inversion complex in *Drosophila*
1144 *pseudoobscura*: analysis of the Esterase-5 gene region. *Mol. Biol. Evol.* 13:297–308.

1145 Bastide H, Ogereau D, Montchamp-Moreau C, Gérard PR. 2022. The fate of a suppressed X-linked meiotic driver:
1146 experimental evolution in *Drosophila simulans*. *Chromosome Res.* 30:141–150.

1147 Bastide H, Yassin A, Johanning EJ, Pool JE. 2014. Pigmentation in *Drosophila melanogaster* reaches its maximum in
1148 Ethiopia and correlates most strongly with ultra-violet radiation in sub-Saharan Africa. *BMC Evol. Biol.*
1149 14:179.

1150 Bateman AJ. 1948. Intra-sexual selection in *Drosophila*. *Heredity* 2:349–368.

1151 Behrman EL, Watson SS, O'Brien KR, Heschel MS, Schmidt PS. 2015. Seasonal variation in life history traits in two
1152 *Drosophila* species. *J. Evol. Biol.* 28:1691–1704.

1153 Berdan EL, Barton NH, Butlin R, Charlesworth B, Faria R, Fragata I, Gilbert KJ, Jay P, Kapun M, Lotterhos KE, et al.
1154 2023. How chromosomal inversions reorient the evolutionary process. J. Evol. Biol. 36:1761–1782.

1155 Betrán E, Santos M, Ruiz A. 1998. Antagonistic Pleiotropic Effect of Second-Chromosome Inversions on Body Size
1156 and Early Life-History Traits in *Drosophila buzzatii*. *Evolution* 52:144–154.

1157 Blackmon H, Brandvain Y. 2017. Long-Term Fragility of Y Chromosomes Is Dominated by Short-Term Resolution of
1158 Sexual Antagonism. *Genetics* 207:1621–1629.

1159 Blows MW, Brooks R. 2003. Measuring Nonlinear Selection. *Am. Nat.* 162:815–820.

1160 Broad Institute. 2019. Picard Toolkit. Available from: http://broadinstitute.github.io/picard/

1161 Charlesworth B. 1974. Inversion polymorphism in a two-locus genetic system. *Genet. Res.* 23:259–280.

1162 Charlesworth B. 1997. The evolution of lethals in the t-haplotype system of the mouse. *Proc. Royal Soc. B* 258:101–
1163 107.

1164 Charlesworth B. 2024. The fitness consequences of genetic divergence between polymorphic gene arrangements.
1165 Genetics 226:iyad218.

1166 Charlesworth B, Barton NH. 2018. The Spread of an Inversion with Migration and Selection. *Genetics* 208:377–382.

1167 Charlesworth B, Charlesworth D. 1973. Selection of new inversions in multi-locus genetic systems. *Genet. Res.*
1168 21:167–183.

1169 Charlesworth B, Charlesworth D. 2010. Elements of Evolutionary Genetics. W. H. Freeman

1170 Charlesworth B, Hughes KA. 2000. The maintenance of genetic variation in life-history traits. In: Singh RS, Krimbas
1171 CB, editors. Evolutionary Genetics: from Molecules to Morphology. Cambridge, U.K.; New York, NY, USA:
1172 Cambridge University Press. p. 369–392.

1173 Charlesworth D, Charlesworth B. 1980. Sex differences in fitness and selection for centric fusions between sex-
1174 chromosomes and autosomes. *Genet. Res.* 35:205–214.

1175 Chen KH-C, Chatla K, Bachtrog D. 2024. Single-cell RNA-seq of *Drosophila miranda* testis reveals the evolution and
1176 trajectory of germline sex chromosome regulation. *PLOS Genet.* 22(4): e3002605.

1177 Cheng C, Kirkpatrick M. 2016. Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies. *PLOS*
1178 *Genet.* 12:e1006170.

1179 Clarke CA, Sheppard PM. 1960. Super-genes and mimicry. *Heredity* 14:175–185.

1180 Comeron JM, Ratnappan R, Bailin S. 2012. The Many Landscapes of Recombination in *Drosophila melanogaster*.
1181 *PLOS Genet.* 8:e1002905.

1182 Connallon T, Chenoweth SF. 2019. Dominance reversals and the maintenance of genetic variation for fitness. *PLOS*
1183 *Biol.* 17:e3000118.

1184 Connallon T, Clark AG. 2011. The resolution of sexual antagonism by gene duplication. *Genetics* 187:919–937.

1185 Connallon T, Clark AG. 2014. Balancing Selection in Species with Separate Sexes: Insights from Fisher's Geometric
1186 Model. *Genetics* 197:991–1006.

1187 Cook LM. 2013. Selection and disequilibrium in Cepaea nemoralis. *Biol. J. Linn. Soc.* 108:484–493.

1188 Corbett-Detig RB. 2016. Selection on Inversion Breakpoints Favors Proximity to Pairing Sensitive Sites in Drosophila
1189 melanogaster. *Genetics* 204:259–265.

1190 Corbett-Detig RB, Hartl DL. 2012. Population Genomics of Inversion Polymorphisms in *Drosophila melanogaster*.
1191 *PLOS Genet.* 8:e1003056.

1192 Courret C, Chang C-H, Wei KH-C, Montchamp-Moreau C, Larracuente AM. 2019. Meiotic drive mechanisms: lessons
1193 from Drosophila. *Proc. Royal Soc. B* 286:20191430.

1194 Coyne JA, Meyers W, Crittenden AP, Sniegowski P. 1993. The Fertility Effects of Pericentric Inversions in *Drosophila*
1195 *melanogaster*. *Genetics* 134:487–496.

1196 Cox RM, Calsbeek R. 2009. Sexually Antagonistic Selection, Sexual Dimorphism, and the Resolution of Intralocus
1197 Sexual Conflict. *Am. Nat.* 173:176–187.

1198 da Silva Ribeiro T, Galván JA, Pool JE. 2022. Maximum SNP FST Outperforms Full-Window Statistics for Detecting
1199 Soft Sweeps in Local Adaptation. *Genome Biol. Evol.* 14:evac143.

1200 Dickerson GE. 1955. Genetic slippage in response to selection for multiple objectives. *Cold Spring Harbor Symposia*
1201 *on Quantitative Biology* 20:213–224.

1202 Dobzhansky TG. 1943. Genetics of Natural Populations IX. Temporal Changes in the Composition of Populations of
1203 *Drosophila pseudoobscura*. *Genetics* 28:162–186.

1204 Dobzhansky TG. 1949. Observations and Experiments on Natural Selection in *Drosophila*. *Hereditas* 35:210–224.

1205 Dobzhansky TG. 1950. Genetics of natural populations. XIX. Origin of heterosis through natural selection in
1206 populations of *Drosophila pseudoobscura*. *Genetics* 35:288–302.

1207 Dobzhansky TG. 1970. Genetics of the Evolutionary Process. Columbia University Press

1208 van Doorn GS, Kirkpatrick M. 2007. Turnover of sex chromosomes induced by sexual conflict. *Nature* 449:909–912.

1209 Erlandson E, Theisen T. 2018. htcondor: HTCondor source repository, formerly the Condor Project. Available from:
1210 https://github.com/htcondor/htcondor

1211 Fang S, Takahashi A, Wu C-I. 2002. A Mutation in the Promoter of *Desaturase 2* Is Correlated With Sexual Isolation
1212 Between *Drosophila* Behavioral Races. *Genetics* 162:781-784.

1213 Faria R, Johannesson K, Butlin RK, Westram AM. 2019. Evolving Inversions. *Trends Ecol. Evol.* 34:239–248.

1214 Feldman MW. 1972. Selection for Linkage Modification: I. Random Mating Populations. *Theor. Pop. Biol.* 3:324–
1215 346.

1216 Feldman MW, Balkau B. 1973. Selection for Linkage Modification: II. A Recombination Balance for Neutral
1217 Modifiers. *Genetics* 74:713–726.

1218 Feldman MW, Christiansen FB, Brooks LD. 1980. Evolution of recombination in a constant environment. *Proc. Natl.*
1219 *Acad. Sci. U.S.A.* 77:4838–4841.

1220 Flatt T. 2020. Life-History Evolution and the Genetics of Fitness Components in *Drosophila melanogaster*. *Genetics*
1221 214:3–48.

1222 Frydenberg O. 1963. Population Studies of a Lethal Mutant in *Drosophila melanogaster*. *Hereditas* 50:89–116.

1223 Gilbert KJ, Pouyet F, Excoffier L, Peischl S. 2020. Transition from Background Selection to Associative
1224 Overdominance Promotes Diversity in Regions of Low Recombination. *Curr. Biol.* 30:101-107.e3.

1225 Gillespie JH. 1998. Population Genetics: A Concise Guide. Johns Hopkins University Press

1226 Giraldo-Deck LM, Loveland JL, Goymann W, Tschirren B, Burke T, Kempenaers B, Lank DB, Küpper C. 2022.
1227 Intralocus conflicts associated with a supergene. *Nat. Commun.* 13:1384.

1228 Glaser-Schmitt A, Ramnarine TJS, Parsch J. 2023. Rapid evolutionary change, constraints and the maintenance of
1229 polymorphism in natural populations of *Drosophila melanogaster*. *Mol. Ecol.*

1230 Gong WJ, McKim KS, Hawley RS. 2005. All Paired Up with No Place to Go: Pairing, Synapsis, and DSB Formation in a
1231 Balancer Heterozygote. *PLOS Genet.* 1:e67.

1232 Gowaty PA, Kim Y-K, Anderson WW. 2012. No evidence of sexual selection in a repetition of Bateman's classic
1233 study of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 109:11740–11745.

1234 Grieshop K, Arnqvist G. 2018. Sex-specific dominance reversal of genetic variation for fitness. *PLOS Biol.*
1235 16:e2006810.

1236 Grillet M, Everaerts C, Houot B, Ritchie MG, Cobb M, Ferveur J-F. 2012. Incipient speciation in *Drosophila*
1237 *melanogaster* involves chemical signals. *Sci. Rep.* 2:224.

1238 Hazel LN. 1943. The Genetic Basis for Constructing Selection Indexes. *Genetics* 28:476–490.

1239 Hearn KE, Koch EL, Stankowski S, Butlin RK, Faria R, Johannesson K, Westram AM. 2022. Differing associations
1240 between sex determination and sex-linked inversions in two ecotypes of Littorina saxatilis. *Evolution*
1241 *Letters* 6:358–374.

1242 Hoffmann AA, Rieseberg LH. 2008. Revisiting the Impact of Inversions in Evolution: From Population Genetic
1243 Markers to Drivers of Adaptive Shifts and Speciation? *Annu. Rev. Ecol. Evol. Syst.* 39:21–42.

1244 Hoquet T, Bridges WC, Gowaty PA. 2020. Bateman's Data: Inconsistent with "Bateman's Principles." *Ecol. Evol.*
1245 10:10325–10342.

1246 Huang AM, Rehm EJ, Rubin GM. 2009. Quick Preparation of Genomic DNA from Drosophila. *Cold Spring Harb.*
1247 *Protoc.* 2009:pdb.prot5198.

1248 Huang W, Lyman RF, Lyman RA, Carbone MA, Harbison ST, Magwire MM, Mackay TF. 2016. Spontaneous
1249 mutations and the origin and maintenance of quantitative genetic variation. Przeworski M, editor. *eLife*
1250 5:e14625.

1251 Illumina. 2013. 16S metagenomic sequencing library preparation guide. Available from:
1252 https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-
1253 metagenomic-library-prep-guide-15044223-b.pdf

1254 Ingleby FC, Flis I, Morrow EH. 2015. Sex-Biased Gene Expression and Sexual Conflict throughout Development. *Cold*
1255 *Spring Harb. Perspect. Biol.* 7:a017632.

1256 Innocenti P, Morrow EH. 2010. The Sexually Antagonistic Genes of Drosophila melanogaster. *PLOS Biol.*
1257 8:e1000335.

1258 Inoue Y. 1979. The Fate Of Polymorphic Inversions Of *Drosophila melanogaster* Transferred To Laboratory
1259 Conditions. *Jpn. J. Genet.* 54:83–96.

1260 John A, Vinayan K, Varghese J. 2016. Achiasmy: Male Fruit Flies Are Not Ready to Mix. *Front. Cell Dev. Biol.* 4:75.

1261 Jones AG, Arguello JR, Arnold SJ. 2002. Validation of Bateman's principles: a genetic study of sexual selection and
1262 mating patterns in the rough-skinned newt. *Proc. R. Soc. B Biol. Sci.* 269:2533–2539.

1263 Kapun M, Fabian DK, Goudet J, Flatt T. 2016. Genomic Evidence for Adaptive Inversion Clines in *Drosophila*
1264 *melanogaster*. *Mol. Biol. Evol.* 33:1317–1336.

1265 Kapun M, Flatt T. 2019. The adaptive significance of chromosomal inversion polymorphisms in *Drosophila*
1266 *melanogaster*. *Mol. Ecol.* 28:1263–1282.

1267 Kapun M, Mitchell ED, Kawecki TJ, Schmidt P, Flatt T. 2023. An Ancestral Balanced Inversion Polymorphism Confers
1268 Global Adaptation. *Mol. Biol. Evol.* 40:msad118.

1269 Kapun M, van Schalkwyk H, McAllister B, Flatt T, Schlötterer C. 2014. Inference of chromosomal inversion dynamics
1270 from Pool-Seq data in natural and laboratory populations of *Drosophila melanogaster*. *Mol. Ecol.*
1271 23:1813–1827.

1272 Kaufmann P, Howie JM, Immonen E. 2023. Sexually antagonistic selection maintains genetic variance when sexual
1273 dimorphism evolves. *Proc. Royal Soc. B* 290:20222484.

1274 Kirkpatrick M, Barton N. 2006. Chromosome Inversions, Local Adaptation and Speciation. *Genetics* 173:419–434.

1275 Kopp M, Hermisson J. 2006. The evolution of genetic architecture under frequency-dependent disruptive selection.
1276 *Evolution* 60:1537–1550.

1277 Koury SA. 2023. Predicting recombination suppression outside chromosomal inversions in *Drosophila*
1278 *melanogaster* using crossover interference theory. *Heredity* 130:196–208.

1279 Koressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M. 2018. Primer3_masker: integrating masking
1280 of template sequence with primer design software. *Bioinforma. Oxf. Engl.* 34:1937–1938.

1281 Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinforma. Oxf.*
1282 *Engl.* 23:1289–1291.

1283 Krimbas CB, Powell JR. 1992. Introduction. In: Krimbas CB, Powell JR, editors. *Drosophila* Inversion Polymorphism.
1284 Boca Raton, FL: CRC Press. p. 1–52.

1285 Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR. 2014. doublesex
1286 is a mimicry supergene. *Nature* 507:229–232.

1287 Küpper C, Stocks M, Risse JE, Remedios N dos, Farrell LL, McRae SB, Morgan TC, Karlionova N, Pinchuk P, Verkuil YI,
1288 et al. 2016. A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.*
1289 48:79–83.

1290 Lack JB, Lange JD, Tang AD, Corbett-Detig RB, Pool JE. 2016. A Thousand Fly Genomes: An Expanded *Drosophila*
1291 Genome Nexus. *Mol. Biol. Evol.* 33:3308–3313.

1292 Lack JB, Yassin A, Sprengelmeyer QD, Johanning EJ, David JR, Pool JE. 2016. Life history evolution and cellular
1293 mechanisms associated with increased size in high-altitude *Drosophila*. *Ecol. Evol.* 6:5893–5906.

1294 Lange JD, Bastide H, Lack JB, Pool JE. 2022. A Population Genomic Assessment of Three Decades of Evolution in a
1295 Natural *Drosophila* Population. *Mol. Biol. Evol.* 39:msab368.

1296 Langley CH, Stevens K, Cardeno C, Lee YCG, Schrider DR, Pool JE, Langley SA, Suarez C, Corbett-Detig RB,
1297 Kolaczkowski B, et al. 2012. Genomic Variation in Natural Populations of *Drosophila melanogaster*.
1298 *Genetics* 192:533–598.

1299 Lemeunier F, Aulard S (1992). Inversion polymorphism in *Drosophila melanogaster*. In: Krimbas CB, Powell JR
1300 (eds) *Drosophila Inversion Polymorphism*. CRC Press: Boca Raton, FL pp 339–405.

1301 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Available from:
1302 http://arxiv.org/abs/1303.3997

1303 Libbrecht R, Kronauer DJC. 2014. Convergent Evolution: The Genetics of Queen Number in Ants. *Curr. Biol.*
1304 24:R1083–R1085.

1305 Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al.
1306 2015. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer
1307 Interactions. *Cell* 161:1012–1025.

1308 Machado HE, Bergland AO, Taylor R, Tilk S, Behrman E, Dyer K, Fabian DK, Flatt T, González J, Karasov TL, et al.
1309 2021. Broad geographic sampling reveals the shared basis and environmental correlates of seasonal
1310 adaptation in Drosophila. *eLife* 10:e67577.

1311 McBroome J, Liang D, Corbett-Detig RB. 2020. Fine-Scale Position Effects Shape the Distribution of Inversion
1312 Breakpoints in *Drosophila melanogaster*. *Genome Biol. Evol.* 12:1378–1391.

1313 Mérot C, Llaurens V, Normandeau E, Bernatchez L, Wellenreuther M. 2020. Balancing selection via life-history
1314 trade-offs maintains an inversion polymorphism in a seaweed fly. *Nat. Commun.* 11:670.

1315 Mettler E, Voelker RA, Mukai T. 1977. Inversion Clines in Populations of *Drosophila melanogaster*. *Genetics* 87:8.

1316 Morgan TH. 1910. Sex Limited Inheritance in *Drosophila*. *Science* 32:120–122.

1317 Mroczek RJ, Melo JR, Luce AC, Hiatt EN, Dawe RK. 2006. The Maize Ab10 Meiotic Drive System Maps to
1318 Supernumerary Sequences in a Large Complex Haplotype. *Genetics* 174:145–154.

1319 Mukai T, Yamaguchi O. 1974. The genetic structure of natural populations of *Drosophila melanogaster*. XI. Genetic
1320 variability in a local population. *Genetics* 76:339–366.

1321 Munasinghe M, Brandvain Y. 2023. Together Inbreeding and Reproductive Compensation Favor Lethal t-
1322 Haplotypes. *J. Hered.*, Accepted.

1323 Nassar R, Muhs HJ, Cook RD. 1973. Frequency-dependent selection at the Payne inversion in *Drosophila*
1324 *melanogaster. Evolution* 27:558-564.

1325 O'Donald P. 1973. Frequency-dependent sexual selection as a result of variations in fitness at breeding time.
1326 *Heredity* 30:351–368.

1327 O'Donald P, Majerus MEN, Clarke BC, Partridge L, Robertson A, Clarke BC, Partridge L. 1997. Frequency-dependent
1328 sexual selection. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 319:571–586.

1329 Ohta T, Kimura M. 1970. Development of associative overdominance through linkage disequilibrium in finite
1330 populations. *Genet. Res.* 16:165–177.

1331 Otto SP, Lenormand T. 2002. Resolving the paradox of sex and recombination: Evolution of sex. *Nat. Rev. Genet.*
1332 3:252–261.

1333 Otwinowski J, McCandlish DM, Plotkin JB. 2018. Inferring the shape of global epistasis. *Proc. Natl. Acad. Sci. U.S.A.*
1334 115:E7550–E7558.

1335 Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadía-Cardoso A, Anderson EC, Rundio DE, Williams TH,
1336 Naish KA, et al. 2019. Sex-dependent dominance maintains migration supergene in rainbow trout. *Nat.*
1337 *Ecol. Evol.* 3:1731–1742.

1338 Pei Y, Forstmeier W, Knief U, Kempenaers B. 2023. Weak antagonistic fitness effects can maintain an inversion
1339 polymorphism. *Mol. Ecol.* 32:3575–3585.

1340 Pennell MW, Kirkpatrick M, Otto SP, Vamosi JC, Peichel CL, Valenzuela N, Kitano J. 2015. Y Fuse? Sex Chromosome
1341 Fusions in Fishes and Reptiles. *PLOS Genet.* 11:e1005237.

1342 Pool JE. 2015. The Mosaic Ancestry of the *Drosophila* Genetic Reference Panel and the *D. melanogaster* Reference
1343 Genome Reveals a Network of Epistatic Fitness Interactions. *Mol. Biol. Evol.* 32:3236–3251.

1344 Pool JE, Braun DT, Lack JB. 2017. Parallel Evolution of Cold Tolerance within *Drosophila melanogaster*. *Mol. Biol.*
1345 *Evol.* 34:349–360.

1346 Poul YL, Whibley A, Chouteau M, Prunier F, Llaurens V, Joron M. 2014. Evolution of dominance mechanisms at a
1347 butterfly mimicry supergene. *Nat. Commun.* 5:5644.

1348 Ragland GJ, Armbruster PA, Meuti ME. 2019. Evolutionary and functional genetics of insect diapause: a call for
1349 greater integration. *Curr. Op. Insect Sci.* 36:74–81.

1350 Reinhardt JA, Kolaczkowski B, Jones CD, Begun DJ, Kern AD. 2014. Parallel Geographic Variation in *Drosophila*
1351 *melanogaster*. *Genetics* 197:361–373.

1352 Reinhart M, Carney T, Clark AG, Fiumera AC. 2015. Characterizing Male–Female Interactions Using Natural Genetic
1353 Variation in Drosophila melanogaster. *J. Hered.* 106:67–79.

1354 Rest JS, Morales CM, Waldron JB, Opulente DA, Fisher J, Moon S, Bullaughey K, Carey LB, Dedousis D. 2013.
1355 Nonlinear Fitness Consequences of Variation in Expression Level of a Eukaryotic Gene. *Mol. Biol. Evol.*
1356 30:448–456.

1357 Rice WR. 1992. Sexually antagonistic genes: experimental evidence. *Science* 256:1436–1439.

1358 Rice WR. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature*
1359 381:232–234.

1360 Rose MR. 1982. Antagonistic pleiotropy, dominance, and genetic variation. *Heredity* 48:63–78.

1361 Rowe L, Chenoweth SF, Agrawal AF. 2018. The Genomics of Sexual Conflict. *Am. Nat.* 192:274–286.

1362 Ruzicka F, Hill MS, Pennell TM, Flis I, Ingleby FC, Mott R, Fowler K, Morrow EH, Reuter M. 2019. Genome-wide
1363 sexually antagonistic variants reveal long-standing constraints on sexual dimorphism in fruit flies. *PLOS*
1364 *Biol.* 17:e3000244.

1365 Said I, Byrne A, Serrano V, Cardeno C, Vollmers C, Corbett-Detig RB. 2018. Linked genetic variation and not genome
1366 structure causes widespread differential expression associated with chromosomal inversions. *Proc. Natl.*
1367 *Acad. Sci. U.S.A.* 115:5492–5497.

1368 Saunders DS, Richard DS, Applebaum SW, Ma M, Gilbert LI. 1990. Photoperiodic diapause in *Drosophila*
1369 *melanogaster* involves a block to the juvenile hormone regulation of ovarian maturation. *Gen. Comp.*
1370 *Endocrinol.* 79:174–184.

1371 Saunders PA, Neuenschwander S, Perrin N. 2019. Impact of deleterious mutations, sexually antagonistic selection,
1372 and mode of recombination suppression on transitions between male and female heterogamety. *Heredity*
1373 123:419–428.

1374 Schaeffer SW, Goetting-Minesky MP, Kovacevic M, Peoples JR, Graybill JL, Miller JM, Kim K, Nelson JG, Anderson
1375 WW. 2003. Evolutionary genomics of inversions in *Drosophila pseudoobscura*: Evidence for epistasis. *Proc.*
1376 *Natl. Acad. Sci. U.S.A.* 100:8319–8324.

1377 Schmidt PS, Paaby AB, Heschel MS. 2005. Genetic Variance For Diapause Expression And Associated Life Histories
1378 In *Drosophila melanogaster*. *Evolution* 59:2616–2625.

1379 Sprengelmeyer QD, Mansourian S, Lange JD, Matute DR, Cooper BS, Jirle EV, Stensmyr MC, Pool JE. 2020.
1380 Recurrent Collection of *Drosophila melanogaster* from Wild African Environments and Genomic Insights
1381 into Species History. *Mol. Biol. Evol.* 37:627–638.

1382 Stearns FW. 2010. One Hundred Years of Pleiotropy: A Retrospective. *Genetics* 186:767–773.

1383 Sturtevant AH. 1921. A Case of Rearrangement of Genes in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 7:235–237.

1384 Sturtevant AH, Beadle GW. 1936. The Relations of Inversions in the X Chromosome of *Drosophila melanogaster* to
1385 Crossing Over and Disjunction. *Genetics* 21:51.

1386 Sturtevant AH, Mather K. 1938. The Interrelations of Inversions, Heterosis and Recombination. *Am. Nat.* 72:447–
1387 452.

1388 Surmacki A, Ożarowska-Nowicka A, Rosin ZM. 2013. Color polymorphism in a land snail Cepaea nemoralis
1389 (Pulmonata: Helicidae) as viewed by potential avian predators. *Naturwissenschaften* 100:533–540.

1390 Sved JA. 1968. The Stability of Linked Systems of Loci with a Small Population Size. *Genetics* 59:543–563.

1391 Tatarenkov A, Healey CIM, Grether GF, Avise JC. 2008. Pronounced reproductive skew in a natural population of
1392 green swordtails, *Xiphophorus helleri*. *Mol. Ecol.* 17:4522–4534.

1393 Thain D, Tannenbaum T, Livny M. 2005. Distributed computing in practice: the Condor experience. *Concurr.*
1394 *Comput. Pract. Exp.* 17:323–356.

1395 Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities
1396 and interfaces. *Nucleic Acids Res.* 40:e115.

1397 Van der Auwera GA, O'Connor BD. 2020. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly
1398 Media, Incorporated.

1399 Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang Y-C, Shoemaker D, Keller L. 2013. A Y-like social
1400 chromosome causes alternative colony organization in fire ants. *Nature* 493:664–668.

1401 Wellenreuther M, Bernatchez L. 2018. Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends Ecol. Evol.*
1402 33:427–440.

1403 White MJD. 1973. Animal Cytology and Evolution. 3rd ed. Cambridge University Press. Available from:
1404 https://wellcomecollection.org/works/meqhhay5

1405 Wittmann MJ, Bergland AO, Feldman MW, Schmidt PS, Petrov DA. 2017. Seasonally fluctuating selection can
1406 maintain polymorphism at many loci via segregation lift. *Proc. Natl. Acad. Sci. U.S.A.* 114:E9932–E9941.

1407 Wright AE, Dean R, Zimmer F, Mank JE. 2016. How to make a sex chromosome. *Nat. Commun.* 7:1–8.

1408 Yamamoto A, Huang W, Carbone MA, Anholt RRH, Mackay TFC. The Genetic Basis of Incipient Sexual Isolation in
1409 *Drosophila melanogaster*. *Proc. Royal Soc. B* 291:20240672.

1410 Yassin A, Bastide H, Chung H, Veuille M, David JR, Pool JE. 2016. Ancient balancing selection at tan underlies female
1411 colour dimorphism in Drosophila erecta. *Nat. Commun.* 7:10400.

1412 Zajitschek F, Connallon T. 2018. Antagonistic pleiotropy in species with separate sexes, and the maintenance of
1413 genetic variation in life-history traits and fitness. *Evolution* 72:1306–1316.

1414 Zhao L, Charlesworth B. 2016. Resolving the Conflict Between Associative Overdominance and Background
1415 Selection. *Genetics* 203:1315–1334.

1416