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 also that a potential effect of inversions in heterokaryotypic individuals is the deleterious generation of aneuploid gametes, it seems reasonable to expect inversion rearrangements to confer either deleterious or neutral fitness consequences in the absence of association with other variants. However, when a new inversion mutates, it samples a single haplotype from a population, which is likely to have some non-neutral fitness effect, even if mild. Greater natural variation in fitness within a population increases the likelihood of sampling haplotypes with exceptionally deleterious or beneficial associations of alleles, even when there are no other effects on fitness (Berdan et al. 2023).

In contrast to the above predictions, a considerable number of well-studied inversion variants are maintained at an intermediate frequency within a population, presumably due to the indirect effect inversions have on reducing recombination and maintaining fit haplotypes that are themselves under selection (reviewed in Berdan et al. 2023). The evolutionary forces maintaining both haplotypes are thought to be quite varied between populations and arrangements, but a common first hypothesis for a frequently sampled inversion is that local adaptation in one region or environment is balanced with migration from another (Kirkpatrick and Barton 2006; Charlesworth and Barton 2018). This association may be maintained more strongly when the alleles have beneficial epistatic interactions, and assumed beneficial epistasis between loci is often paired with another mechanism of balancing selection operating at a locus to explain inversion (Dobzhansky 1949, 1950, 1970; Schaeffer et al. 2003; Hoffmann and Rieseberg 2008). However, unless migration rates are very high, local adaptation models predict inversions to be at high frequency in their favored geographic range but at low frequency elsewhere, and many inversions do not have sampled locations at which they reach very high frequency.

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

Balanced polymorphic loci may be maintained in other ways, under various models and selective pressures. For example, multiple overdominant loci that are each most beneficial as heterozygotes may benefit from linkage maintaining complementary haplotypes (Sved 1968; Charlesworth and Charlesworth 1973; Charlesworth 1974). Several recent studies have proposed that functionally overdominant loci may be more common than previously considered. Specifically, the dominance relation between alleles at a locus may change between different sexes,

seasons, life history stages, or other trait or fitness challenge contexts, generally termed 'reversal of dominance', which in some cases can lead heterozygotes to have greatest average fitness (Rose 1982; Wittman et al. 2017; Grieshop and Arnqvist 2018; Connallon and Chenoweth 2019). Other specific biological dynamics may predict the maintenance of inversion polymorphism in particular ways. For example, inversions may also link drivers and enhancers in meiotic drive systems (Babcock and Anderson 1996; Mroczek et al. 2006; Courret et al. 2019), and these drive-associated inversions may persist at equilibrium frequencies in the presence of repressors (*e.g.* Bastide et al. 2022).

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

1.2 Inversions in *Drosophila melanogaster*

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

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

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

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

Given the relevance of sexual selection and sexual antagonism to *Drosophila* evolution, we propose that polymorphic *D. melanogaster* inversion and standard arrangements may help maintain haplotypes of alleles that specify opposite ends of a sexually antagonistic tradeoff. We first explore this model through forward simulation, to establish that it may in theory arise in a population and maintain inversion polymorphism over a significant number of generations. Then we examine the results of an experiment with an outbred laboratory population of *D. melanogaster*, to test whether common inversions show changes in frequency over the course of a generation in a manner consistent with a trade-off between survival and male reproductive success. Many such antagonistically

203 pleiotropic scenarios are possible under different selective tradeoffs, and indeed some have been proposed for
204 other model inversions with varying mechanisms of maintenance (Betrán et al. 1998; Pearse et al. 2019; Mérot et
205 al. 2020; Pei et al. 2023; see Discussion), but we focus on a specific scenario of sexual antagonism here.

206 2 Results

207 2.1 A Model of Sexually Antagonistic Inversions

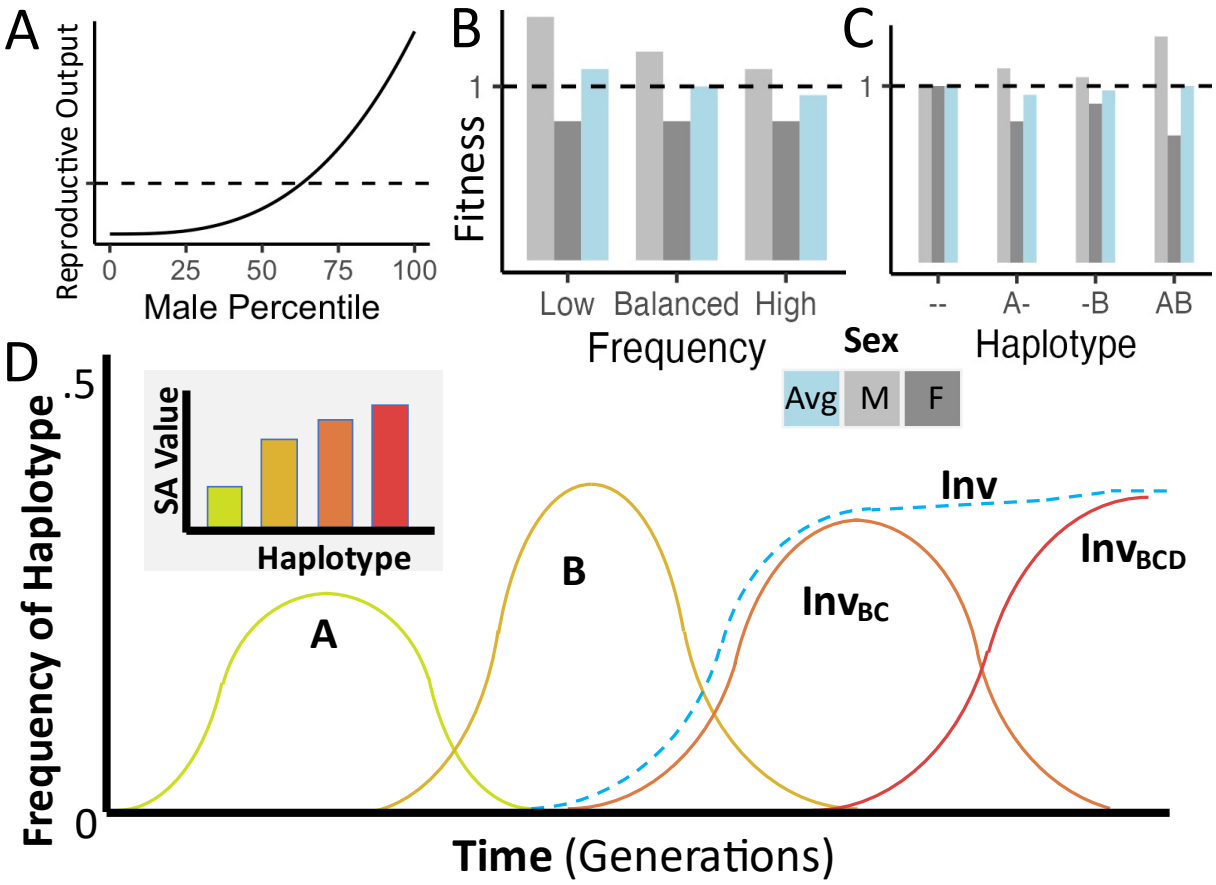
208 Here, we consider a population featuring a competitive mate choice system that results in a skewed distribution of
209 reproductive success among males. Within this population, we model a genomic region containing a number of
210 linked loci, each with alternate antagonistic, pleiotropic alleles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

214 If mate choice is highly competitive, such that each female can choose amongst several or many males, the
215 variance of offspring number among males (Figure 1A) may be much higher than for females - yielding a
216 distribution of male reproductive output that follows a relatively convex trend, in which most males produce few
217 offspring. Conversely, superior display quality can enable males to mate successfully much more often than other
218 display phenotypes they compete against (Figure 1A). An increase in the display of a male with low quality results
219 in a smaller change in number of offspring than an increase in display of a male with high quality (change ii in
220 Figure 1A), due to the reproductive skew towards highly competitive individuals. So, there is an emergent non-
221 additive fitness effect, where multiple variants increasing the reproductive rank of an individual together have a
222 larger fitness benefit than the combined effects of each substitution separately - a positive, synergistic epistatic
223 effect. This effect occurs whether the difference is substituting alleles at two loci, or substituting two homologous
224 alleles at a locus in a diploid or polyploid, which would give an emergent dominance effect. Additionally, fitness
225 conveyed by an allele favoring display quality is also frequency-dependent: since mating success depends on the

display qualities of other males, the relative advantage of a display trait will be diminished as more males carry it, for both haploid and diploid populations (Figure 1B, male fitnesses). We refer to this as a negative frequency dependent selection upon alleles favoring display, though the alternate alleles have a positive frequency dependence. We note that some authors instead use “negative frequency dependence” to specifically describe fitness functions in which any and all alternate alleles are of selective benefit when rare (*e.g.* Charlesworth and Charlesworth 2010). And so, for an antagonistic variant that benefits display and begins under net positive selection, the benefit will decrease as the variant rises in frequency. If a variant only conveys benefit to mate competition success, it is expected to fix given enough time, as even if its evolution is nearly neutral at high population frequency, selection will push the frequency back up if it drifts down (Hazel 1943; Dickerson 1955; Charlesworth and Hughes 2000). However, if the variant is antagonistically pleiotropic for traits not involved in competition between males, such as survival in the face of climate, pathogens, or predators, these fitness costs may typically not be frequency dependent. If the benefit of an antagonistic display-favoring allele overcomes the pleiotropic cost when the variant is rare, its diminished benefit when at higher frequency may reverse its net benefit, while at some intermediate frequency, we predict that these fitness effects should be equal and an equilibrium frequency should be reached and maintained (Figure 1B). In the absence of other forces, this dynamic is expected to maintain the allele under balancing selection indefinitely.

We further predict that multiple such balanced pleiotropic loci should generate interesting emergent behavior. Because of the emergent synergistic epistasis from mate competition (see discussion of Figure 1A above), the haplotype with the greatest male display benefit should increase in frequency relative to other haplotypes until reaching an equilibrium at which its relative advantage in mate choice is reduced enough to be balanced by the pleiotropic, frequency-independent, non-epistatic survival cost (Figure 1B). The haplotypes and their linked alleles conveying intermediate phenotypes will have less than proportional benefits during mate choice due to the convex relationship between quality and success described above (Figure 1A), but a larger, proportional cost to survival against non-intraspecific survival challenges, giving them lower average fitness in the equilibrium state than either of the extreme haplotypes (Figure 1C). Therefore, recombination between the two extreme haplotypes will generate less fit haplotypes in the gametes, and a local modifier of recombination that reduces recombination and

its associated costs (such as an inversion) will be favored by indirect selection (Altenberg and Feldman 1987; Otto and Lenormand 2002). And so, selection will favor the association of these haplotypes with the alternate standard and inverted karyotypes, which generate fewer recombinant gametes from a heterokaryotypic parent. We further predict that in a population that is close to the balanced equilibrium state for such a pleiotropic trade-off, the population of adults will contain a higher frequency of survival-focused haplotypes than the population of zygotes, but the display-focused haplotypes will increase in frequency in zygotes relative to adults, in line with the life stage-specific fitness differences between these haplotypes.

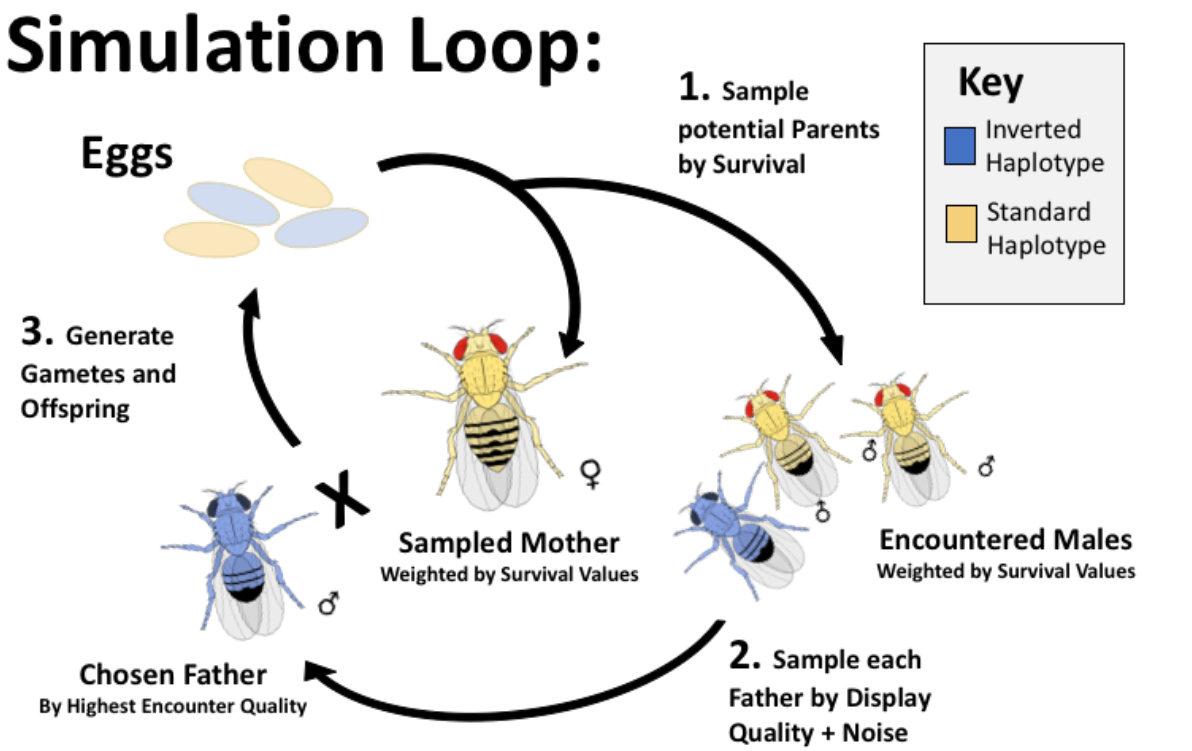


[Figure 1: Conceptual representations of the proposed model of inversion-associated balanced sexual antagonism. (A) Hypothetical skewed distribution of male mating success, yielding a greater variance in reproductive success for

263 males than for females in our model. (B) Fitnesses of display-favoring alleles under sexually antagonistic balancing
264 selection, illustrating a pleiotropic variant that should rise in frequency when rare, but decline if frequency exceeds
265 the balanced equilibrium value. (C) Fitnesses of display-favoring alleles at two haploid loci under the same model,
266 illustrating synergistic epistasis between displaying-favoring alleles. (D) A hypothetical trajectory of four such
267 mutations, in which B outcompetes A, an inversion links B and C to create a more strongly display-favoring
268 haplotype, and then the addition of D to that haplotype furthers the accumulation of antagonistic variants,
269 reaching an equilibrium frequency while displacing less extreme display-favoring haplotypes.]

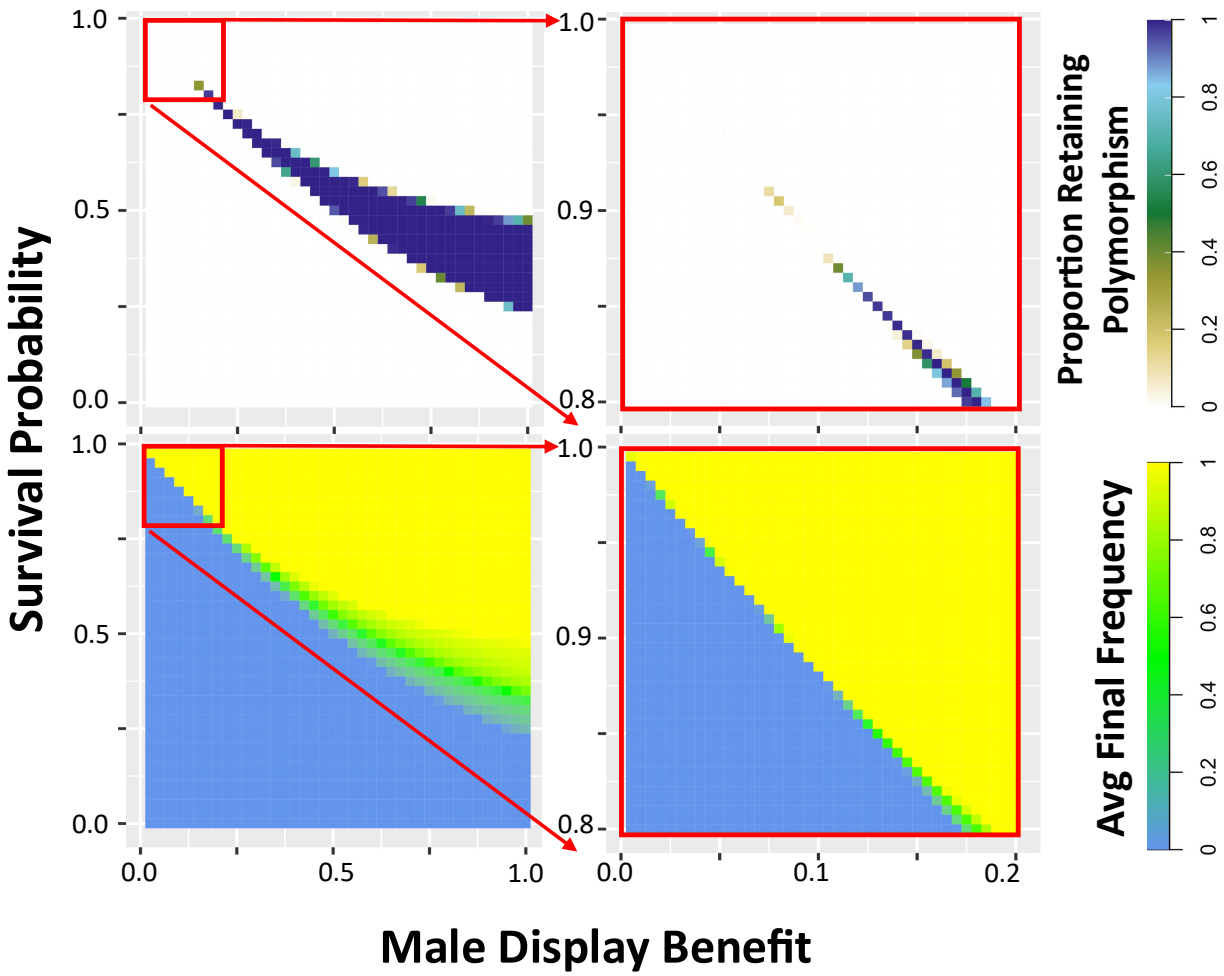
270 While Figure 1 illustrates a haploid scenario for the sake of simplicity, this model extends naturally to a diploid
271 case. The main difference is that heterokaryotypic diploids – those heterozygous for the arrangement – would
272 have intermediate trait values and be less fit than either homokaryotype, for the same selective reasons that an
273 intermediate haplotype with two alleles of opposing effects at different loci is least fit in a haploid context. This is
274 therefore, interestingly, an example of a scenario in which the two most common haplotypes in the population can
275 exhibit underdominant fitness, and yet may still be maintained at a stable balanced equilibrium. By comparison,
276 the simplest models of underdominant fitness predict an unstable equilibrium that disfavors polymorphism
277 (Gillespie 2004, Charlesworth and Charlesworth 2010), a characteristic shared with non-frequency-dependent
278 models of disruptive selection. The cost of generating underdominant offspring might be compensated for by a
279 strong co-directional dominance of the variants involved, such that heterokaryotypic individuals exhibit a
280 phenotype similar to one of the homokaryotypes. However, dominance is not a necessary feature of our model,
281 and it is not invoked in our simulations. Whereas, investigating a wider range of genetic architectures would be a
282 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 - without replacement for each pool and with replacement between pools. For each encounter pool, each male is assigned an observed display quality as the sum of the display effects of its alleles, plus a normally distributed noise effect for each encounter. The highest quality male in the encounter pool is selected as the father, and the offspring genome is generated. The female gamete is generated with crossover and gene conversion events. Odd numbers of crossovers in the interior of an inversion are resampled from the same parent, to model the biology of an organism that eliminates aneuploid gametes during oogenesis, or exhibits reproductive

compensation. *D. melanogaster* males do not cross-over, and so males do not recombine in the simulations presented.]



[**Figure 3**: Simulations show that a balanced equilibrium frequency exists for certain sexually antagonistic variants. The proportion of simulations retaining polymorphism (top) and the average final frequency (bottom) are plotted for a single locus at which one antagonistic allele with the indicated effects reducing survival proportion (of both sexes) and increasing male reproductive quality score is initiated at equal frequency with an alternative allele that has no such effects. Larger effect sizes are plotted in the left panels, with values incrementing by 0.025, and

304 smaller effect sizes (a subset of the larger) are plotted in greater detail in the right panels, with values
305 incrementing by 0.005. Most parameters result in either all simulations retaining polymorphism (purple), or else all
306 simulations losing polymorphism (white). However, simulations retaining polymorphism have a broad range of
307 equilibrium frequencies (shades of green). For each parameter combination, 500 replicates were simulated for a
308 diploid population of 1,000 individuals evolving for 20,000 generations, with further details as indicated in the
309 Materials and Methods.]

310

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

322 With this simulator, we first confirmed that some individual variants with an antagonistic pleiotropy between
323 survival (initially of both sexes) and male reproduction can generate stable selectively balanced polymorphisms *in*
324 *silico*. We simulated a single locus in diploid populations of N = 1000, polymorphic for an allele of a specified
325 survival cost and display benefit at an initial frequency of 0.5, with the alternate allele having no trait effect. One
326 hundred males were sampled for each mating competition. After 20N generations, we calculated the proportion of
327 simulations retaining polymorphism and the average final frequency of the variant across all simulations, revealing
328 that there is a range of paired trait effects for which the mutation remained at intermediate frequency after 20N

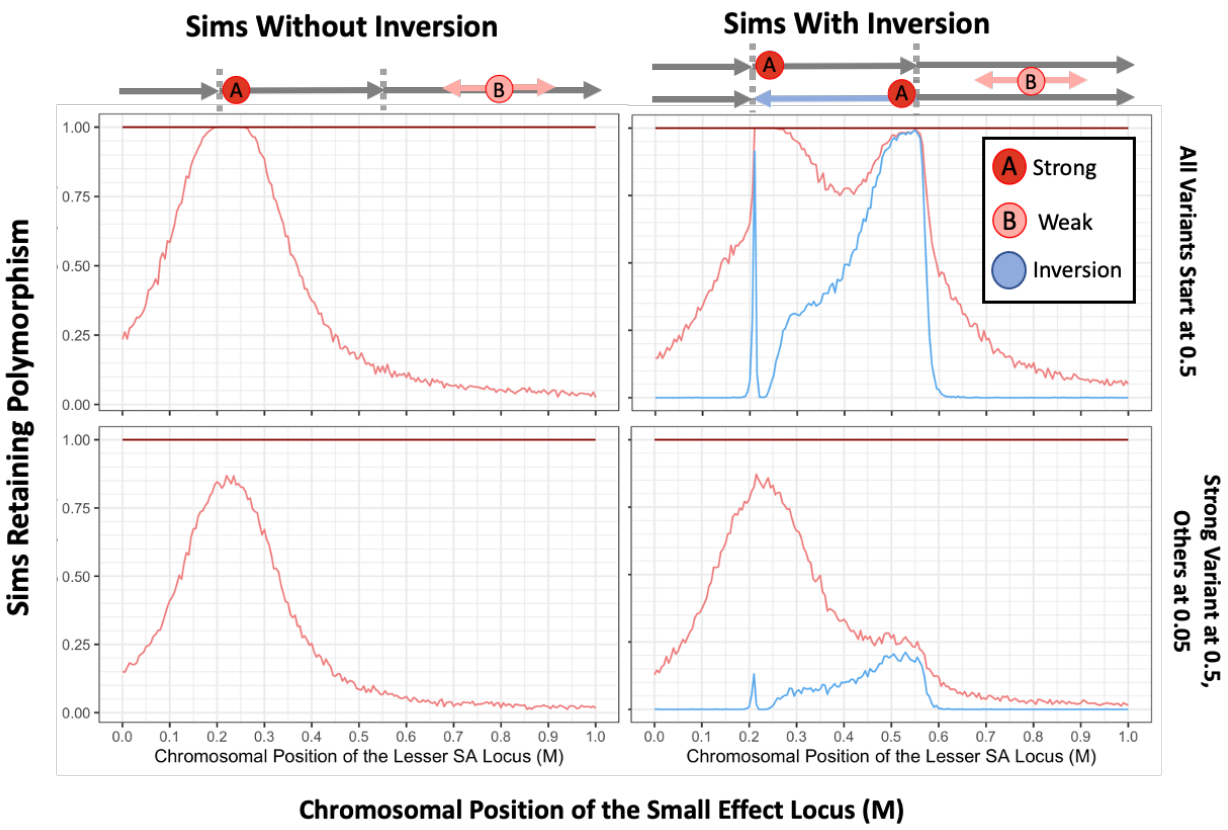
329 generations (Figure 3). Alleles with small effect sizes had only a narrow band of balanced selection, outside of
330 which they were either fixed or lost, while large-effect alleles remained balanced across a larger area of simulated
331 parameter space. While this balanced range is somewhat narrow, such variants may persist under balancing
332 selection indefinitely in the absence of other forces, and many small-effect alleles in linkage blocks may confer
333 larger effects in synergy. In a scenario where only females pay the survival cost, the balanced parameter space of
334 antagonistic trait effects is broader and shifted somewhat (Figure S1). In a scenario where only males pay the
335 survival cost, and variants are neutral in females, essentially no simulations retain polymorphism (Figure S1). The
336 number of males sampled to assess mate competition for each offspring also has a significant effect, with larger
337 values giving broader polymorphic ranges, though simulations with costs only in females show robust
338 polymorphism even when only sampling two males for each competition (Figure S1).

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

352 We then examined the same spacings between antagonistic variants across a region containing a polymorphic
353 inversion, with its left breakpoint close to the stronger variant, and the right breakpoint within the range of the

354 positions in which the weaker variant was simulated (Figure 4, lower panel). When the inversion was added, the
355 weaker variant was then retained when the two antagonistic variants are relatively more distant, due to
356 association of the weaker variant with the right-hand inversion breakpoint (pale red line), at which point this
357 association also preserves the inversion itself in a polymorphic state (light blue line). While more stochastic, a
358 weaker antagonistic variant near the middle of the inversion also appeared to maintain polymorphism more
359 frequently in systems starting with an inversion than without, i.e. across approximate positions 0.30-0.45M in
360 Figure 4, in spite of its weaker linkage with the inversion itself due to double cross-over events. The inversion is not
361 maintained when the weaker antagonistic variant is within the inversion but tightly linked to the stronger variant
362 (see positions 0.22-0.25M); here there is little crossing-over for an inversion to suppress, and so the effect of the
363 inversion approaches neutrality. Interestingly, there is a small window of weaker variant position in which the
364 inversion maintains polymorphism despite the already tight linkage between SA variants, where the inversion
365 seemingly does not modify recombination much (see positions 0.2-0.22M). Here, the weaker antagonistic variant is
366 more tightly linked to the breakpoint than to the stronger variant in this range, and presumably benefits more
367 consistently from the crossover reduction. The maintenance of the weaker variant drops more rapidly outside of
368 the inversion than over similar distances in the simulation without inversions (positions 0-0.2M, bottom versus top
369 panel). This seemingly counter-intuitive result may stem from crossover events that result in aneuploid gametes
370 being resampled in our simulations, which effectively increases the crossover rate outside the inversion
371 breakpoints, a behavior which should be biologically reasonable in taxa with aneuploid chromosome elimination or
372 reproductive compensation. So, for a weaker variant outside the left inversion breakpoint, the presence of the
373 inversion can increase the crossover rate between the two antagonistic variants, and thus reduce the probability of
374 the weaker variant surviving by retaining linkage to the stronger one.

375

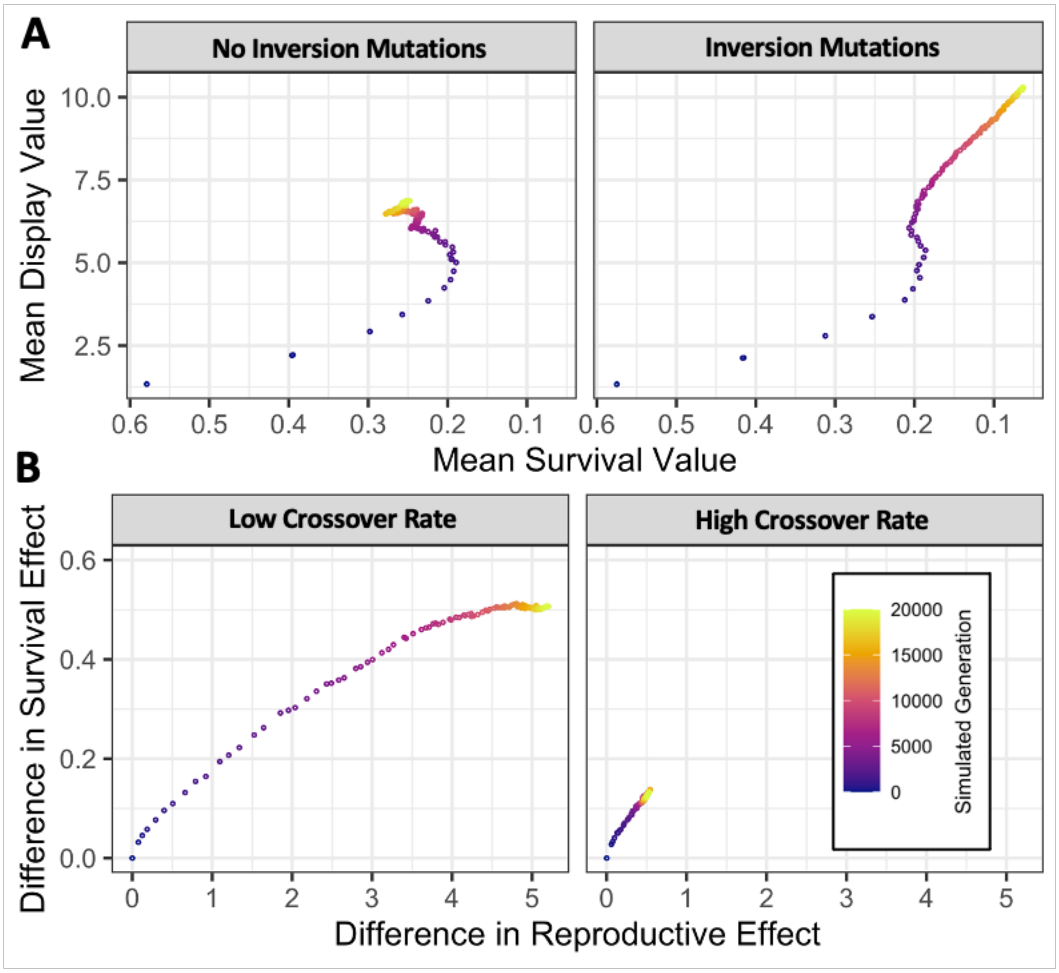


376

377 [Figure 4: Simulated inversions may persist as polymorphisms when linked to sexually antagonistic, pleiotropic
378 variants. In turn, the presence of inversions facilitates the accumulation of antagonistic variation. Rates of long-
379 term persistence are shown for simulations with two defined antagonistic variants alone (left panels) and with the
380 additional presence of a defined inversion (right panels). To focus the potential addition of a second linked variant
381 to an antagonistic haplotype, we begin with the stronger variant already at an intermediate frequency in the
382 population (0.5). The weaker antagonistic variant and the inversion either both start at 0.5 frequency as well
383 (upper row) or both start at a lower frequency of 0.05 (lower row). The 1 Morgan chromosomal segment is
384 diagrammed above the plots, with the inversion breakpoints marked (0.21M, 0.55M), chosen to include space
385 across and beyond the inverted region. The proportion of simulations in which the second, variably-positioned
386 weaker antagonistic variant (survival and reproductive values 0.86, 0.12) maintains polymorphism is shown (light
387 red line), dependent on the recombination distance from a stronger allele (0.75,0.3) at fixed position 0.225M,
388 which always persists (dark red line at 1 across the top of each plot). In the right panels, the proportion of

389 replicates in which the inversion is retained is also plotted (blue line). Values for each parameter combination were
390 averaged over 1000 replicate simulations, for populations of 1000 diploid individuals run for 20*N* generations.
391 Gene conversion occurs for each heterozygous variant at a probability of 10^{-2} per female meiosis with random
392 direction. Crossover arrangements generating aneuploidy, where there are an odd number of crossovers within
393 the inverted region of a heterokaryotypic parent, are resampled from the same parent to represent removal to
394 polar bodies or reproductive compensation.]

395



396

397 [**Figure 5**: Simulated populations with stochastic mutation of new inversions and sexually antagonistic variants
398 generate persistent inversion polymorphisms and accumulate linked sexually antagonistic variants. (A) In simulated

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

408

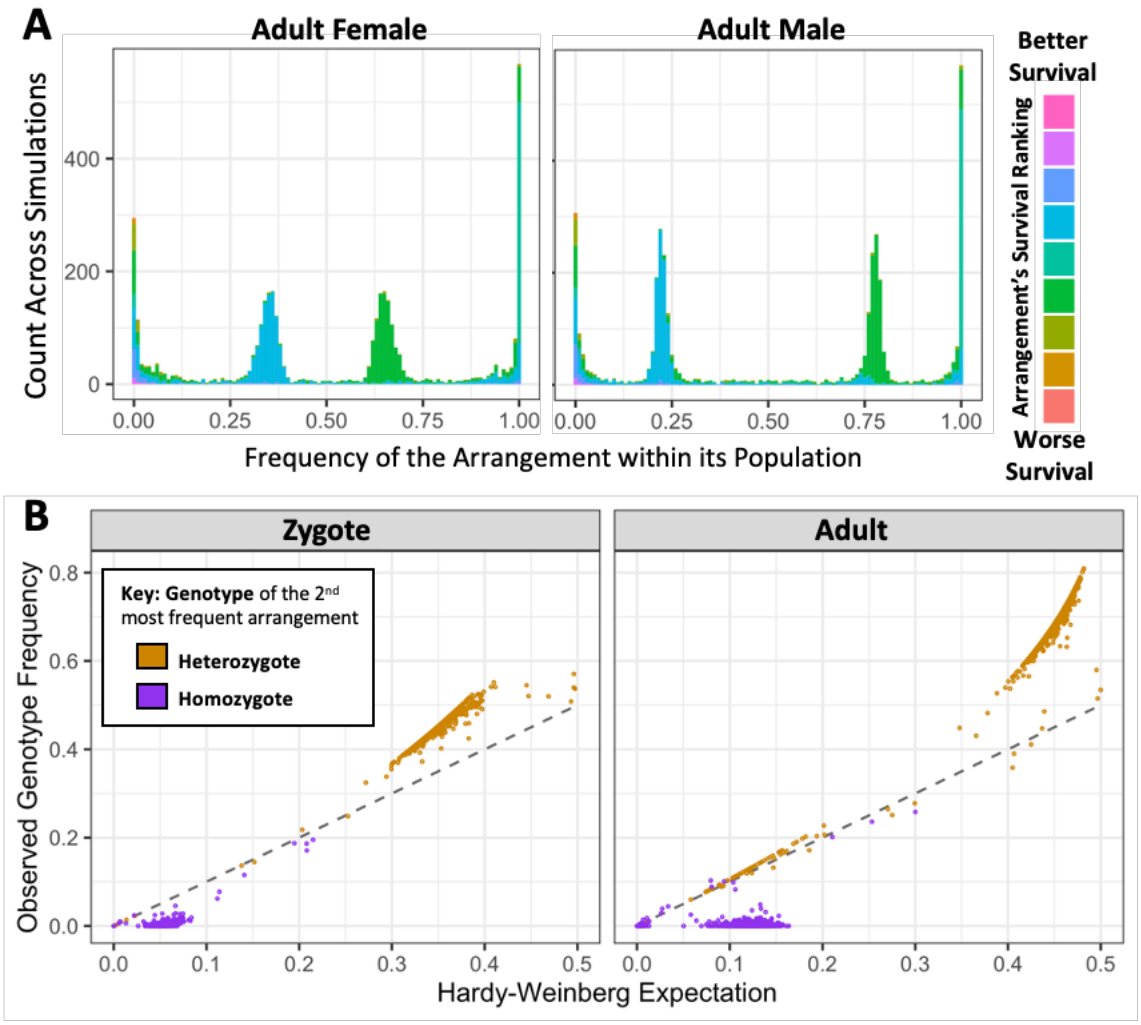
409 In further simulations, we began with a population initially free of variation, in which sexually antagonistic
410 mutations arise stochastically. To assess differences between longer or shorter antagonistic regions, or higher or
411 lower crossover rates, we modeled populations with parameters scaled to either 1 Mb or 10 kb, using estimates
412 from Comeron et al. (2012) and Huang et al. (2016). The mutations had pleiotropic effects increasing reproductive
413 quality score and decreasing survival proportion, each drawn sampled independently (Materials and Methods).
414 Depending on the effect scores drawn, a mutation could be broadly advantageous (potentially leading to a
415 selective sweep), broadly deleterious, or potentially subject to balancing selection (Figure 3). These mutations
416 arose at a rate equal to one thousandth of the scaled single nucleotide mutation rate given in Huang et al. (2016).
417 Inversions (of random lengths and positions) also arose stochastically, at a hundredth of the rate of antagonistic
418 variants. The frequency trajectories of antagonistic variants and inversions are tracked, and the simulation was run
419 for 20*N* generations to allow near-equilibrium patterns to be established. These simulations demonstrated the
420 potential of a population without initial balanced polymorphism or genetic variation to establish alternate
421 karyotypes which link sexually antagonistic, pleiotropic alleles and an inversion, to maintain polymorphism in a
422 balanced manner, and then to subsequently accumulate significant antagonistic character (Figure 5A), although

423 the degree of antagonism accumulated depends strongly on the rate of crossing over (relative to the occurrence of
424 antagonistic mutations; Figure 5B).

425 The simulations that successfully maintained balanced antagonistic chromosomes also had an interesting
426 distribution of allele and genotype frequencies of the balanced haplotypes at equilibrium. Without any selection,
427 the inversion frequency follows a pattern characteristic of the expected neutral site frequency spectrum (Figure
428 S2). But strikingly, in parameter spaces where simulations with shared survival costs between sexes maintained
429 polymorphism under balancing selection, inversion frequency distributions were tightly clustered around 0.25 and
430 0.75 among the zygote population (Figure S2, 6A). We found that this outcome was related to the genotype
431 frequencies of reproducing adults. These simulations evolve toward strong antagonism, and it is predominantly
432 males homozygous for the display-favoring haplotype (whether that is inverted or standard) who succeed in
433 reproducing, and these males therefore contribute one copy of the display-favoring haplotype to each of their
434 male and female offspring. However, this haplotype's significant cost to viability and longevity would select for the
435 heterokaryotypic females over the display-favoring homokaryotypes. Given that successful parents would
436 predominantly be heterokaryotypic females and display-favoring homokaryotypic males, around three quarters of
437 transmitted chromosomes have the display-favoring arrangement and one quarter does not, yielding the 0.25 and
438 0.75 zygote inversion frequencies seen in simulations. Although these are autosomal simulations, selection yields
439 an inheritance pattern is similar to that of a ZW sex chromosome system, where the Z chromosome is
440 homokaryotypic in the fathers, and heterokaryotypic with the W chromosome in the mothers.

441 The sex-specific fitness dynamics present at equilibrium under this model also result in population-wide genotype
442 frequencies that depart from standard Hardy-Weinberg expectations. If the preponderance of matings are
443 between heterokaryotypic females and homokaryotypic display-favoring males, then homokaryotypes for the
444 survival-focused haplotype will rarely be produced (as confirmed in Figure 6B). We note that in cases where the
445 standard arrangement favors male display, then especially if antagonism is strong, this model could explain an
446 observed deficiency or absence of inversion homokaryotypes, even if the inverted arrangement was not associated
447 with any recessive deleterious load. Further, heterokaryotypes may become disproportionately frequent after

viability selection (Figure 6B). And still, this enrichment for heterokaryotypes reflects neither the existence of recessive deleterious load associated with each arrangement, nor overdominance with respect to viability. Hence, our model may offer another alternative explanation for deviations from Hardy-Weinberg genotype frequencies observed for polymorphic inversions in natural populations.



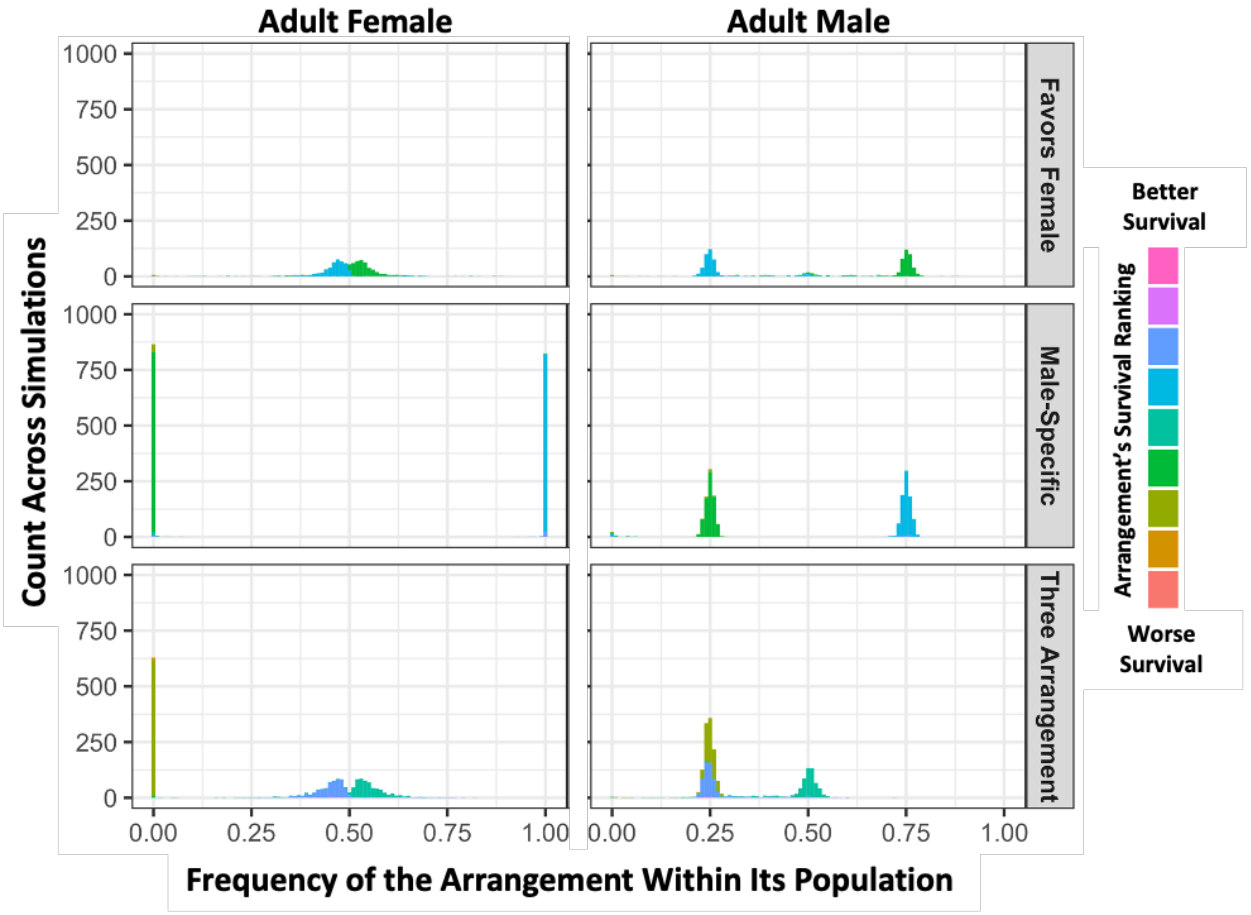
[**Figure 6:** Antagonism-associated karyotypes reach predictable equilibrium frequencies in simulations with antagonistic and inversion mutations, when survival costs are shared between sexes. (A) Histograms of the number of arrangements at a specified frequency across all simulation replicates in females (left) and males (right), colored

457 by how the average survival effect of that arrangement ranks compared to the others in its population. Ranks are
458 normalized to be relative to the median arrangement value, to better account for rare arrangements and ties.
459 These simulations reach an equilibrium with two predominant arrangements, with the more survival-focused
460 arrangement around an overall frequency of 0.25 and more common in females, and the more male-competition-
461 focused arrangement around a frequency of 0.75 and more common in males. (B) Because successful males tend
462 to be homozygotes and successful females heterozygotes, the less frequent of the two major arrangements (which
463 tends to favor survival) has homokaryotype frequencies far below the Hardy-Weinberg expectation, and there is an
464 excess of heterokaryotypes. These genotype frequencies shift somewhat between zygote (left panel) and adult
465 (right panel) populations as the genotype affects survival.]

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

478 Specific properties of our simulations may have influenced the occurrence of these three stable states under the
479 female-limited cost model. To reduce the computational complexity of both crossovers and arrangement-specific
480 trait values, the simulations do not allow two inversions on the same individual chromosome, whether overlapping
481 or not, though two chromosomes sampled from the population may have distinct inversions with overlapping

breakpoints. This means that if two arrangements, each with an inversion, become the prevalent haplotypes and thereby remove all standard arrangement chromosomes from the population, there is no mutational path to generate a third arrangement. However, a portion of the two-state simulations still had the ancestral, non-inverted arrangement as one of the two haplotypes present, which could have been mutated to a third arrangement to establish the three-arrangement pattern. We are not sure what selective or mutational barriers may prevent transitioning between each state once established. One clear barrier is that it is difficult to generate a less extreme arrangement from a Y-like arrangement, as these can accumulate an indefinite number of mutations since males experience no costs.



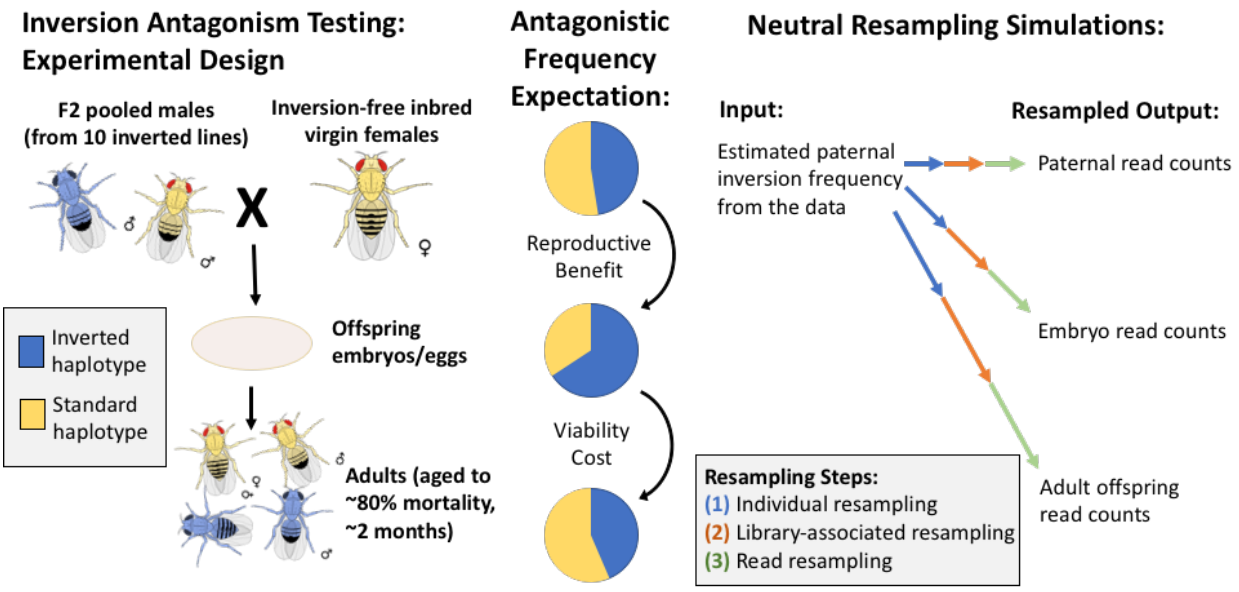
492 [**Figure 7:** In simulations with female-limited survival costs, three distinct outcomes are observed involving either
493 two or three balanced haplotypes. From simulations with randomly occurring antagonistic and inversion
494 mutations, histograms of inversion frequencies across simulations for surviving females (left) and males (right) are
495 partitioned into three categories based on sex-specific haplotype frequency outcomes. The "Favors Female"
496 category (top) includes simulations in which there are exactly two arrangements with frequencies between 0.1 and
497 0.9 in both sexes (20.0% of replicates). The "Male-Specific" category (middle) includes simulations in which there
498 are two arrangements with whole-population frequencies between 0.1 and 0.9, but one arrangement is absent in
499 adult females (37.0% of replicates). The "Three Arrangement" category (bottom) includes simulations in which
500 there are exactly three arrangements with whole-population frequencies between 0.1 and 0.9 frequency (40.6% of
501 replicates). 2.3% of simulations did not fit in these categories and had only one predominant arrangement. As in
502 Figure 6, inversions are colored by how the average survival effect of that arrangement ranks compared to the
503 others in its population, and ranks are normalized to be relative to the median arrangement value, to better
504 account for rare arrangements and ties.]

505

506 2.3 Empirical tests for inversion-associated tradeoffs

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

513

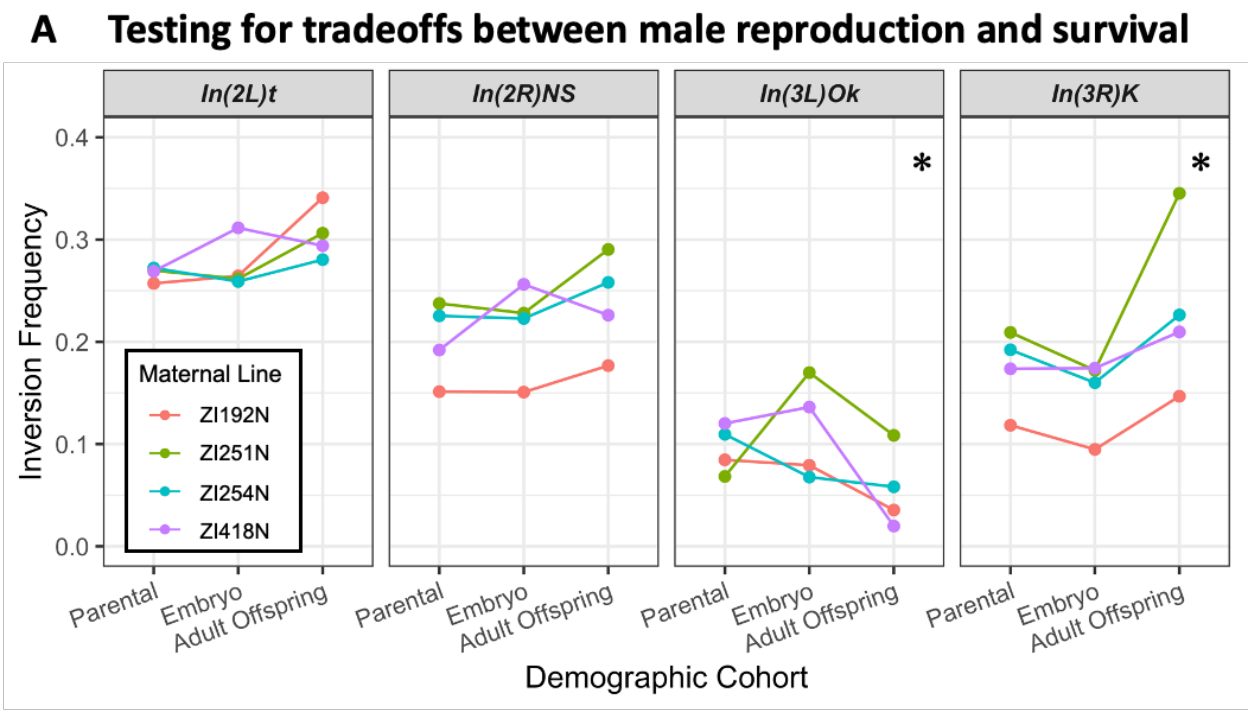


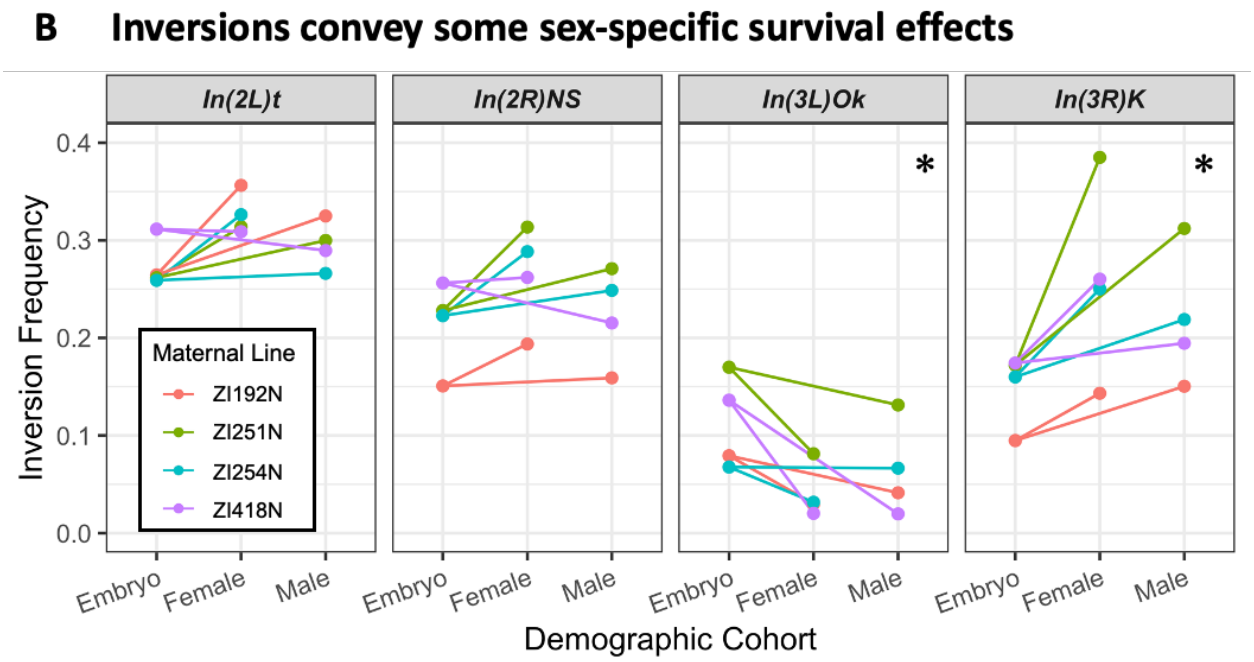
[**Figure 8:** The layout and potential expectations of the laboratory evolution experiment. We cross outbred males from a high-inversion population to inbred, non-inverted females from a specific inbred line, and collect DNA samples from the fathers, embryo offspring, and aged adult offspring (left).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

We assembled an outbred population with moderate frequencies of each inversion to obtain karyotypically diverse males. We set up four crosses using these males along with females from one of four different inversion-free inbred Zambia lines. This crossing scheme is represented in Figure 8. These parental crosses occurred in a single large population cage, with populations of at least 600 of each sex. We developed a PCR and next-generation sequencing-based assay to estimate inversion-associated SNP frequencies in each pool of males and in each corresponding pool of embryos. These SNPs (given in Table S1) were based on our own analysis of variants that showed perfect association with inversions among 197 Zambia genomes (Table S2; see Materials and

530 Methods). We then implemented a resampling method to test for significant inversion frequency changes between
531 parents and offspring (while accounting for the variance observed in technical replicates; see Materials and
532 Methods), in order to determine whether males carrying each inversion were more or less successful in
533 reproducing than the average male. We also compared frequencies between embryos and aged adults that
534 eclosed either in the first or second half of the collection period, to investigate the relationship between inversions
535 and viability/longevity. We calculated p-values from the resampling tests to assess the significance of each
536 individual change in inversion frequency between relevant samples (Table S3). However, in accord with our
537 primary question regarding fitness tradeoffs, we primarily focus here on joint tests of two opposing frequency
538 changes, as further described below. We could therefore test the hypothesis that a given inversion had opposing
539 influences on survival and male reproduction, which would provide evidence consistent with our model of
540 balanced sexual antagonism, while also examining whether any such tradeoffs depend on maternal genetic
541 background.

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





[**Figure 9**: Some *Drosophila* inversions show evidence for antagonistic tradeoffs and sex-specific survival. (A) Testing for antagonistic fitness effects on reproduction and survival. Each inversion's frequency in parents, embryo offspring, and adult offspring are shown for each maternal inbred line. Parental frequency is taken as the paternal

559 frequency divided by two to account for the inversion-free inbred mothers. Adult offspring frequencies are taken
560 from the sum of estimated allele counts across all adult offspring cohorts. * indicates p-value < 0.05 for evidence of
561 significant reversed frequency change from parental to embryo and from embryo to offspring, combined across all
562 four maternal line crosses and corrected for multiple tests, when compared to neutral simulated experiments
563 (details in Materials and Methods). (B) Testing for evidence of sex-specific survival effects. The changes in inversion
564 frequencies between embryos and both female and male adult offspring are illustrated. * indicates p-value < 0.05
565 for significant differences in female versus male inversion frequency among adult offspring, combined across all
566 four maternal crosses and corrected for multiple tests.]

567

568 When we separated the effects of male reproduction (based on father to embryo changes) and survival (embryos
569 versus adults), we found that the effects of inversions on survival were generally consistent across maternal
570 genotypes. Three inversions were associated with increased survival: *In(3R)K* was in all four crosses, while *In(2L)t*
571 and *In(2R)NS* each showed elevated survival with all maternal strains except ZI418N (Figure 9A). In contrast,
572 *In(3L)Ok* was associated with lower survival in each cross (Figure 9A). With regard to male reproduction, the
573 trajectory of inversion frequencies from fathers to embryos to offspring was less consistent between different
574 crosses (Figure 9A), implying that differences in male reproductive success may be dependent on the maternal
575 genotype, which is consistent with past findings in *D. melanogaster* (Reinhart et al. 2015). Despite such variability,
576 each of the four inversions had multiple frequency changes that departed strongly from null expectations (Figure
577 9A; Table S3), as might be expected if selection was present, but varied in strength and direction between crosses
578 and life stages.

579 To test our central hypothesis of a tradeoff between male reproductive success and cumulative survival during
580 development and adult aging, we evaluated the significance of each change in inversion frequency between the
581 paternal and embryo cohorts, alongside the change in inversion frequency between the embryo and combined
582 adult offspring cohorts. Among these results, *In(3R)K* had particularly consistent changes in frequency across
583 replicate lines, with a decrease between parents and embryos, and then an increase between embryos and aged

584 offspring (Figure 9A). When we assessed the likelihood of observing contrasting frequency shifts between life
585 stages of this magnitude, integrating across the four maternal crosses and accounting for multiple test correction,
586 we found that the *In(3R)K* pattern was significantly non-random ($p = 2.70 \times 10^{-3}$; Table S4; see Materials and
587 Methods). The other inversion that gave reversed frequency shifts from cross-averaged data was *In(3L)Ok*, which
588 unlike *In(3R)K*, was associated with (variably) higher male reproductive success and (consistently) lower survival.
589 The frequency reversal observed for *In(3L)Ok* was also found to be statistically significant ($p = 2.97 \times 10^{-7}$). Wright-
590 Fisher estimates of selection coefficients for these and all comparisons are given in Table S8.

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

3 Discussion

Here, we have introduced a new conceptual model in which (1) sexually antagonistic variation is present and subject to balancing selection that maintains extreme genotypes favoring either male reproduction or survival, and (2) inversions are predicted to facilitate the aggregation of progressively more antagonistically differentiated haplotypes. We have tested this model using simulations, confirming the predicted efficacy of balancing selection and demonstrating that inversions indeed facilitate the construction of haplotypes containing multiple antagonistic variants. Finally, we have deployed *Drosophila* lab experiments to uncover some evidence for inversion-associated tradeoffs between reproduction and survival, depending on the specific inversion and strains tested.

3.1 Inversions in *D. melanogaster*

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

As indicated in the Introduction, models like ours in which inversion frequencies are influenced by balancing selection on pleiotropic tradeoffs may help explain multiple aspects of observed *D. melanogaster* inversion frequencies. Adjacent models have been explored in flies, for example an inversion in *Drosophila buzzatii* that was found to trade off between faster development time and larger adults (Betrán et al. 1998), and an inversion

631 polymorphism found to trade off between larval survival and adult reproduction in the seaweed fly *Coelopa frigida*
632 (Mérot et al. 2020). Some other groups found to harbor antagonistic inversions include the rainbow trout (Pearse
633 et al. 2019) and the zebra finch (Pei et al. 2023). These related systems are variously hypothesized to involve
634 maintenance by overdominance (Mérot et al. 2020), linked deleterious recessives (Pei et al. 2023), or reversals of
635 dominance (Pearse et al. 2019), but these mechanisms are not mutually exclusive. This class of models inherently
636 accounts for the intermediate maximal frequency of most *D. melanogaster* inversions in natural populations. To
637 the extent that inversions favor male reproduction, the sexually antagonistic model we propose here could also
638 help account for the predominantly autosomal locations of common inversions in this species, given that male-
639 specific fitness effects have a weaker influence on the frequencies of X-linked variants since the X chromosome
640 spends only one third of its time in males.

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

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

662 More broadly, we observed a trend toward increased inversion frequencies across our full experimental
663 generation. A general benefit for inversions would mirror previous findings in which inversions found at
664 intermediate frequencies in natural *D. melanogaster* populations (Inoue 1979), as well as inversions from seaweed
665 flies known to be under balancing selection in the wild (Mérot et al. 2020), were both inferred to be under
666 directional selection in the lab. However, one laboratory study did report evidence that a *D. melanogaster*
667 inversion not tested here, *In(3R)P*, had frequency-dependent fitness under crowded conditions (Nassar et al.
668 1973). Furthermore, even an inversion found to be under directional selection in the lab might still be subject to a
669 balanced tradeoff a more challenging or complex natural environment. Other studies in *D. melanogaster*, including
670 those on focused seasonal evolution, have found evidence for important tradeoffs between reproduction and
671 robustness to environmental challenges (*e.g.* Behrman et al. 2015). Hence, it would clearly be desirable to study
672 the effects of inversions on a wider range of potential tradeoffs. For example, Said et al. (2018) suggested that
673 inversions *In(2L)t* and *In(3R)K* might be maintained by immunity-related tradeoffs, in light of an over-
674 representation of immune-related genes among differentially expressed transcripts between inverted and
675 standard karyotypes.

676 In addition to a broader array of balanced tradeoffs that could modulate *D. melanogaster* inversion frequencies,
677 drift and other forms of selection may contribute as well. Although we emphasize above the potential interplay
678 between ecological adaptation and balanced tradeoffs, some inversions could be impacted by simpler forms of
679 directional selection. However, we must then invoke secondary explanations to account for their limited maximal
680 frequencies (see Introduction). Other forms of balancing selection may also come into play as well. It is clear that
681 some variation in this species is subject to temporally-varying selection, and some inversions have displayed

682 seasonal frequency differences (Dobzhansky 1943; Kapun et al. 2016; Machado et al. 2021; Lange et al. 2022). As
683 with spatial clines, temporal shifts could reflect either simple directional selection or else environmentally-
684 modulated tradeoffs in a model such as ours. Spatial selective pressures may vary both between populations and
685 within population scales.

686 As indicated in the Introduction, balancing selection on inversions could also be achieved by linkage to recessive
687 deleterious variants, which may be observed as an antagonistic pleiotropic phenotype despite potentially involving
688 strictly deleterious variants in linkage (Pei et al. 2023). In extreme form inversions might be under associative
689 overdominance (Zhao and Charlesworth 2016; Gilbert et al. 2020), where two haplotypes are balanced by
690 recessive cost to both homozygotes. Inversions may be particularly susceptible to carrying recessive deleterious
691 alleles if population size and/or inversion frequency are low, since under these circumstances inversion
692 homozygotes are rarely generated and hence the crossover rate among inverted chromosomes is low, plus there
693 are fewer inverted chromosomes to facilitate rare double recombination events with standard chromosomes.
694 However, this dynamic is not expected to establish new inversions easily (Charlesworth 2024) and is expected to
695 decay with increased age and population size of the inversion, as mutation, gene conversion, or double crossover
696 resolve the negative linkage between recessive deleterious alleles. Most common inversions in *D. melanogaster*
697 are both relatively old and have large effective population sizes. Inversions including *In(2L)t*, *In(2R)NS*, and *In(3R)K*
698 have breakpoints estimated to be several times older than the expansion of the species out of southern-central
699 Africa (Corbett-Detig and Hartl 2012; Sprengelmeyer et al. 2020). Those three inversions, plus *In(3L)Ok*, for which
700 inversion age has not been estimated, all maintain frequencies greater than 0.15 within very large populations
701 (Kapun and Flatt 2019; Sprengelmeyer et al. 2020), and so it is not clear that the persistence of recessive
702 deleterious or lethal alleles in the inverted regions is expected to be high enough that inversions would be
703 maintained by a shared recessive lethal variant preventing fixation. However, sampling of natural populations has
704 demonstrated a somewhat higher frequency of recessive lethal variants in inverted haplotypes, but these are
705 largely unique and heterokaryotypic individuals remain robust, giving low costs to a heavily outbred population
706 (Mukai and Yamaguchi 1974). Many wild-derived inbred strains of *D. melanogaster* are homokaryotypic for
707 inversions (Lack et al. 2016a). Furthermore, as noted in the Results, an observed lack of inversion homozygotes

708 could reflect strong sexual antagonism as opposed to recessive deleterious variation. Nevertheless, further
709 simulation and experimental testing regarding the potential influence of associative overdominance and other
710 processes on inversion frequencies in *D. melanogaster* and other species would be desirable.

711 3.2 Generality of Balanced Pleiotropic Inversions

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

718 For the specific model explored here, of antagonistic pleiotropy between survival and display in a mate choice
719 system, the prevalence of the dynamic across species and populations is unclear. It seems plausible that a similar
720 model of antagonistic pleiotropy and balancing selection under mate choice underlies other systems, such as the
721 balanced inversion system in the European ruff *Calidris pugnax* (formerly *Philomachus pugnax*), where three
722 inverted haplotypes on an autosome determine aggressive, cooperative, and female-mimic male lekking morphs,
723 albeit with a dominance relationship and homozygous inviability for an inverted haplotype (Küpper et al. 2016). A
724 recent study on the female-mimic Faeder morph suggests that it may be maintained in part by a balance between
725 its effects lowering female reproductive output while increasing male mating success, and the system is likely more
726 complex as a whole due to the three interacting morphs (Giraldo-Deck et al. 2022). More generally, a number of
727 factors contribute to the potential prevalence of this dynamic. Mate competition and choice during reproduction is
728 rather common among animals, and this may mean sexually antagonistic inversions are plausibly also common.
729 The model depends on the distribution of fitness effects of new mutations and the availability of sexually
730 antagonistic variation, but such variation does seem prevalent (Zajitschek and Connallon 2018; Ruzicka et al. 2019).
731 More likely to limit the prevalence of such inversions are the meiotic costs experienced by inversion

732 heterokaryotypes in some systems (White 1973), whether sexually antagonistic variants are clustered in a region
733 that would allow inversions to increase linkage, and whether other mechanisms like sex-biased expression (Ingleby
734 et al. 2015), gene duplication (Connallon and Clark 2011), and linkage to the sex chromosome or sex-determining
735 locus would resolve the antagonism more frequently (though see following subsection on sex chromosome
736 evolution).

737 Perhaps more importantly, there are a number of well-studied cases of inversion polymorphism that may follow
738 this dynamic of balancing selection between haplotypes simultaneously demonstrating antagonistic pleiotropy and
739 trait-specific frequency dependence, but which do not involve sexual antagonism. *Papilio* butterfly populations
740 harbor well-studied mimicry polymorphisms that segregate between inverted segments and may often experience
741 frequency-dependent selection (Clarke and Sheppard 1960; Kunte et al. 2014; Poul et al. 2014). The snail *Cepaea*
742 *nemoralis* harbors diverse and distinct shell patterning haplotypes associated with inversions that remain balanced
743 within single sampling locations (Cook 2013; Surmacki et al. 2013). Several inversions in different ant species
744 segregate a single-queen colony strategy from a cooperative multi-queen strategy within the same range (Wang et
745 al. 2013; Libbrecht and Kronauer 2014).

746 Essentially, any population with antagonistic pleiotropy involving some form of negative frequency-dependent
747 balancing selection on one of the phenotypes might exhibit the same dynamic. The population would experience
748 similar divergent selection towards alternate phenotypic extremes, with the frequency-dependence maintaining
749 the presence of two genotypes. Importantly, this model involves a different mechanism for the maintenance of
750 antagonistically pleiotropic variation than models based on reversals in dominance, where dominance effects
751 generate a net heterozygote advantage at a locus (Connallon and Chenoweth 2019). The model proposed here
752 integrates well with existing thought on the evolution of divergent selection on balanced systems. For example,
753 Kopp and Hermisson (2006) have proposed a specific form of frequency dependent divergent selection (FDDS),
754 involving competitive exclusion between similar genotypes for a trait otherwise under stabilizing selection, and this
755 model was expected to result in single or few loci of large effect. Their model did not consider pleiotropy or linkage
756 modifiers, but still had comparable dynamics in that selection in the FDDS regime favored small numbers of large

757 effect loci, while the inversion haplotypes in the model presented here simplify the recombination architecture to
758 approximate a single locus. This simplification of the genetic architecture follows the reduction principle in
759 recombination modifier theory (Feldman 1972; Feldman and Balkau 1973; Feldman et al. 1980; Altenberg and
760 Feldman 1987; Altenberg et al. 2017). When a population experiences selection towards two (or more) divergent
761 optima, under a selection regime that provides a balancing mechanism to maintain both, indirect selection will
762 favor a reduction of the genetic architectures to prevent the generation of intermediate phenotypes. The
763 likelihood of either model likely rests on the inherent polygenicity of the trait under selection, with simpler traits
764 being more likely to resolve by single locus FDDS dynamics than by inversion association (*e.g.* Yassin et al. 2016).

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

773 3.3 Antagonistic autosomal inversions may precede novel sex chromosomes

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

781 Here, we discovered that strong inversion-linked antagonism could yield a scenario where essentially all highly-
782 successful males were homokaryotypic for a display-favoring karyotype, whereas reproducing females were
783 overwhelmingly heterokaryotypic (Figure 5). This situation is similar to the ZW sex chromosome system, suggesting
784 that this particular pleiotropic trade-off may favor resolution by linkage to a female sex determination factor.
785 Further, simulations in which only females experienced antagonistic survival cost usually generated populations
786 with a chromosome essentially lethal in females, thus exclusively transmitted in males similar to a Y chromosome
787 (Figure 7). These simulations sometimes had three states that appeared stable at 20*N* generations, including Y
788 chromosome-like, W chromosome-like, and shared intermediate arrangements. This three-arrangement stable
789 polymorphism seems unusual for a natural population, but might be interesting to explore in populations with
790 polymorphic sex determination.

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

797 A basic prediction of the above hypothesis is that at least one inversion difference between sex chromosomes
798 should already exist at the time of new sex chromosome formation, and genomic differentiation therefore begins
799 between inversion karyotypes on an autosome even before sex-specific transmission of the region. This scenario
800 contrasts with current models of sex chromosome evolution in which inversion accumulation and genomic
801 differentiation accrue only after sex chromosome formation, due to advantages in suppressing recombination
802 involving sexual antagonistic variants already associated with a sex-determining locus (Wright et al. 2016). In this
803 context, we note that the neo-X and neo-Y homologs of Muller element C in *Drosophila miranda*, which became
804 sex-linked about 1.5 million years ago, already appear to have inversion differences between them (Chen et al.
805 2024 – Figure 5A). Further, a recent study in *Littorina saxatalis* snails has discovered an ecotype-specific sex

806 determination system that may be founded on such tradeoffs, though selection across ecotypes involves discrete
807 geographic boundaries and is only partly comparable (Hearn et al. 2022). Still, additional genomic studies of new
808 sex chromosomes are needed to assess the potential role of sexually antagonistic inversions in sex chromosome
809 evolution.

810 Summary and Future Prospects

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

822 In light of the persistent mystery of how *D. melanogaster* inversion frequencies are determined, and how readily
823 we detected tradeoffs involving male reproduction and survival, we suggest that further studies investigating the
824 potential fitness tradeoffs of inversions in this model system are strongly warranted. While some such studies have
825 previously been conducted, the applicability of our amplicon sequencing approach to estimate inversion
826 frequencies in large pools of flies may improve the sensitivity of such experiments. Our understanding of the
827 phenotypic impacts of *D. melanogaster* inversions would benefit from investigations of a wider range of fitness-
828 related traits, including female fecundity, immunity, and performance in challenging environments such as cool
829 temperatures and high population densities.

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

4 Materials and Methods

4.1 Simulations

We developed a new individual-level forward simulator ('SAIsim' for Sexually Antagonistic Inversion simulator) written in python 3.7 for the simulation of inversions and sexually antagonistic alleles at infinite loci, with uniform rates of mutation, crossing over, and gene conversion (per bp, per generation) over a number of independent chromosome arms. The dynamic we were interested in modeling requires direct accounting of pleiotropy between different components of fitness, crossing over and gene conversion within inversions (as well as the possibility of inversion fixation), and directly modeling male display and female choice, and no existing simulator provided this functionality. These simulations were used to establish the potential for the model to occur in a population generally, and not to compare against specific empirical observations. We have used literature estimates of mutation rate and conversion rate of Zambian *D. melanogaster* for these simulations (Comeron et al. 2012; Huang et al. 2016) but have no information on the joint distribution of effects of mutations on our model of survival and display in *D. melanogaster*. So, we sample from relatively broad but arbitrary combinations of mutation effects, only assuming that there is likely some degree of antagonistic pleiotropy naturally present in genomic variation.

852 In SAIsim, a population is instantiated as a python object, and populated with individuals (also python objects)
853 based on provided (or initially empty) genomes. In each generation, as diagrammed in Figure 5, the simulator
854 calculates a survival probability for each individual as the product of the survival values of each variant they carry,
855 which range [0,1]. Here, multiplicative selection for survival prevents biologically unreasonable negative survival
856 probabilities from occurring. Male and female survival costs can each be included in a sex-specific manner, but
857 only globally in the simulation, not per mutation. For each offspring in the next generation, a mother is selected
858 randomly, weighted by individual survival probabilities, with replacement. To model the encountered males that
859 might become the father of an offspring individual from this female, a set number of males (*i.e.* the uniformly
860 specified encounter number) is randomly sampled from the full population of males, weighted by their survival
861 probabilities. The males each have a genetic display quality given as the sum of the reproductive values of their
862 alleles. A noise parameter, normally distributed around zero with a standard deviation of one, is added to each
863 male display value during each encounter. The male with the highest noise-adjusted display quality among the
864 encountered males is chosen to be the father. As offspring are generated, crossover locations are sampled, and
865 resampled in cases where the resultant gamete would be aneuploid. Resampling aneuploids removes the fecundity
866 cost that may exist in some taxa, but was chosen to reflect the lack of observed fecundity effects for most
867 paracentric inversions of *D. melanogaster*. Gene conversion is modeled as a constant probability per heterozygous
868 variant, with each event independently affecting just one locus, rather than modeling a conversion tract length
869 (given that gene conversion tracts are generally short and each would typically contain only a single modeled
870 variant in our simulations). This probability reflects the likelihood of any gene conversion event affecting a single
871 nucleotide site, rather than the per-bp rate of initiation of gene conversion tracts. The dynamics of gene
872 conversion are expected to have separate effects on the assembly and the maintenance of arrangements. The
873 assembly of an arrangement is positively dependent on the population-scaled rate of recombination of alleles as
874 new polymorphisms start at low frequency and often not on their optimal haplotype. However, the indirect fitness
875 effect of the inversion is dependent on its effect on recombinant offspring, and therefore negatively depends on
876 the per-individual rate of conversion. The simulation is intended to represent *D. melanogaster* life history and
877 allows removal of recombination in males, which also removes gene conversion in males by default. Finally,

878 inversions and new mutations are potentially added to the resultant gamete, with inversions resampled when
879 overlapping the coordinates of an inversion already present on the same individual chromosome. SAIsim also
880 allows the specification of mutation rates and the distribution of effect sizes of new mutations.

881 The simulations presented in Figures 5, 6, and 7 rescale parameter values by a factor of 2,000 to model an
882 estimated ancestral Ne of around 2×10^6 (Sprengelmeyer et al. 2020) while using a smaller simulated population of
883 1,000 individuals. The low crossover rate simulations model a 0.0218 cM region and the high crossover rate
884 simulations model 2.18 cM, which scale to 43.6 cM and 43.6 M female meiotic maps in the simulated populations.
885 Based on an average autosomal recombination rate in females of 2.18×10^{-8} for *D. melanogaster* (Comeron et al.
886 2012), these represent 10kb and 1Mb regions. Gene conversion occurs at each heterozygous variant with a
887 probability of 0.1295, based on scaling $\gamma=1.25 \times 10^{-7}$ events/bp/female meiosis with a conversion track length of
888 518 bp (Comeron et al. 2012) to a per-single-nucleotide-variant estimate of 6.475×10^{-5} conversions/bp/female
889 meiosis. Sexually antagonistic mutations were modeled at a rate corresponding to one per thousand mutations,
890 and so based on an autosomal estimate of 5.21×10^{-9} mutations/bp/gen on autosomes (Huang et al. 2016), and
891 modeling 10kb and 1Mb regions, these were scaled to 1.042×10^{-4} and 1.042×10^{-2} events per gamete
892 respectively. Inversions were approximated as arising once per hundred sexually antagonistic mutations, for 1.042
893 $\times 10^{-6}$ and 1.042×10^{-4} events per gamete. Both of these rates are empirically underdetermined. The simulations
894 use mutations with uniformly distributed quality values in the range [0,1]; although distributions skewed toward
895 small effects may be more realistic, our uniform approach avoids the need to invoke a specific but empirically
896 unknown distribution. SAIsim's object-oriented design allows specification of a population model by a new user
897 familiar with the basics of python scripting. A population object can be parameterized and populated in a few lines
898 of python script, and then parameters can be changed between periods of simulation, allowing bottlenecks and
899 demographic changes as well as changes in mutational parameters between generations. Replicate SAIsim
900 simulations were performed via the UW-Madison Center For High Throughput Computing (CHTC), which manages
901 cycle servers using the HTCondor distributed computing project (Thain et al. 2005; Erlandson and Theisen 2018).

4.2 Experimental Populations and Sequencing

To act as a high-inversion source population for the potential fathers in our experiment, we combined 15 males and 15 females from each of 10 inbred strains derived from wild flies collected in Zambia. These 10 strains were chosen to collectively carry the following common inversions: *In(2L)t*, *In(2R)NS*, *In(3L)Ok, In(3R)K*. We then collected at least 500 F2 males from this high-inversion population to use in each of our independent maternal line experiments, and these males were expected to carry a wide range of inversion genotypes. Each batch of high-inversion males was crossed to an equal number of virgin females from a single inbred Zambian strain (Figure 6). Four such experiments were performed, each using one inversion-free maternal strain. Since no inversions could be inherited through the mothers, inversion frequencies of successfully reproducing F2 males could be inferred from their pooled offspring. Unless otherwise noted, all flies were kept in a lab space of 23°C with around a degree of temperature fluctuation and without a controlled day/night cycle. All fly food used was prepared in batches of 4.5 L water, 500 mL cornmeal, 500 mL molasses, 200 mL yeast, 54 g agar, 20 mL propionic acid, and 45 mL tegosept 10% (in 95% ethanol). The four parental populations were each allowed to mate freely in clear plexiglass cages of dimension 19.3 by 19.3 by 30.0 cm, with a loose mesh covering on one end for access by hand, twisted and clipped shut with a binder clip. The parents were given either fly food bottles or grape agar in 7.9 cm petri dishes as food or laying substrate on alternating days over two weeks, after which all surviving male parents were counted and frozen. The grape agar was prepared in batches of 500 mL from FlyStuff.com grape agar powder, and wet Red Star yeast was brushed over the agar plates before use to encourage laying. Embryos were collected from the agar, counted, and frozen for sequencing. Embryos laid in the bottles were allowed to hatch and develop into adults. These adult offspring were separated by sex and by the time after laying at which they eclosed. In our husbandry conditions, the earliest eclosion was 11 days after laying, but this was rare and a first eclosion at 12 days was more common among the bottles. Most flies took several days longer to develop, and we considered flies collected in the first 6 days of a 10 day collection window to be the early-eclosing pool, in order to split the pool sample sizes closer to evenly. The collected adult flies were then aged by sequentially transferring them between bottles weekly for approximately 2.5 months, and then the adult offspring were sexed, counted, and collected. A 10-week

927 aging window was chosen based on the time taken for a control set of four bottles of flies from the paternal
928 outbred population to reach 80% mortality. All adult and embryo samples were frozen at -80°C before DNA
929 extraction. This sampling scheme was chosen to identify inversion frequencies in the set of potential parents, in
930 embryonic offspring, and in older adults, and so to assess potentially selected changes in inversion frequency
931 across male reproduction or in egg to adult (encompassing viability plus longevity). We chose to age the
932 populations to include longevity as long term survival is relevant both in the estimated reproductive age of the
933 species (Pool 2015), and the need to survive during periods of unsuitable environments, often while under
934 reproductive diapause (Saunders et al. 1990; Ragland et al. 2019). In seasonally varying climates, either temperate
935 or with long dry season, survival through challenging conditions is expected to require several months. In many
936 such cases females are in reproductive diapause, and so longevity is the main selective pressure.

937 DNA for each replicate age group was extracted by a phenol-chloroform extraction protocol used previously in
938 preparing DNA for the Drosophila Genome Nexus and *Drosophila* Population Genomics Project. The protocol is
939 similar to protocol 'BPC' used in DPGP phase 1 (Langley et al. 2012), except with the addition of proteinase-K
940 during homogenization of adult flies, and generally based on the common quick preparation protocol presented in
941 *Drosophila Protocols* (Huang et al. 2009). Inversion genotyping was performed by amplicon sequencing, where
942 amplicons associated with each inversion were designed to target at least one single nucleotide polymorphism in
943 the inversion interior that is fixed for alternate alleles between the inversion karyotypes, while still sharing
944 conserved primer regions across karyotypes to prevent primer bias. Genetic variation in these inversions was
945 assessed using previously sequenced and inversion-called haploid genomes from the Zambian *D. melanogaster*
946 population, available from the *Drosophila* Genome Nexus (Lack et al. 2016a). The amplicon region and primer
947 selection was performed with a custom script. The script calls Primer3 version 2.4.0 for primer candidate
948 generation (Koressaar and Remm 2007; Untergasser et al. 2012; Koressaar et al. 2018). Primers thus selected were
949 ordered from IDT as oligos ligated to Illumina stubby adaptors [Table S1]. Illumina sequencing libraries were then
950 prepared for each sample following the Illumina metagenomic amplicon sequencing protocol (Illumina 2013), with
951 50 ng/uL diluted DNA samples. and sequenced 2x150 on an Illumina NovaSeq to at least a depth of 20000 reads.

952 Depth was selected to ensure high likelihood of detecting a 0.025 change in inversion frequency from an initial
953 frequency of 0.10, based on simulated resampling of the collected flies, DNA extraction, and sequencing.

954 4.3 Sequence Analysis and Statistics

955 Sequence preparation and analysis were performed using custom pipeline scripts. Read trimming, alignment, and
956 QC were performed using bwa version 0.7.17-r1188 (Li 2013), GATK version 3.4-46 (Van der Auwera and O'Connor
957 2020), and Picard version 1.79 (Broad Institute 2019). Reads were trimmed at the ends to remove segments with
958 base quality less than 20, and read pairs were filtered to retain only those pairs where both reads aligned with
959 mapping quality >20. SNP identity at the sites of interest was called from those remaining reads with base
960 quality >20 at the SNP of interest. Due to the presence of chimeric reads in the amplicon library sequencing, we
961 chose to use a single discriminant SNP to determine each inversion's frequency in the read pool, instead of
962 incorporating haplotype or additional SNP information.

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

972 To generate the neutral distribution of the chi-square statistic, we took the observed paternal inversion
973 frequencies and then sampled from them to obtain new paternal and offspring read counts to represent a single
974 neutral data set as follows. First, we resampled the collected allelic counts (given in Tables S6 and S7) to account
975 for sampling variance between replicates. For the paternal samples, we needed to account for pre-sequencing

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

999 For a pairwise comparison of inversion frequency between two samples, we calculated the same allele count chi-
1000 square statistic as described above for the empirical data for each resampled set of allele counts, generating the
1001 full distribution of neutral chi-square test statistics from which we could identify extreme value cutoffs. We then

1002 obtained a p-value from the proportion of resampled replicates in which the chi-square statistic was greater than
1003 or equal to the empirical value. If the test was directional, we took the proportion of resampled replicates with
1004 both a chi-square statistic greater than or equal to the empirical value and a frequency difference in the expected
1005 direction.

1006 In order to identify a frequency change reversal between life history stages that could reflect a fitness tradeoff, we
1007 tested for the joint significance of two opposing frequency changes using three cohorts. A tradeoff would require a
1008 frequency change in one direction between paternal and embryo samples, followed by a frequency change in the
1009 opposite direction between embryo and adult offspring samples. These changes both depend on the shared
1010 embryonic cohort in the test, and so are not independent. Therefore, we evaluated the two inversion frequency
1011 changes jointly from a set of resampled neutral data sets in which the three cohorts arose from the same
1012 resampling process. Here we considered how likely it was that a sample of the three cohorts from the associated
1013 estimated neutral model had chi-square values for both the comparison of the fathers to embryos and the
1014 embryos to combined adults that were equivalent or more extreme in the test direction than the observed data.
1015 For example, when testing the significance of an observed paternal-embryo increase and then embryo-adult
1016 decrease in frequency, our p-value would be defined by the proportion of resampled replicates in which the
1017 frequency changes were in the tested directions of increase then decrease, and their chi-square values were both
1018 more extreme than the empirical values.

1019 To assess the significance of inversion frequency changes considered across all four maternal line crosses at once,
1020 we used Fisher's combined probability test. We applied this to both particular three-cohort frequency change
1021 reversals and to the set of two-cohort frequency changes. The inversion trajectory is independent between
1022 maternal crosses, but the differences between each cross in maternal line and initial paternal genetic variation
1023 make the replicates difficult to compare in a more structured way, so we felt this was an appropriate test for
1024 combined significance. In assessing our tradeoff hypothesis, we applied this test only for inversions where it was
1025 well-motivated; *i.e.* where the averaged frequency change across maternal crosses was in opposite directions
1026 between paternal/embryo and embryo/adult comparisons.

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

Data Availability

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

Acknowledgements

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

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

This research was performed using the compute resources and assistance of the UW-Madison Center for High Throughput Computing (CHTC) in the Department of Computer Sciences. The CHTC is supported by UW-Madison,

the Advanced Computing Initiative, the Wisconsin Alumni Research Foundation, the Wisconsin Institutes for Discovery, and the National Science Foundation, and is an active member of the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

1052 ## Works Cited

1053 Altenberg L, Feldman MW. 1987. Selection, generalized transmission and the evolution of modifier genes. I. The
1054 reduction principle. *Genetics* 117:559–572.

1055 Altenberg L, Liberman U, Feldman MW. 2017. Unified reduction principle for the evolution of mutation, migration,
1056 and recombination. *Proc. Natl. Acad. Sci.* 114:E2392–E2400.

1057 Audet T, Krol J, Pelletier K, Stewart AD, Dworkin I. 2024. Sexually discordant selection is associated with trait
1058 specific morphological changes and a complex genomic response. *Evolution*:qpae071.

1059 Aulard S, David JR, Lemeunier F. 2002. Chromosomal inversion polymorphism in Afrotropical populations of
1060 *Drosophila melanogaster*. *Genet. Res.* 79:49–63.

1061 Babcock CS, Anderson WW. 1996. Molecular evolution of the Sex-Ratio inversion complex in *Drosophila*
1062 *pseudoobscura*: analysis of the Esterase-5 gene region. *Mol. Biol. Evol.* 13:297–308.

1063 Bastide H, Ogereau D, Montchamp-Moreau C, Gérard PR. 2022. The fate of a suppressed X-linked meiotic driver:
1064 experimental evolution in *Drosophila simulans*. *Chromosome Res.* 30:141–150.

1065 Bastide H, Yassin A, Johanning EJ, Pool JE. 2014. Pigmentation in Drosophila melanogaster reaches its maximum in
1066 Ethiopia and correlates most strongly with ultra-violet radiation in sub-Saharan Africa. *BMC Evol. Biol.*
1067 14:179.

1068 Bateman AJ. 1948. Intra-sexual selection in *Drosophila*. *Heredity* 2:349–368.

1069 Behrman EL, Watson SS, O'Brien KR, Heschel MS, Schmidt PS. 2015. Seasonal variation in life history traits in two
1070 *Drosophila* species. *J. Evol. Biol.* 28:1691–1704.

1071 Berdan EL, Barton NH, Butlin R, Charlesworth B, Faria R, Fragata I, Gilbert KJ, Jay P, Kapun M, Lotterhos KE, et al.
1072 2023. How chromosomal inversions reorient the evolutionary process. J. Evol. Biol. 36:1761–1782.

1073 Betrán E, Santos M, Ruiz A. 1998. Antagonistic Pleiotropic Effect of Second-Chromosome Inversions on Body Size
1074 and Early Life-History Traits in *Drosophila buzzatii*. *Evolution* 52:144–154.

1075 Blackmon H, Brandvain Y. 2017. Long-Term Fragility of Y Chromosomes Is Dominated by Short-Term Resolution of
1076 Sexual Antagonism. *Genetics* 207:1621–1629.

1077 Blows MW, Brooks R. 2003. Measuring Nonlinear Selection. *Am. Nat.* 162:815–820.

1078 Broad Institute. 2019. Picard Toolkit. Available from: http://broadinstitute.github.io/picard/

1079 Charlesworth B. 1974. Inversion polymorphism in a two-locus genetic system. *Genet. Res.* 23:259–280.

1080 Charlesworth B. 1997. The evolution of lethals in the t-haplotype system of the mouse. *Proc. R. Soc. B Biol. Sci.*
1081 258:101–107.

1082 Charlesworth B. 2024. The fitness consequences of genetic divergence between polymorphic gene arrangements.
1083 Genetics 226:iyad218.

1084 Charlesworth B, Barton NH. 2018. The Spread of an Inversion with Migration and Selection. *Genetics* 208:377–382.

1085 Charlesworth B, Charlesworth D. 1973. Selection of new inversions in multi-locus genetic systems. *Genet. Res.*
1086 21:167–183.

1087 Charlesworth B, Charlesworth D. 2010. Elements of Evolutionary Genetics. W. H. Freeman

1088 Charlesworth B, Hughes KA. 2000. The maintenance of genetic variation in life-history traits. In: Singh RS, Krimbas
1089 CB, editors. Evolutionary Genetics: from Molecules to Morphology. Cambridge, U.K.; New York, NY, USA:
1090 Cambridge University Press. p. 369–392.

1091 Charlesworth D, Charlesworth B. 1980. Sex differences in fitness and selection for centric fusions between sex-
1092 chromosomes and autosomes. *Genet. Res.* 35:205–214.

1093 Chen KH-C, Chatla K, Bachtrog D. 2024. Single-cell RNA-seq of *Drosophila miranda* testis reveals the evolution and
1094 trajectory of germline sex chromosome regulation. *PLOS Genet.* 22(4): e3002605.

1095 Cheng C, Kirkpatrick M. 2016. Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies. *PLOS*
1096 *Genet.* 12:e1006170.

1097 Clarke CA, Sheppard PM. 1960. Super-genes and mimicry. *Heredity* 14:175–185.

1098 Comeron JM, Ratnappan R, Bailin S. 2012. The Many Landscapes of Recombination in Drosophila melanogaster.
1099 *PLOS Genet.* 8:e1002905.

1100 Connallon T, Chenoweth SF. 2019. Dominance reversals and the maintenance of genetic variation for fitness. *PLOS*
1101 *Biol.* 17:e3000118.

1102 Connallon T, Clark AG. 2011. The resolution of sexual antagonism by gene duplication. *Genetics* 187:919–937.

1103 Cook LM. 2013. Selection and disequilibrium in Cepaea nemoralis. *Biol. J. Linn. Soc.* 108:484–493.

1104 Corbett-Detig RB. 2016. Selection on Inversion Breakpoints Favors Proximity to Pairing Sensitive Sites in Drosophila
1105 melanogaster. *Genetics* 204:259–265.

1106 Corbett-Detig RB, Hartl DL. 2012. Population Genomics of Inversion Polymorphisms in Drosophila melanogaster.
1107 *PLOS Genet.* 8:e1003056.

1108 Courret C, Chang C-H, Wei KH-C, Montchamp-Moreau C, Larracuente AM. 2019. Meiotic drive mechanisms: lessons
1109 from Drosophila. *Proc. R. Soc. B Biol. Sci.* 286:20191430.

1110 Coyne JA, Meyers W, Crittenden AP, Sniegowski P. 1993. The Fertility Effects of Pericentric Inversions in Drosophila
1111 Melanogaster. *Genetics* 134:487–496.

1112 Cox RM, Calsbeek R. 2009. Sexually Antagonistic Selection, Sexual Dimorphism, and the Resolution of Intralocus
1113 Sexual Conflict. *Am. Nat.* 173:176–187.

1114 da Silva Ribeiro T, Galván JA, Pool JE. 2022. Maximum SNP FST Outperforms Full-Window Statistics for Detecting
1115 Soft Sweeps in Local Adaptation. *Genome Biol. Evol.* 14:evac143.

1116 Dickerson GE. 1955. Genetic slippage in response to selection for multiple objectives. *Cold Spring Harbor Symposia*
1117 *on Quantitative Biology* 20:213–224.

1118 Dobzhansky TG. 1943. Genetics of Natural Populations IX. Temporal Changes in the Composition of Populations of
1119 *Drosophila pseudoobscura*. *Genetics* 28:162–186.

1120 Dobzhansky TG. 1949. Observations and Experiments on Natural Selection in Drosophila. *Hereditas* 35:210–224.

1121 Dobzhansky TG. 1950. Genetics of natural populations. XIX. Origin of heterosis through natural selection in
1122 populations of *Drosophila pseudoobscura*. *Genetics* 35:288–302.

1123 Dobzhansky TG. 1970. Genetics of the Evolutionary Process. Columbia University Press

1124 van Doorn GS, Kirkpatrick M. 2007. Turnover of sex chromosomes induced by sexual conflict. *Nature* 449:909–912.

1125 Erlandson E, Theisen T. 2018. htcondor: HTCondor source repository, formerly the Condor Project. Available from:
1126 https://github.com/htcondor/htcondor

1127 Faria R, Johannesson K, Butlin RK, Westram AM. 2019. Evolving Inversions. *Trends Ecol. Evol.* 34:239–248.

1128 Feldman MW. 1972. Selection for Linkage Modification: I. Random Mating Populations. *Theor. Pop. Biol.* 3:324–
1129 346.

1130 Feldman MW, Balkau B. 1973. Selection for Linkage Modification: II. A Recombination Balance for Neutral
1131 Modifiers. *Genetics* 74:713–726.

1132 Feldman MW, Christiansen FB, Brooks LD. 1980. Evolution of recombination in a constant environment. *Proc. Natl.*
1133 *Acad. Sci.* 77:4838–4841.

1134 Frydenberg O. 1963. Population Studies of a Lethal Mutant in *Drosophila melanogaster*. *Hereditas* 50:89–116.

1135 Gilbert KJ, Pouyet F, Excoffier L, Peischl S. 2020. Transition from Background Selection to Associative
1136 Overdominance Promotes Diversity in Regions of Low Recombination. *Curr. Biol.* 30:101-107.e3.

1137 Gillespie JH. 1998. Population Genetics: A Concise Guide. Johns Hopkins University Press

1138 Giraldo-Deck LM, Loveland JL, Goymann W, Tschirren B, Burke T, Kempenaers B, Lank DB, Küpper C. 2022.
1139 Intralocus conflicts associated with a supergene. *Nat. Commun.* 13:1384.

1140 Glaser-Schmitt A, Ramnarine TJS, Parsch J. 2023. Rapid evolutionary change, constraints and the maintenance of
1141 polymorphism in natural populations of *Drosophila melanogaster*. *Mol. Ecol.*

1142 Gong WJ, McKim KS, Hawley RS. 2005. All Paired Up with No Place to Go: Pairing, Synapsis, and DSB Formation in a
1143 Balancer Heterozygote. *PLOS Genet.* 1:e67.

1144 Gowaty PA, Kim Y-K, Anderson WW. 2012. No evidence of sexual selection in a repetition of Bateman's classic
1145 study of Drosophila melanogaster. *Proc. Natl. Acad. Sci. U. S. A.* 109:11740–11745.

1146 Grieshop K, Arnqvist G. 2018. Sex-specific dominance reversal of genetic variation for fitness. *PLOS Biol.*
1147 16:e2006810.

1148 Hazel LN. 1943. The Genetic Basis for Constructing Selection Indexes. *Genetics* 28:476–490.

1149 Hearn KE, Koch EL, Stankowski S, Butlin RK, Faria R, Johannesson K, Westram AM. 2022. Differing associations
1150 between sex determination and sex-linked inversions in two ecotypes of Littorina saxatilis. *Evolution*
1151 *Letters* 6:358–374.

1152 Hoffmann AA, Rieseberg LH. 2008. Revisiting the Impact of Inversions in Evolution: From Population Genetic
1153 Markers to Drivers of Adaptive Shifts and Speciation? *Annu. Rev. Ecol. Evol. Syst.* 39:21–42.

1154 Hoquet T, Bridges WC, Gowaty PA. 2020. Bateman's Data: Inconsistent with "Bateman's Principles." *Ecol. Evol.*
1155 10:10325–10342.

1156 Huang AM, Rehm EJ, Rubin GM. 2009. Quick Preparation of Genomic DNA from Drosophila. *Cold Spring Harb.*
1157 *Protoc.* 2009:pdb.prot5198.

1158 Huang W, Lyman RF, Lyman RA, Carbone MA, Harbison ST, Magwire MM, Mackay TF. 2016. Spontaneous
1159 mutations and the origin and maintenance of quantitative genetic variation. Przeworski M, editor. *eLife*
1160 5:e14625.

1161 Illumina. 2013. 16S metagenomic sequencing library preparation guide. Available from:
1162 https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-
1163 metagenomic-library-prep-guide-15044223-b.pdf

1164 Ingleby FC, Flis I, Morrow EH. 2015. Sex-Biased Gene Expression and Sexual Conflict throughout Development. *Cold*
1165 *Spring Harb. Perspect. Biol.* 7:a017632.

1166 Innocenti P, Morrow EH. 2010. The Sexually Antagonistic Genes of Drosophila melanogaster. *PLOS Biol.*
1167 8:e1000335.

1168 Inoue Y. 1979. THE FATE OF POLYMORPHIC INVERSIONS OF *DROSOPHILA MELANOGASTER* TRANSFERRED TO
1169 LABORATORY CONDITIONS. *Jpn. J. Genet.* 54:83–96.

1170 John A, Vinayan K, Varghese J. 2016. Achiasmy: Male Fruit Flies Are Not Ready to Mix. *Front. Cell Dev. Biol.* 4:75.

1171 Jones AG, Arguello JR, Arnold SJ. 2002. Validation of Bateman's principles: a genetic study of sexual selection and
1172 mating patterns in the rough-skinned newt. *Proc. R. Soc. B Biol. Sci.* 269:2533–2539.

1173 Kapun M, Fabian DK, Goudet J, Flatt T. 2016. Genomic Evidence for Adaptive Inversion Clines in Drosophila
1174 melanogaster. *Mol. Biol. Evol.* 33:1317–1336.

1175 Kapun M, Flatt T. 2019. The adaptive significance of chromosomal inversion polymorphisms in Drosophila
1176 melanogaster. *Mol. Ecol.* 28:1263–1282.

1177 Kapun M, Mitchell ED, Kawecki TJ, Schmidt P, Flatt T. 2023. An Ancestral Balanced Inversion Polymorphism Confers
1178 Global Adaptation. *Mol. Biol. Evol.* 40:msad118.

1179 Kapun M, van Schalkwyk H, McAllister B, Flatt T, Schlötterer C. 2014. Inference of chromosomal inversion dynamics
1180 from Pool-Seq data in natural and laboratory populations of Drosophila melanogaster. *Mol. Ecol.*
1181 23:1813–1827.

1182 Kirkpatrick M, Barton N. 2006. Chromosome Inversions, Local Adaptation and Speciation. *Genetics* 173:419–434.

1183 Kopp M, Hermisson J. 2006. The evolution of genetic architecture under frequency-dependent disruptive selection.
1184 *Evolution* 60:1537–1550.

1185 Koury SA. 2023. Predicting recombination suppression outside chromosomal inversions in Drosophila
1186 melanogaster using crossover interference theory. *Heredity* 130:196–208.

1187 Koressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M. 2018. Primer3_masker: integrating masking
1188 of template sequence with primer design software. *Bioinforma. Oxf. Engl.* 34:1937–1938.

1189 Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinforma. Oxf.*
1190 *Engl.* 23:1289–1291.

1191 Krimbas CB, Powell JR. 1992. Introduction. In: Krimbas CB, Powell JR, editors. Drosophila Inversion Polymorphism.
1192 Boca Raton, FL: CRC Press. p. 1–52.

1193 Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR. 2014. doublesex
1194 is a mimicry supergene. *Nature* 507:229–232.

1195 Küpper C, Stocks M, Risse JE, Remedios N dos, Farrell LL, McRae SB, Morgan TC, Karlionova N, Pinchuk P, Verkuil YI,
1196 et al. 2016. A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.*
1197 48:79–83.

1198 Lack JB, Lange JD, Tang AD, Corbett-Detig RB, Pool JE. 2016. A Thousand Fly Genomes: An Expanded Drosophila
1199 Genome Nexus. *Mol. Biol. Evol.* 33:3308–3313.

1200 Lack JB, Yassin A, Sprengelmeyer QD, Johanning EJ, David JR, Pool JE. 2016. Life history evolution and cellular
1201 mechanisms associated with increased size in high-altitude Drosophila. *Ecol. Evol.* 6:5893–5906.

1202 Lange JD, Bastide H, Lack JB, Pool JE. 2022. A Population Genomic Assessment of Three Decades of Evolution in a
1203 Natural Drosophila Population. *Mol. Biol. Evol.* 39:msab368.

1204 Langley CH, Stevens K, Cardeno C, Lee YCG, Schrider DR, Pool JE, Langley SA, Suarez C, Corbett-Detig RB,
1205 Kolaczkowski B, et al. 2012. Genomic Variation in Natural Populations of Drosophila melanogaster.
1206 *Genetics* 192:533–598.

1207 Lemeunier F, Aulard S (1992). Inversion polymorphism in *Drosophila melanogaster*. In: Krimbas CB, Powell JR
1208 (eds) *Drosophila Inversion Polymorphism*. CRC Press: Boca Raton, FL pp 339–405.

1209 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Available from:
1210 http://arxiv.org/abs/1303.3997

1211 Libbrecht R, Kronauer DJC. 2014. Convergent Evolution: The Genetics of Queen Number in Ants. *Curr. Biol.*
1212 24:R1083–R1085.

1213 Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al.
1214 2015. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer
1215 Interactions. *Cell* 161:1012–1025.

1216 Machado HE, Bergland AO, Taylor R, Tilk S, Behrman E, Dyer K, Fabian DK, Flatt T, González J, Karasov TL, et al.
1217 2021. Broad geographic sampling reveals the shared basis and environmental correlates of seasonal
1218 adaptation in Drosophila. *eLife* 10:e67577.

1219 McBroome J, Liang D, Corbett-Detig RB. 2020. Fine-Scale Position Effects Shape the Distribution of Inversion
1220 Breakpoints in Drosophila melanogaster. *Genome Biol. Evol.* 12:1378–1391.

1221 Mérot C, Llaurens V, Normandeau E, Bernatchez L, Wellenreuther M. 2020. Balancing selection via life-history
1222 trade-offs maintains an inversion polymorphism in a seaweed fly. *Nat. Commun.* 11:670.

1223 Mettler E, Voelker RA, Mukai T. 1977. Inversion Clines in Populations of Drosophila melanogaster. *Genetics* 87:8.

1224 Morgan TH. 1910. Sex Limited Inheritance in Drosophila. *Science* 32:120–122.

1225 Mroczek RJ, Melo JR, Luce AC, Hiatt EN, Dawe RK. 2006. The Maize Ab10 Meiotic Drive System Maps to
1226 Supernumerary Sequences in a Large Complex Haplotype. *Genetics* 174:145–154.

1227 Mukai T, Yamaguchi O. 1974. The genetic structure of natural populations of Drosophila melanogaster. XI. Genetic
1228 variability in a local population. *Genetics* 76:339–366.

1229 Munasinghe M, Brandvain Y. 2023. Together Inbreeding and Reproductive Compensation Favor Lethal t-
1230 Haplotypes. bioRxiv 2023.07.26.550691. Available from:
1231 https://www.biorxiv.org/content/10.1101/2023.07.26.550691v1

1232 Nassar R, Muhs HJ, Cook RD. 1973. Frequency-dependent selection at the Payne inversion in *Drosophila*
1233 *melanogaster. Evolution* 27:558-564.

1234 O'Donald P, Majerus MEN, Clarke BC, Partridge L, Robertson A, Clarke BC, Partridge L. 1997. Frequency-dependent
1235 sexual selection. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 319:571–586.

1236 Ohta T, Kimura M. 1970. Development of associative overdominance through linkage disequilibrium in finite
1237 populations. *Genet. Res.* 16:165–177.

1238 Otto SP, Lenormand T. 2002. Resolving the paradox of sex and recombination: Evolution of sex. *Nat. Rev. Genet.*
1239 3:252–261.

1240 Otwinowski J, McCandlish DM, Plotkin JB. 2018. Inferring the shape of global epistasis. *Proc. Natl. Acad. Sci.*
1241 115:E7550–E7558.

1242 Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadía-Cardoso A, Anderson EC, Rundio DE, Williams TH,
1243 Naish KA, et al. 2019. Sex-dependent dominance maintains migration supergene in rainbow trout. *Nat.*
1244 *Ecol. Evol.* 3:1731–1742.

1245 Pei Y, Forstmeier W, Knief U, Kempenaers B. 2023. Weak antagonistic fitness effects can maintain an inversion
1246 polymorphism. *Mol. Ecol.* 32:3575–3585.

1247 Pennell MW, Kirkpatrick M, Otto SP, Vamosi JC, Peichel CL, Valenzuela N, Kitano J. 2015. Y Fuse? Sex Chromosome
1248 Fusions in Fishes and Reptiles. *PLOS Genet.* 11:e1005237.

1249 Pool JE. 2015. The Mosaic Ancestry of the Drosophila Genetic Reference Panel and the D. melanogaster Reference
1250 Genome Reveals a Network of Epistatic Fitness Interactions. *Mol Biol Evol* 32:3236–3251.

1251 Pool JE, Braun DT, Lack JB. 2017. Parallel Evolution of Cold Tolerance within Drosophila melanogaster. *Mol. Biol.*
1252 *Evol.* 34:349–360.

1253 Poul YL, Whibley A, Chouteau M, Prunier F, Llaurens V, Joron M. 2014. Evolution of dominance mechanisms at a
1254 butterfly mimicry supergene. *Nat. Commun.* 5:5644.

1255 Ragland GJ, Armbruster PA, Meuti ME. 2019. Evolutionary and functional genetics of insect diapause: a call for
1256 greater integration. *Curr Opin Insect Sci* 36:74–81.

1257 Reinhardt JA, Kolaczkowski B, Jones CD, Begun DJ, Kern AD. 2014. Parallel Geographic Variation in Drosophila
1258 melanogaster. *Genetics* 197:361–373.

1259 Reinhart M, Carney T, Clark AG, Fiumera AC. 2015. Characterizing Male–Female Interactions Using Natural Genetic
1260 Variation in Drosophila melanogaster. *J. Hered.* 106:67–79.

1261 Rest JS, Morales CM, Waldron JB, Opulente DA, Fisher J, Moon S, Bullaughey K, Carey LB, Dedousis D. 2013.
1262 Nonlinear Fitness Consequences of Variation in Expression Level of a Eukaryotic Gene. *Mol. Biol. Evol.*
1263 30:448–456.

1264 Rice WR. 1992. Sexually antagonistic genes: experimental evidence. *Science* 256:1436–1439.

1265 Rice WR. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nat.*
1266 *Lond.* 381:232–234.

1267 Rose MR. 1982. Antagonistic pleiotropy, dominance, and genetic variation. *Heredity* 48:63–78.

1268 Rowe L, Chenoweth SF, Agrawal AF. 2018. The Genomics of Sexual Conflict. *Am. Nat.* 192:274–286.

1269 Ruzicka F, Hill MS, Pennell TM, Flis I, Ingleby FC, Mott R, Fowler K, Morrow EH, Reuter M. 2019. Genome-wide
1270 sexually antagonistic variants reveal long-standing constraints on sexual dimorphism in fruit flies. *PLOS*
1271 *Biol.* 17:e3000244.

1272 Said I, Byrne A, Serrano V, Cardeno C, Vollmers C, Corbett-Detig RB. 2018. Linked genetic variation and not genome
1273 structure causes widespread differential expression associated with chromosomal inversions. *Proc. Natl.*
1274 *Acad. Sci.* 115:5492–5497.

1275 Saunders DS, Richard DS, Applebaum SW, Ma M, Gilbert LI. 1990. Photoperiodic diapause in *Drosophila*
1276 *melanogaster* involves a block to the juvenile hormone regulation of ovarian maturation. *Gen Comp*
1277 *Endocrinol* 79:174–184.

1278 Saunders PA, Neuenschwander S, Perrin N. 2019. Impact of deleterious mutations, sexually antagonistic selection,
1279 and mode of recombination suppression on transitions between male and female heterogamety. *Heredity*
1280 123:419–428.

1281 Schaeffer SW, Goetting-Minesky MP, Kovacevic M, Peoples JR, Graybill JL, Miller JM, Kim K, Nelson JG, Anderson
1282 WW. 2003. Evolutionary genomics of inversions in *Drosophila pseudoobscura*: Evidence for epistasis. *Proc.*
1283 *Natl. Acad. Sci.* 100:8319–8324.

1284 Sprengelmeyer QD, Mansourian S, Lange JD, Matute DR, Cooper BS, Jirle EV, Stensmyr MC, Pool JE. 2020.
1285 Recurrent Collection of Drosophila melanogaster from Wild African Environments and Genomic Insights
1286 into Species History. *Mol. Biol. Evol.* 37:627–638.

1287 Stearns FW. 2010. One Hundred Years of Pleiotropy: A Retrospective. *Genetics* 186:767–773.

1288 Sturtevant AH. 1921. A Case of Rearrangement of Genes in Drosophila. *Proc. Natl. Acad. Sci.* 7:235–237.

1289 Sturtevant AH, Beadle GW. 1936. The Relations of Inversions in the X Chromosome of *Drosophila melanogaster* to
1290 Crossing Over and Disjunction. *Genetics* 21:51.

1291 Sturtevant AH, Mather K. 1938. The Interrelations of Inversions, Heterosis and Recombination. *Am. Nat.* 72:447–
1292 452.

1293 Surmacki A, Ożarowska-Nowicka A, Rosin ZM. 2013. Color polymorphism in a land snail Cepaea nemoralis
1294 (Pulmonata: Helicidae) as viewed by potential avian predators. *Naturwissenschaften* 100:533–540.

1295 Sved JA. 1968. The Stability of Linked Systems of Loci with a Small Population Size. *Genetics* 59:543–563.

1296 Tatarenkov A, Healey CIM, Grether GF, Avise JC. 2008. Pronounced reproductive skew in a natural population of
1297 green swordtails, Xiphophorus helleri. *Mol. Ecol.* 17:4522–4534.

1298 Thain D, Tannenbaum T, Livny M. 2005. Distributed computing in practice: the Condor experience. *Concurr.*
1299 *Comput. Pract. Exp.* 17:323–356.

1300 Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities
1301 and interfaces. *Nucleic Acids Res.* 40:e115.

1302 Van der Auwera GA, O'Connor BD. 2020. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly
1303 Media, Incorporated

1304 Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang Y-C, Shoemaker D, Keller L. 2013. A Y-like social
1305 chromosome causes alternative colony organization in fire ants. *Nature* 493:664–668.

1306 Wellenreuther M, Bernatchez L. 2018. Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends Ecol. Evol.*
1307 33:427–440.

1308 White MJD. 1973. Animal Cytology and Evolution. 3rd ed. Cambridge University Press. Available from:
1309 https://wellcomecollection.org/works/meqhhay5

1310 Wittmann MJ, Bergland AO, Feldman MW, Schmidt PS, Petrov DA. 2017. Seasonally fluctuating selection can
1311 maintain polymorphism at many loci via segregation lift. *Proc. Natl. Acad. Sci.* 114:E9932–E9941.

1312 Wright AE, Dean R, Zimmer F, Mank JE. 2016. How to make a sex chromosome. *Nat. Commun.* 7:1–8.

1313 Yamamoto A, Huang W, Carbone MA, Anholt RRH, Mackay TFC. The Genetic Basis of Incipient Sexual Isolation in
1314 *Drosophila melanogaster*. bioRxiv 2024.03.01.582979. Available from:
1315 https://www.biorxiv.org/content/10.1101/2024.03.01.582979v1

1316 Yassin A, Bastide H, Chung H, Veuille M, David JR, Pool JE. 2016. Ancient balancing selection at tan underlies female
1317 colour dimorphism in Drosophila erecta. *Nat. Commun.* 7:10400.

1318 Zajitschek F, Connallon T. 2018. Antagonistic pleiotropy in species with separate sexes, and the maintenance of
1319 genetic variation in life-history traits and fitness. *Evolution* 72:1306–1316.

1320 Zhao L, Charlesworth B. 2016. Resolving the Conflict Between Associative Overdominance and Background
1321 Selection. *Genetics* 203:1315–1334.

1322