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. Additionally, all inversions show survival differences between sexes, and paternal success depends on maternal genotype. 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 parent heterozygous for an inversion; 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. Alternatively, crossovers may be inhibited directly in regions heterozygous for an inversion, or avoided by achiasmy or segregation of the aneuploid gametes 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 by inversion heterozygotes.

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), but many inversion arrangements are expected to be neutral in and of themselves (Said et al. 2018). So, in the absence of indirect selection on the recombination effects of inversions, inversion variants are expected to evolve similarly to single nucleotide mutations: to be selectively removed when the net direct effects are deleterious, to sweep when the effects are beneficial, to be fixed or lost in proportion to frequency when neutral, or to be selectively maintained when associated with a balanced trait. Given that a common effect of inversion heterozygotes is the deleterious generation of aneuploid gametes, and that studies of extant inversions and engineered inversions suggest that inversions may commonly have minimal effect on regulation or coding regions, many inversion rearrangements may confer either deleterious or neutral fitness consequences in the absence of association with other variants.

In contrast to the above predictions, several well-studied examples are known in which inversion variants are maintained at an intermediate frequency within a population, presumably due to the indirect effects of inversions on recombination. Simulation and analytic theory have demonstrated that inversions associated with a local population can reduce recombination between locally adapted variants and unfit migrants, and thereby increase in frequency by association with the fittest haplotype locally (Kirkpatrick and Barton 2006; Charlesworth and Barton 2018). Such locally adaptive inversions could potentially also link together favorable combinations of variants between epistatically interacting loci (Dobzhansky 1970; Schaeffer et al. 2003; Hoffmann and Rieseberg 2008). Unless migration rates are high, these models predict inversions to be at high frequency in their favored geographic range but at low frequency elsewhere. Alternately, chance associations of both an inversion and its standard arrangement counterpart with different recessive deleterious alleles may cause a form of overdominant selection, where heterozygotes carrying both main haplotypes experience none of the deleterious effects of a homozygous recessive. This associative overdominance can persist for extended periods, particularly when the rate of recombination between inversion arrangements is low (Zhao and Charlesworth 2016; Gilbert et al. 2020). This model predicts that such haplotypes will most readily accumulate in small populations, or when inversions that are initially rare accumulate recessive deleterious variants before the arrival of a linked positive variant or change in selection pressure. This may extend the polymorphic lifespan of an inversion, and remain stable in small populations, but is not expected to remain indefinitely stable as an intermediate frequency polymorphism in sufficiently large populations due to the greater population recombination rates allowing gene conversion and/or double-crossover events to generate haplotypes lacking deleterious variants. 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, and they are not always mutually exclusive.

1.2 Inversions in *Drosophila melanogaster*

Drosophila 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* (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) could allow inversion rearrangements to have limited fitness consequences. Further, genome editing tools have allowed for the creation of synthetic inversions, and the finding that these structural rearrangements themselves are not 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), the primary evolutionary relevance of these rearrangements appears to be the resulting reduction of recombinant offspring between otherwise highly recombining loci. Hence, *D. melanogaster* inversions would appear to be maintained for the sake of maintaining linkage associations between distant variants.

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

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

Second, even in the locations where inversions would hypothetically be ecologically advantageous, they rarely surpass intermediate frequencies (Pool et al. 2017; Sprengelmeyer et al. 2020). Under a local adaptation model, that observation would require a secondary explanation, such as very high levels of long distance gene flow or the fixation of recessive deleterious variants on inverted haplotypes, generating an associative overdominance at linked neutral loci (Frydenberg 1963; Sved 1968; Ohta and Kimura 1970; Zhao and Charlesworth 2016; Gilbert et al. 2020) that resists fixation. While migration may well limit inversion frequencies in some populations, we note that some putatively adaptive Single Nucleotide Polymorphism (SNP) variants do achieve much greater frequency differentials between populations (da Silva Ribeiro et al. 2022), and thus we would have to suppose that inversions are under relatively weaker selection to explain their lesser differentiation. Regarding the hypothesis of associative overdominance, we note that *D. melanogaster* populations do contain substantial recessive genetic load (Greenberg and Crow 1960), and it has been speculated that associative overdominance might boost inversion frequencies in *D. melanogaster* populations during founder events (Pool et al. 2012) or in isolated wilderness environments (Sprengelmeyer et al. 2020). However, it seems likely that in large human-commensal populations, *D. melanogaster* inversions would escape from linked deleterious variants through gene conversion or double crossover events with standard chromosomes, in light of both the antiquity of most of these inversions (Corbett-Detig and Hartl 2012) and the very high population recombination rates of *D. melanogaster*. Homozygous lines

have also been generated for each common inversion, demonstrating a lack of perfect linkage of the inversions with recessive lethal or infertile variants (*e.g.* Lack et al. 2016a).

While local adaptation, gene flow, and associative overdominance may all contribute to geographic variation in inversion frequency among *D. melanogaster* populations, it is not clear that these processes are collectively sufficient to explain observed patterns. Alternatively, the allele frequencies and diversity patterns of at least some of these inversions might be primarily shaped by some form of balancing selection acting on inversion-linked variation, in which the balanced frequency is dependent on the environment (Kapun et al. 2023). Balancing selection on inversions has received increasing attention in recent literature (Wellenreuther and Bernatchez 2018; Faria et al. 2019). However, there has not been a proposed mechanism whereby alleles at multiple linked loci undergoing balancing selection would directly benefit from linkage and thereby maintain an associated inversion polymorphism under indirect selection. Here we propose one such mechanism, in which alleles that contribute to naturally frequency-dependent traits like competitive mating display simultaneously contribute antagonistically to another trait, such as the survival to reproductive maturity.

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 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. The tendency of selection to favor the reduction of recombination when the haplotypes present are at a fitness optimum is well studied, and has been termed the 'reduction principle' (Altenberg et al. 2017).

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 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.

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. 2023) 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 may be the simplest case of antagonism to consider for a model of balancing selection, and one highly relevant to the biology of *D. melanogaster*.

Given the relevance of sexual selection and sexual antagonism to *Drosophila* evolution, we propose that polymorphic *D. melanogaster* inversion and standard arrangements may help maintain haplotypes of alleles that specify opposite ends of a sexually antagonistic tradeoff. We first explore this model through forward simulation,

to establish that it may in theory arise in a population and maintain inversion polymorphism over a significant number of generations. Then we examine the results of an experiment with an outbred laboratory population of *D. melanogaster*, to test whether common inversions show changes in frequency over the course of a generation in a manner consistent with a trade-off between survival and male reproductive success. Many such antagonistically pleiotropic scenarios are possible under different selective tradeoffs, but we focus on a specific scenario of sexual antagonism here.

2 Results

2.1 A Model of Sexually Antagonistic Inversions

Here, we consider a population featuring a competitive mate choice system that results in a skewed distribution of reproductive success among males. Within this population, we model a genomic region containing a number of linked loci, each with alternate antagonistic, pleiotropic alleles that trade off between male display success and survival likelihood (of females or both sexes). If mate choice is highly competitive, then superior display quality can enable males to mate successfully in excess of other display phenotypes they compete against. This gives a disproportionate benefit to additional mate competition quality if it allows a male to out-compete others, while additional survival costs stay proportional. Yet, fitness conveyed by 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. 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, until eventually it is equal in magnitude (across both sexes) to its pleiotropic survival cost – at which point an equilibrium frequency is reached. At this stage, the highest fitness among haplotypes may favor either of two extreme strategies - either producing among the most reproductively competitive males (by carrying a maximal number of antagonistic variants and paying the necessary cost of survival likelihood) or else paying none of the survival cost (at the expense of male reproductive competitiveness, carrying none of the antagonistic variants). In these conditions, we propose that

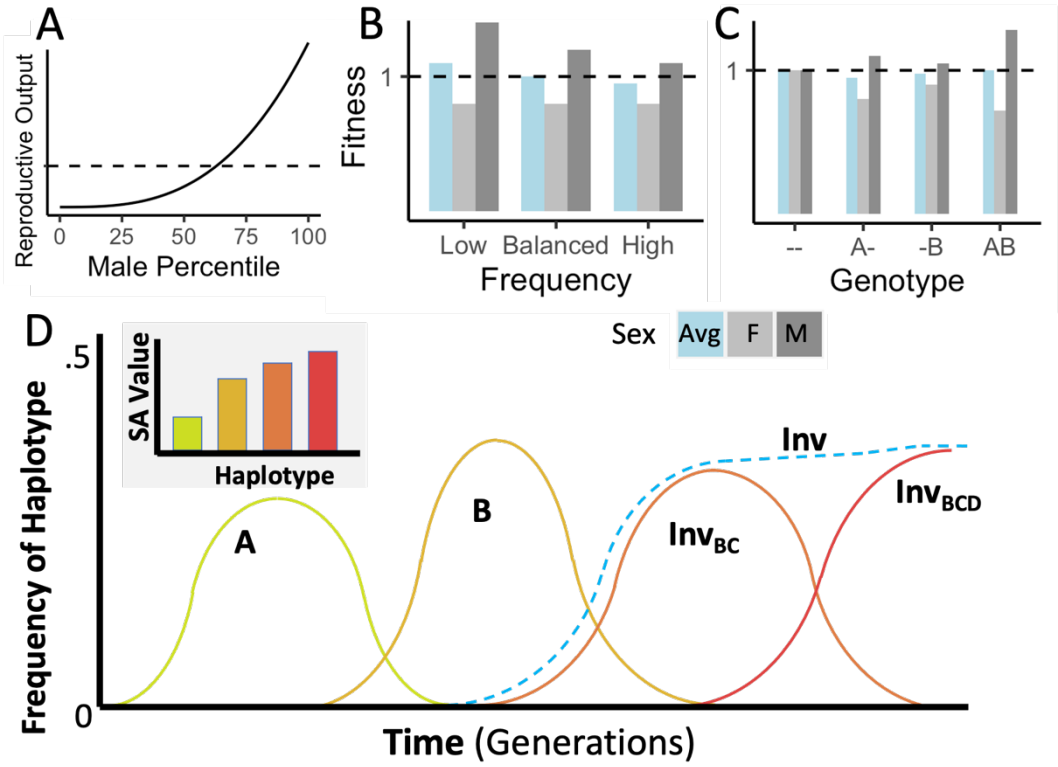
197 selection will generate linkage disequilibrium between sets of associated alleles that together strongly favor either
198 display or survival, and selection will favor the association of these haplotypes with alternate inversion karyotypes
199 to reduce the recombination load experienced in heterozygous parents carrying one haplotype each from the
200 extreme survival- and display-focused classes. We further predict that in a population that is close to the balanced
201 equilibrium state, the population of adults will contain a higher frequency of survival-focused haplotypes than the
202 population of zygotes, but the display-focused haplotypes will increase in frequency in zygotes relative to adults.

203 To build an intuition for the model, let us consider a simplistic model population with mate choice and phenotypic
204 variation in male display. Particularly if each female can choose amongst several or many males, the variance of
205 offspring number among males (Fig 1A) may be much higher than for females - yielding a relatively convex
206 distribution in which the least and most successful males produce no or many offspring respectively. A genetic
207 change that increases reproductive quality in a male can increase competitiveness and offspring count, but an
208 increase in the display of a male with low quality (change i in Figure 1A) results in a smaller change in number of
209 offspring than an increase in display of a male with high quality (change ii in Figure 1A), due to the reproductive
210 skew towards highly competitive individuals. Therefore, multiple variants increasing the reproductive rank of an
211 individual together have a larger fitness benefit than the combination of each substitution separately - a positive,
212 synergistic epistatic effect.

213 And yet, as pointed out above, this positive epistasis is accompanied by a negative frequency dependence due to
214 an antagonistic variant's decreasing advantage over other males as its frequency rises and more competing males
215 carry it. Display-enhancing variants with little or no pleiotropic cost may ultimately fix in spite of their diminishing
216 advantage as a sweep progresses, whereas those with relatively high pleiotropic costs (modeled here as reduced
217 survival probability) may be unconditionally deleterious and lost quickly. However, pleiotropic variants with more
218 comparable fitness effects on display enhancement and survival reduction may end up at a balanced equilibrium
219 frequency: when rare the allele experiences a large benefit from its display effects and is expected to increase in
220 frequency, but above a certain frequency the benefit becomes less than the antagonistic cost (Figure 1B) and the

allele is expected to decrease in frequency. In the absence of other forces, this dynamic is expected to maintain the allele under balancing selection indefinitely.

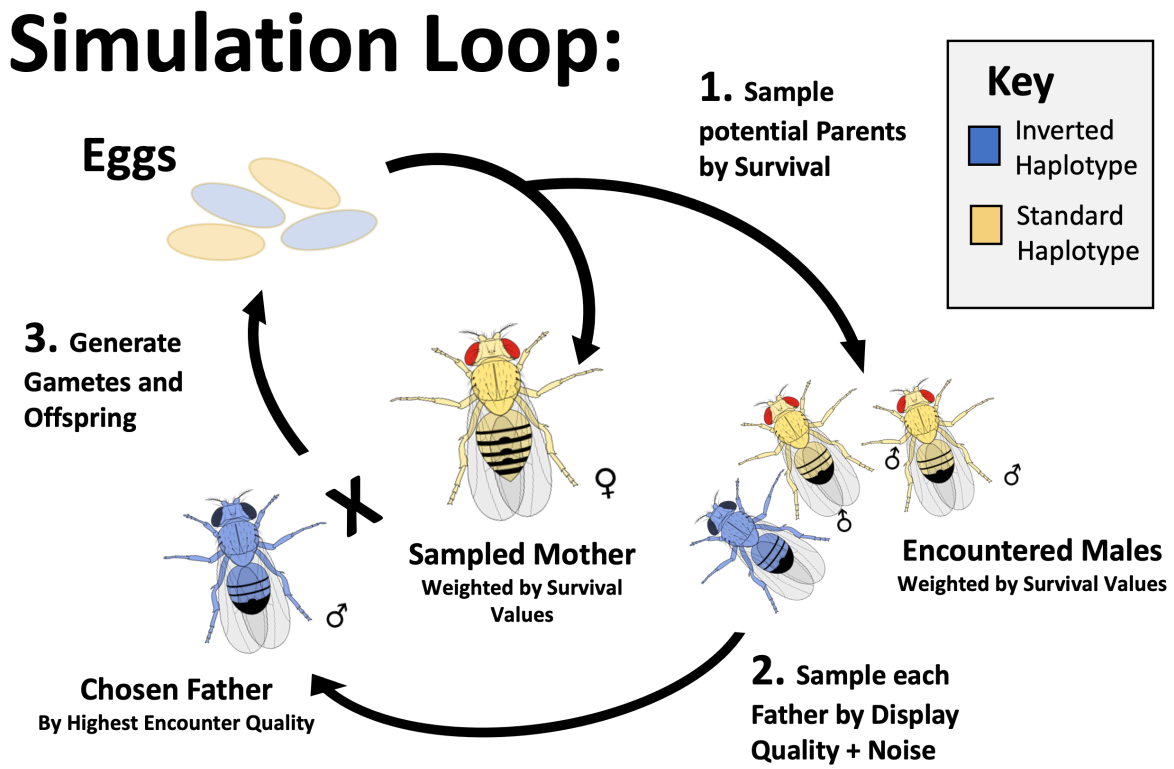
Consider a haploid population with two linked loci that carry such antagonistically pleiotropic alleles. Further consider the unrealistic circumstance that all alleles are at a frequency of 0.5 with perfect linkage equilibrium, and so all possible haplotypes are present at a frequency of 0.25. The haplotypes with all display-favoring alleles will then tend to be disproportionately chosen over other haplotype classes during male reproductive competition, yielding a positive epistatic interaction as seen in Figure 1A. If selection proceeds without recombination, the most display-heavy haplotype increases in frequency until reaching an equilibrium at which its relative advantage in mate choice is reduced enough to be balanced by the pleiotropic survival cost, which is assumed to be frequency-independent and non-epistatic among loci (Figure 1B). In this new state, the fitness of the most display-heavy haplotype and its opposite survival-focused haplotype are equally fit. However, haplotypes of intermediate composition all have lower fitness: they experience some pleiotropic cost, but due to the reproductive skew of mate choice they are selected disproportionately less often than the display-heavy haplotypes, so the net fitness contribution is negative (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 this cost of recombination, like an inversion, will be favored by indirect selection (Feldman et al. 1980; Otto and Lenormand 2002). As inversions reduce recombinant gametes between haplotypes heterozygous for the inversion, an inversion introduced to this population should benefit from associating with one or the other extreme haplotype and should then increase to the balanced frequency of that haplotype.



240

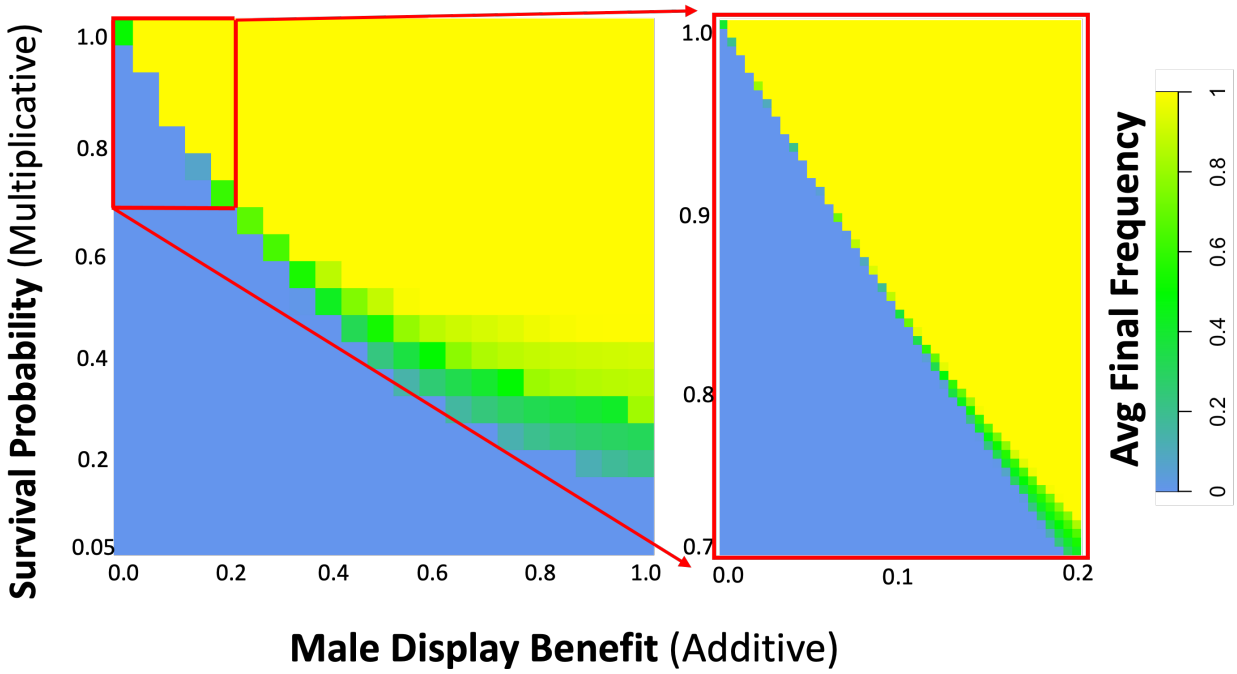
241 [Figure 1: Conceptual representations of the proposed model of inversion-associated balanced sexual antagonism.
242 (A) Hypothetical skewed distribution of male mating success, yielding a greater variance in reproductive success for
243 males than for females in our model. (B) Fitnesses of display-favoring alleles under sexually antagonistic balancing
244 selection, illustrating a pleiotropic variant that should rise in frequency when rare, but decline if frequency exceeds
245 the balanced equilibrium value. (C) Fitnesses of display-favoring alleles at two haploid loci under the same model,
246 illustrating synergistic epistasis between displaying-favoring alleles. (D) A hypothetical trajectory of four such
247 mutations, in which B outcompetes A, an inversion links B and C to create a more strongly display-favoring
248 haplotype, and then the addition of D to that haplotype furthers the accumulation of antagonistic variants,
249 reaching an equilibrium frequency while displacing less extreme display-favoring haplotypes.]

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 with crossover and gene conversion events, where odd numbers of crossovers in the interior of an inversion are thrown out and resampled from the same parent. In the simulations presented, males do not recombine when generating gametes, in light of the biology of *D. melanogaster*.]

263



264

265 [**Figure 3**: Simulations show that a balanced equilibrium frequency exists for certain sexually antagonistic variants.
266 The average final frequency of a single mutation with given effects on survival proportion (of both sexes) and male
267 reproductive quality score is shown, with green indicating intermediate balanced frequencies. These simulations
268 involve a single antagonistic variant with initial frequency 0.5, in a population of 1,000 individuals simulated for
269 20,000 generations, with further details as indicated in the Materials and Methods. The left panel depicts trait
270 values at 0.05 increments, whereas the insert presents a finer-scale set of simulations, every 0.005 trait units.]

271

272 We investigated the predictions of the above model of inversion-associated balanced sexual antagonism using a
273 purpose-built forward "Sexually Antagonistic Inversion simulator" (SAIsim; see Materials and Methods). This
274 simulation program incorporates mutation of new nucleotide and inversion polymorphisms, crossing over and
275 gene conversion (including between chromosomes heterozygous for inversions), drift, and sexually antagonistic

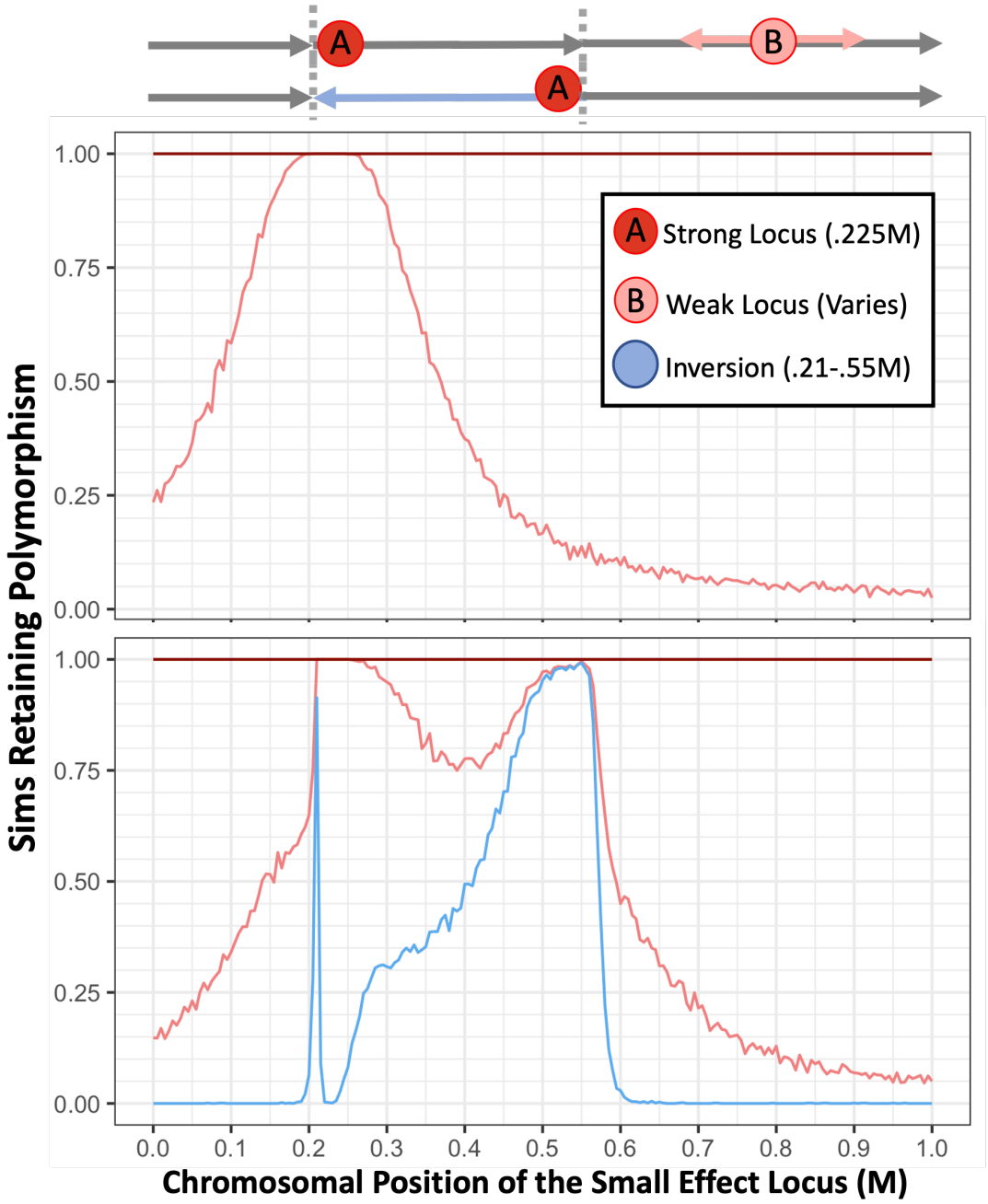
276 selection (Materials and Methods). SAIsim allowed us to model the influence of polymorphic inversions on
277 recombination, as well as the changes in linkage between sites entailed by the mutation and fixation of an
278 inversion. It explicitly models mate selection, in that each female randomly selected to reproduce is assumed to
279 encounter a number of randomly selected males, and the display values of each male are assigned based on the
280 sum of genetic and environmental effects (with genetic effects based on genotypes at antagonistic loci), and the
281 male with the highest display value then succeeds in this mating (Figure 2). We found it easiest to write a new
282 simulator as no extant simulator accurately managed mutation and recombination of multiple inversion
283 polymorphisms, as well as an explicit mate competition in generating offspring.

284 With this simulator, we first confirmed that individual variants with an antagonistic pleiotropy between survival
285 (initially of both sexes) and male reproduction can generate stable selectively balanced polymorphisms *in silico*.
286 We simulated populations of N = 1000 carrying a single variant of a specified survival and display quality at an
287 initial frequency of 0.5. After 20N generations, we calculated the average final frequency of the variant. From this
288 analysis we can see that there is a range of paired trait effects for which the mutation remained at intermediate
289 frequency after 20N generations (Figure 3). In a scenario where only females pay the survival cost, the balanced
290 parameter space is shifted somewhat, but qualitatively the results are similar (Figure S1). Alleles with small effect
291 sizes had only a narrow band of balanced selection, outside of which they were either fixed or lost, while large-
292 effect alleles remained balanced across a larger area of simulated parameter space. While this balanced range is
293 narrow, such variants may persist under balancing selection indefinitely in the absence of other forces.

294 To test the effects of association with an inversion, we initially ran two sets of simulations, considering two
295 antagonistic variants found to undergo balancing selection separately. First, we spaced the two antagonistic loci at
296 a varying genomic distances, in the absence of an inversion (Figure 4). In each case, the stronger antagonistic
297 variant (modelled at fixed position 0.225M) persists, and so we focus on whether the weaker variant (at varying
298 positions) also persists with it, or else is outcompeted by the more antagonistic variant. As seen in the upper panel
299 of Figure 4, the weaker variant (pale red line) is retained at close linkage, as the haplotype containing both
300 antagonistic variants acts as a single larger effect supergene. However, as the weaker variant moves farther from

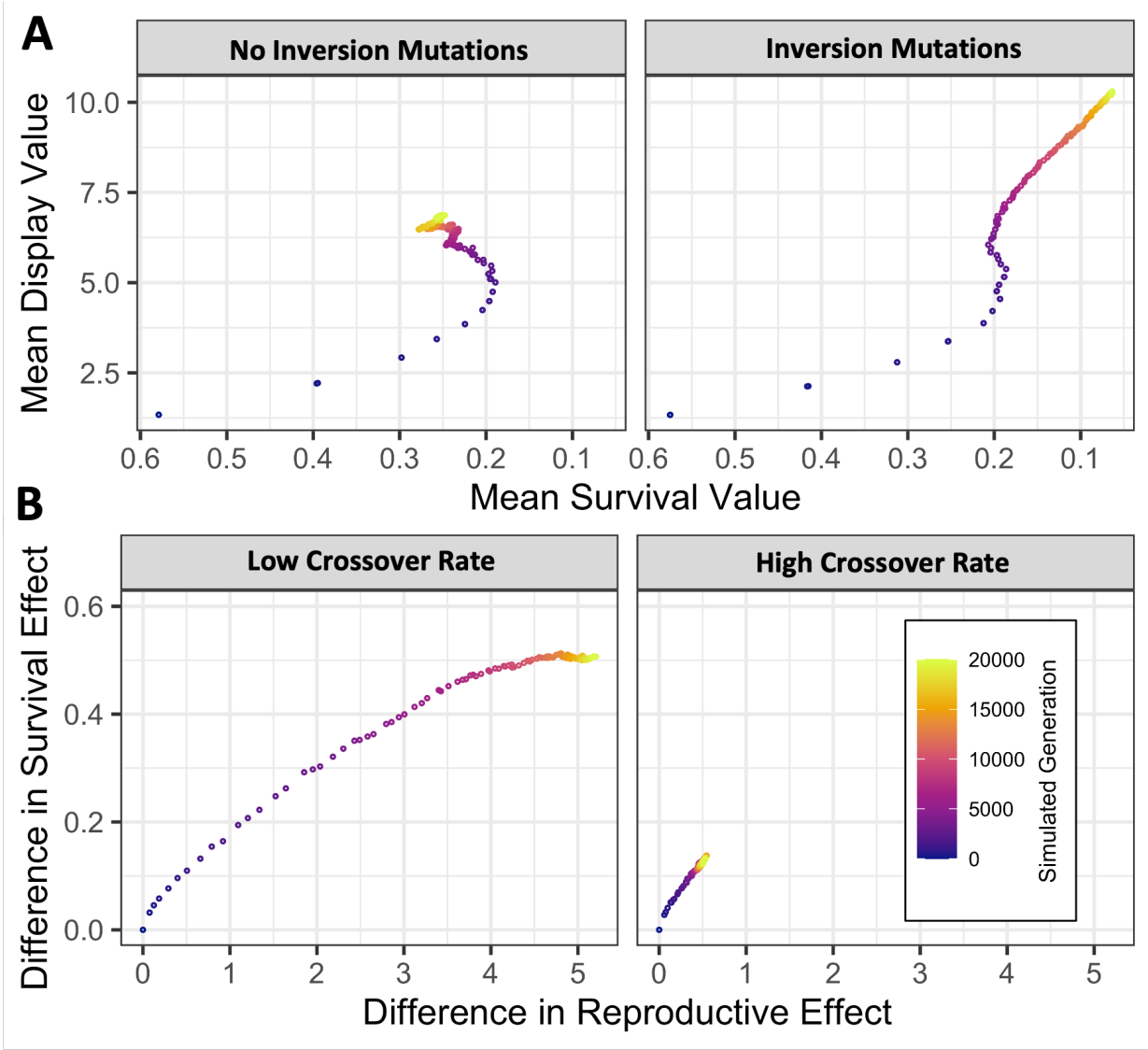
the stronger variant, recombination increasingly generates unfit intermediate genotypes compared to the extremes of the reproduction-survival tradeoff, and the stronger variant outcompetes the weaker variant among these haplotypes, leading to the loss of the weaker variant.

We then examined the same spacings between antagonistic variants across a region containing a polymorphic inversion, with its left breakpoint close to the stronger variant, and the right breakpoint within the range of the positions in which the weaker variant was simulated (Figure 4, lower panel). When the inversion is added, the weaker variant is now retained when the two antagonistic variants are relatively more distant, due to association of the weaker variant with the right-hand inversion breakpoint (pale red line), at which point this association also preserves the inversion itself in a polymorphic state (light blue line). While more stochastic, a weaker antagonistic variant near the middle of the inversion also appears to maintain polymorphism more frequently in systems starting with an inversion than without, i.e. across approximate positions 0.30-0.45M in Figure 4, in spite of its weaker linkage with the inversion itself due to double cross-over events. The inversion is not maintained when the weaker antagonistic variant is within the inversion but tightly linked to the stronger variant (see positions 0.22-0.25M); here there is little crossing-over for an inversion to suppress, and so the effect of the inversion approaches neutrality. Interestingly, there is a small window of weaker variant position in which the inversion maintains polymorphism even in the face of relatively tight linkage to the stronger variant (see positions 0.2-0.22M). Here, the weaker antagonistic variant is more tightly linked to the breakpoint than to the stronger variant in this range, and presumably benefits more consistently from the crossover reduction. The maintenance of the weaker variant drops more rapidly outside of the inversion than over similar distances in the simulation without inversions (positions 0-0.2M, bottom versus top panel). This seemingly counter-intuitive result may stem from crossover events that result in aneuploid gametes being resampled in our simulations, which effectively increases the crossover rate outside the inversion breakpoints, a behavior which should be biologically reasonable in many clades. Hence, for a weaker variant outside the left inversion breakpoint, the presence of the inversion can increase the crossover rate between the two antagonistic variants, and thus reduce the probability of of the weaker variant surviving by retaining linkage to the stronger one.



[Figure 4: Simulated inversions persist as polymorphisms when linked to sexually antagonistic, pleiotropic variants. In turn, the presence of inversions facilitates the accumulation of antagonistic variation. Results of simulations with two defined antagonistic variants alone (upper panel) and in the presence of a defined inversion (lower panel). The 1 Morgan chromosomal segment is diagrammed above the plots, with the inversion breakpoints marked. The proportion of simulations in which the second, variably-positioned weaker antagonistic variant

(survival and reproductive values 0.86, 0.12) maintains polymorphism is shown (light red line), dependent on the recombination distance from a stronger allele (0.75,0.3) at fixed position 0.225M, which always persists (dark red line). In the lower panel, the proportion of replicates in which the inversion is retained is also plotted (blue line). Populations start with an equal frequency of all haplotypes in a population of 1000 individuals. 1000 replicate simulations were run for 20*N* generations in each parameter combination.]



339 [**Figure 5**: Simulated randomly mutating inversions persist as polymorphisms and accumulate sexually antagonistic
340 variation when linked to sexually antagonistic, pleiotropic variants. (A) In simulated populations with sexually
341 antagonistic mutations, simulations which also allow inversion mutations accumulated more average display
342 benefit and survival cost across generations (right panel) than those without inversions (left panel). (B) When
343 inversion mutation is enabled, a low crossover rate (left panel) allows populations to accumulate larger differences
344 between the most common karyotype and all others than when the rate of crossover is particularly high (right
345 panel). 1000 replicate simulations were run for 20*N* generations in each parameter combination.]

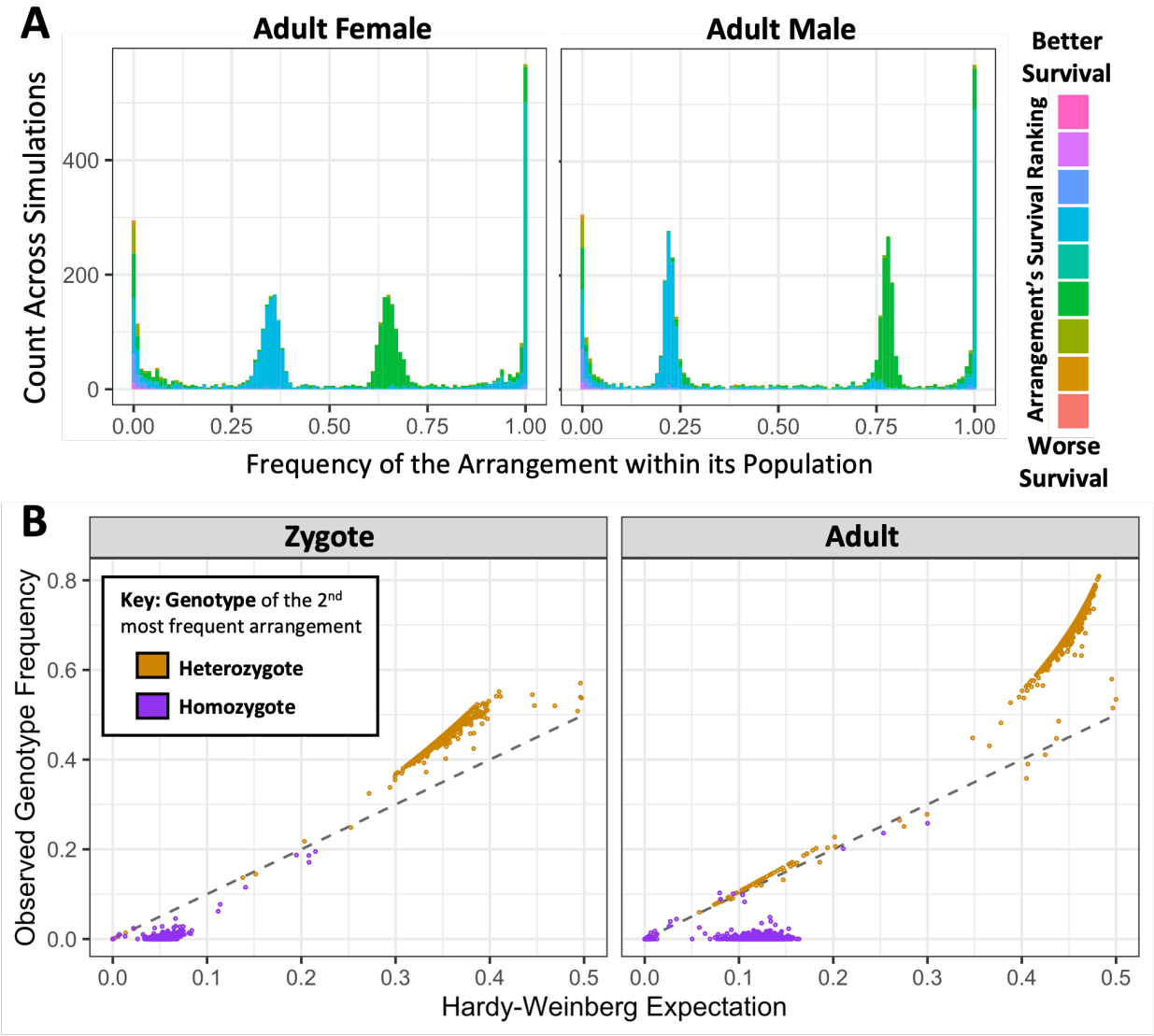
346

347 In further simulations, we began with a population initially free of variation, in which sexually antagonistic
348 mutations arise at a specified rate, randomly located across a 1M simulated region. These mutations had
349 pleiotropic effects increasing reproductive quality score and decreasing survival proportion, each drawn separately
350 from distributions (Materials and Methods). Depending on the effect scores drawn, a mutation could be broadly
351 advantageous (potentially leading to a selective sweep), broadly deleterious, or potentially subject to balancing
352 selection (Figure 3). Inversions (of random lengths and positions) also arose at a specified rate. The frequency
353 trajectories of antagonistic variants and inversions are tracked, and the simulation was run for 20*N* generations to
354 allow near-equilibrium patterns to be established. These simulations established the potential of a population to
355 mutate and evolve an inversion that links sexually antagonistic, pleiotropic alleles to maintain polymorphism in a
356 balanced manner (Figure 5B), although the degree of antagonism accumulated depends strongly on the rate of
357 crossing over (relative to the occurrence of antagonistic mutations; Figure 5C).

358 The simulations that successfully maintained balanced antagonistic chromosomes also had an interesting
359 distribution of allele and genotype frequencies of the balanced haplotypes at equilibrium. Without any selection,
360 the inversion frequency follows a pattern characteristic of the expected neutral site frequency spectrum (Figure
361 S2). However, in parameter spaces where simulations with shared survival costs between sexes maintained
362 polymorphism under balancing selection, inversion frequency distributions were tightly clustered around an
363 intermediate frequency (Figure S2, 6A). Strikingly, the frequencies of alternate haplotypes consistently clustered

364 around 0.25 and 0.75 among the zygote population. We found that this outcome was related to the genotype
365 frequencies of reproducing adults. These simulations evolve toward strong antagonism, and it is predominantly
366 males homozygous for the display-favoring haplotype (whether that is inverted or standard) who succeed in
367 reproducing, and these males therefore contribute one copy of the display-favoring haplotype to each of their
368 male and female offspring. However, this haplotype's significant cost to viability and longevity would select for the
369 heterozygous females over the display-favoring homozygotes. Given that successful parents would predominantly
370 be heterozygous females and display-homozygous males, three quarters of transmitted chromosomes have the
371 display-favoring haplotype and one quarter does not, yielding the 0.25 and 0.75 inversion frequencies seen in
372 simulations.

373 The sex-specific fitness dynamics present at equilibrium under this model also result in population-wide
374 genotype frequencies that depart from standard Hardy-Weinberg expectations. If the preponderance of matings
375 are between heterozygous females and display-favoring males, then homozygotes for the survival-focused
376 haplotype will rarely be produced (as confirmed in Figure 6B). We note that in cases where the standard
377 arrangement favors male display, then especially if antagonism is strong, this model could explain an observed
378 deficiency or absence of inversion homozygotes, even if the inversion was not associated with any recessive
379 deleterious variation. In contrast, heterozygotes may become disproportionately frequent after viability selection
380 (Figure 6B). And yet, this heterozygote abundance reflects neither the existence of recessive deleterious variants
381 on each haplotype, nor overdominance with respect to viability. Hence, our model may offer an alternative
382 explanation for deviations from Hardy-Weinberg genotype frequencies observed for polymorphic inversions in
383 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 by how the average survival effect of that arrangement ranks compared to the others in its population. Ranks are normalized to be relative to the median arrangement value, to better account for rare arrangements and ties. These simulations reach an equilibrium with two predominant arrangements, with the more survival-focused arrangement around an overall frequency of 0.25 and more common in females, and the more male-competition-focused arrangement around a frequency of 0.75 and more common in males. (B) Because successful males tend

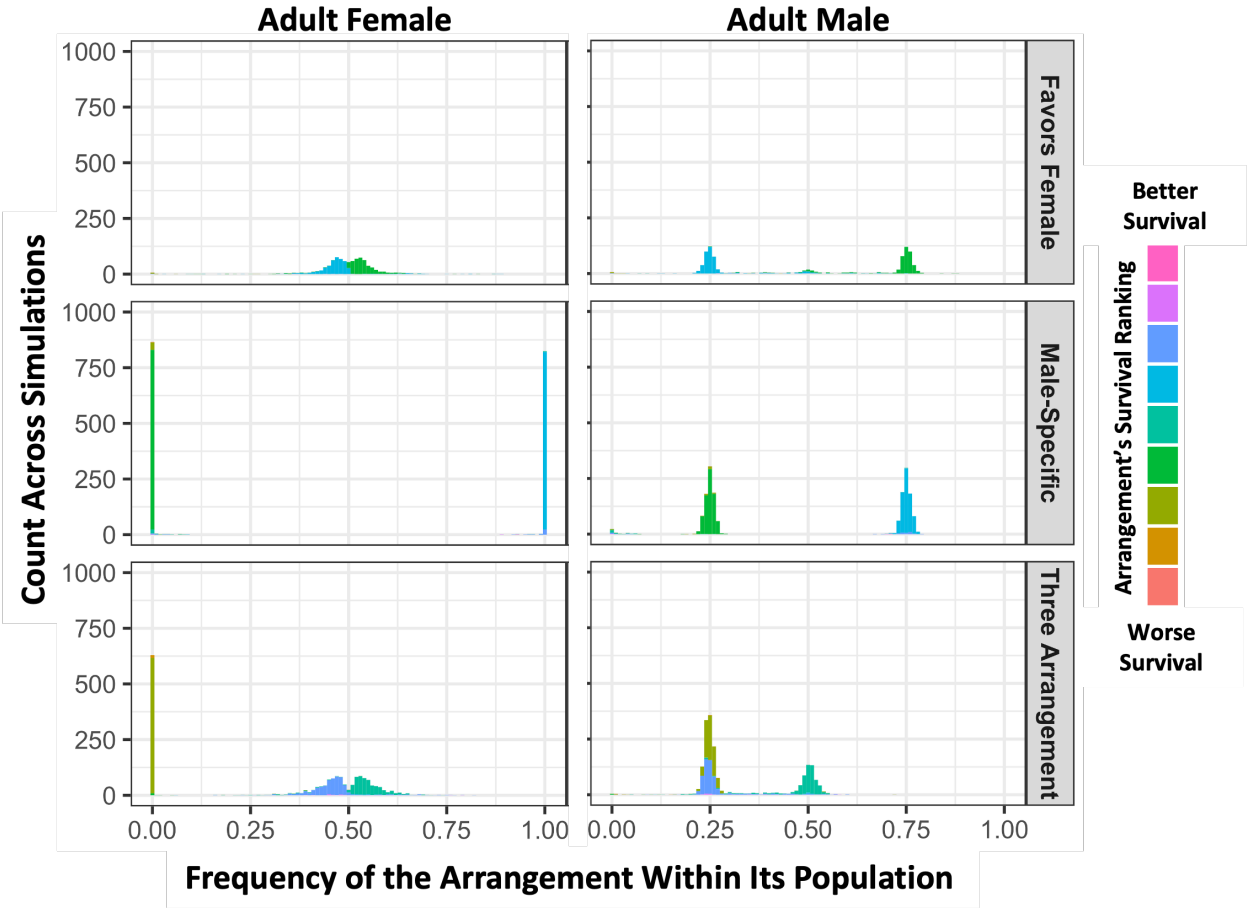
to be homozygotes and successful females heterozygotes, the less frequent of the two major arrangements (which tends to favor survival) has homozygote frequencies far below the Hardy-Weinberg expectation, and there is an excess of heterozygotes. These genotype frequencies shift somewhat between zygote (left panel) and adult (right panel) populations as the genotype affects survival.]

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

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

418 transitioning between each state once established. One clear barrier is that it is difficult to generate a less extreme
419 arrangement from a Y-like arrangement, as these can accumulate an indefinite number of mutations since males
420 experience no costs.

421



422

423 [**Figure 7:** In simulations with female-limited survival costs, three distinct outcomes are observed involving either
424 two or three balanced haplotypes. From simulations with randomly occurring antagonistic and inversion
425 mutations, histograms of inversion frequencies across simulations for surviving females (left) and males (right) are
426 partitioned into three categories based on sex-specific haplotype frequency outcomes. The "Favors Female"
427 category (top) includes simulations in which there are exactly two arrangements with frequencies between 0.1 and
428 0.9 in both sexes (20.0% of replicates). The "Male-Specific" category (middle) includes simulations in which there

429 are two arrangements with whole-population frequencies between 0.1 and 0.9, but one arrangement is absent in
430 adult females (37.0% of replicates). The "Three Arrangement" category (bottom) includes simulations in which
431 there are exactly three arrangements with whole-population frequencies between 0.1 and 0.9 frequency (40.6% of
432 replicates). 2.3% of simulations did not fit in these categories and had only one predominant arrangement. As in
433 Figure 6, inversions are colored by how the average survival effect of that arrangement ranks compared to the
434 others in its population, and ranks are normalized to be relative to the median arrangement value, to better
435 account for rare arrangements and ties.]

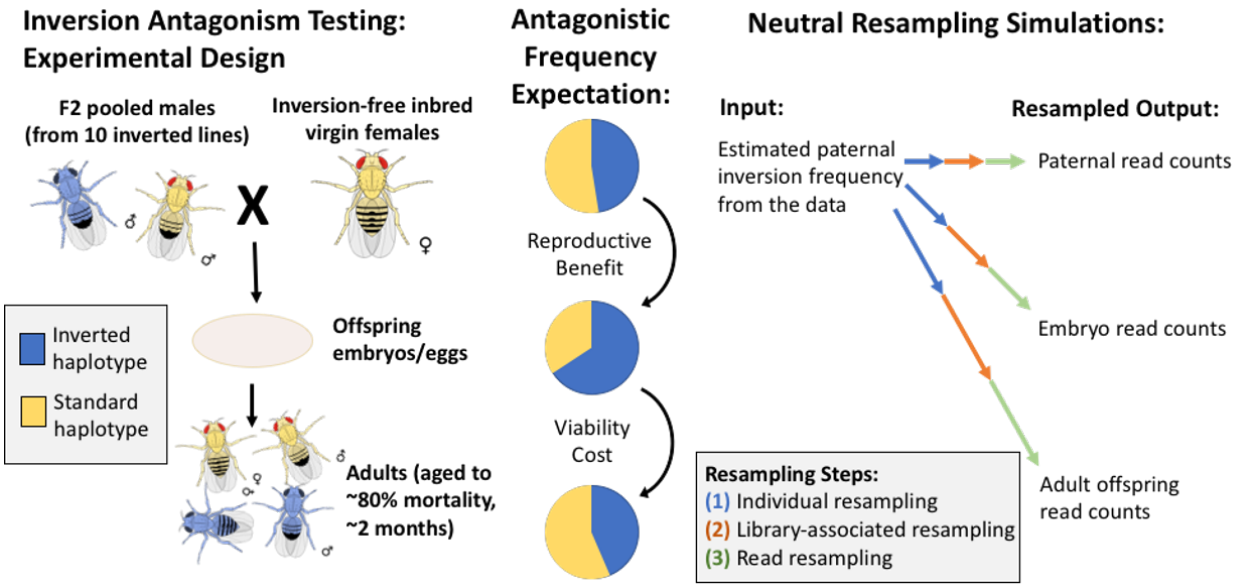
436

437 2.3 Empirical tests for inversion-associated tradeoffs

438 A specific motivation for developing and evaluating the model of inversion-associated, balanced sexual antagonism
439 was to potentially better explain the pattern of inversion frequencies observed in *D. melanogaster*, in which
440 multiple inversions reach some of their highest frequencies in populations from the ancestral range or with
441 similarly tropical/subtropical climates. We therefore set out to test experimentally whether any of four inversions
442 common in a Zambia population of this species show evidence for tradeoffs between male reproductive success
443 and survival. We assembled an outbred population with moderate frequencies of each inversion to obtain
444 karyotypically diverse males. We set up four crosses using these males along with females from one of four
445 different inversion-free inbred Zambia lines. This crossing scheme is represented in Figure 8. These parental
446 crosses occurred in a single large population cage, with populations of at least 600 of each sex. We developed a
447 PCR and next-generation sequencing-based assay to estimate inversion-associated SNP frequencies in each pool of
448 males and in each corresponding pool of embryos. These SNPs (given in Table S1) were based on our own analysis
449 of variants that showed perfect association with inversions among 197 Zambia genomes (Table S2; see Materials
450 and Methods). We then implemented a resampling method to test for significant inversion frequency changes
451 between parents and offspring (while accounting for the variance observed in technical replicates; see Materials
452 and Methods), in order to determine whether males carrying each inversion were more or less successful in

453 reproducing than the average male. We also compared frequencies between embryos and aged adults that
454 eclosed either in the first or second half of the collection period, to investigate the relationship between inversions
455 and viability/longevity. We calculated p-values from the resampling tests to assess the significance of each
456 individual change in inversion frequency between relevant samples (Table S3). However, in accord with our
457 primary question regarding fitness tradeoffs, we primarily focus here on joint tests of two opposing frequency
458 changes, as further described below. We could therefore test the hypothesis that a given inversion had opposing
459 influences on survival and male reproduction, which would provide evidence consistent with our model of
460 balanced sexual antagonism, while also examining whether any such tradeoffs depend on maternal genetic
461 background.

462



463

464 [**Figure 8:** The layout and potential expectations of the laboratory evolution experiment. We cross outbred males
465 from a high-inversion population to inbred, non-inverted females from a specific inbred line, and collect DNA
466 samples from the fathers, embryo offspring, and aged adult offspring (left). We hypothesize from our antagonistic
467 pleiotropy model that the inversions will experience opposing selection between fathers and embryos versus

468 between embryos and aged adults (middle). To test the significance of observed frequency changes, we designed a
469 resampling model to represent the sources of neutral variation expected in the experiment (right).]

470

471 When we initially examined the joint effects of male reproduction and survival, we found that in a majority of
472 cases, a given inversion increased in frequency in adult offspring compared to the null expectation based on the
473 parental male pool (in 13 out of 16 inversion-line trials; binomial $P = 0.02$; Figure 9). Given that inversion
474 homozygotes are not generated by our crossing design, these results are likely to indicate that inversion
475 heterozygotes often had higher fitness than standard homozygotes (across male reproduction and survival jointly).
476 We note that our standard homozygotes are not identical-by-descent but are instead heterozygous for unique
477 standard-arrangement chromosomes from two independent Zambia strains, and that long genomic tracts of
478 identity-by-descent between Zambia strains are quite rare in this large, ancestral-range population (Lack et al.
479 2016a). Hence, we do not expect that inversion heterozygosity is buffering against deleterious effects of
480 inbreeding. Instead, cases where inversions are associated with fitness advantages suggest either a simple
481 directional advantage of inversion-associated alleles for a given fitness trait, or else an advantage for heterozygous
482 inversion/standard genotypes in particular. In the former case, the inversions may still be under balancing
483 selection in the wild, if the equilibrium frequency of the inversion was shifted due to lab conditions such that one
484 karyotype was always favored in the lab environment tested. This scenario would mirror previous findings in which
485 inversions found at intermediate frequencies in natural *D. melanogaster* populations (Inoue 1979), as well as
486 inversions from seaweed flies known to be under balancing selection in the wild (Mérot et al. 2020), were both
487 inferred to be under purely directional selection in the lab. We might still expect some difference in benefit across
488 the life history of a single generation if there is a balanced tradeoff in some alternate environment or at a different
489 frequency, consistent with the differences observed in our populations.

490 When we separated the effects of male reproduction (based on father to embryo changes) and survival (embryos
491 versus adults), we found that the effects of inversions on survival were generally consistent across maternal
492 genotypes. Three inversions were associated with increased survival: *In(3R)K* was in all four crosses, while *In(2L)t*

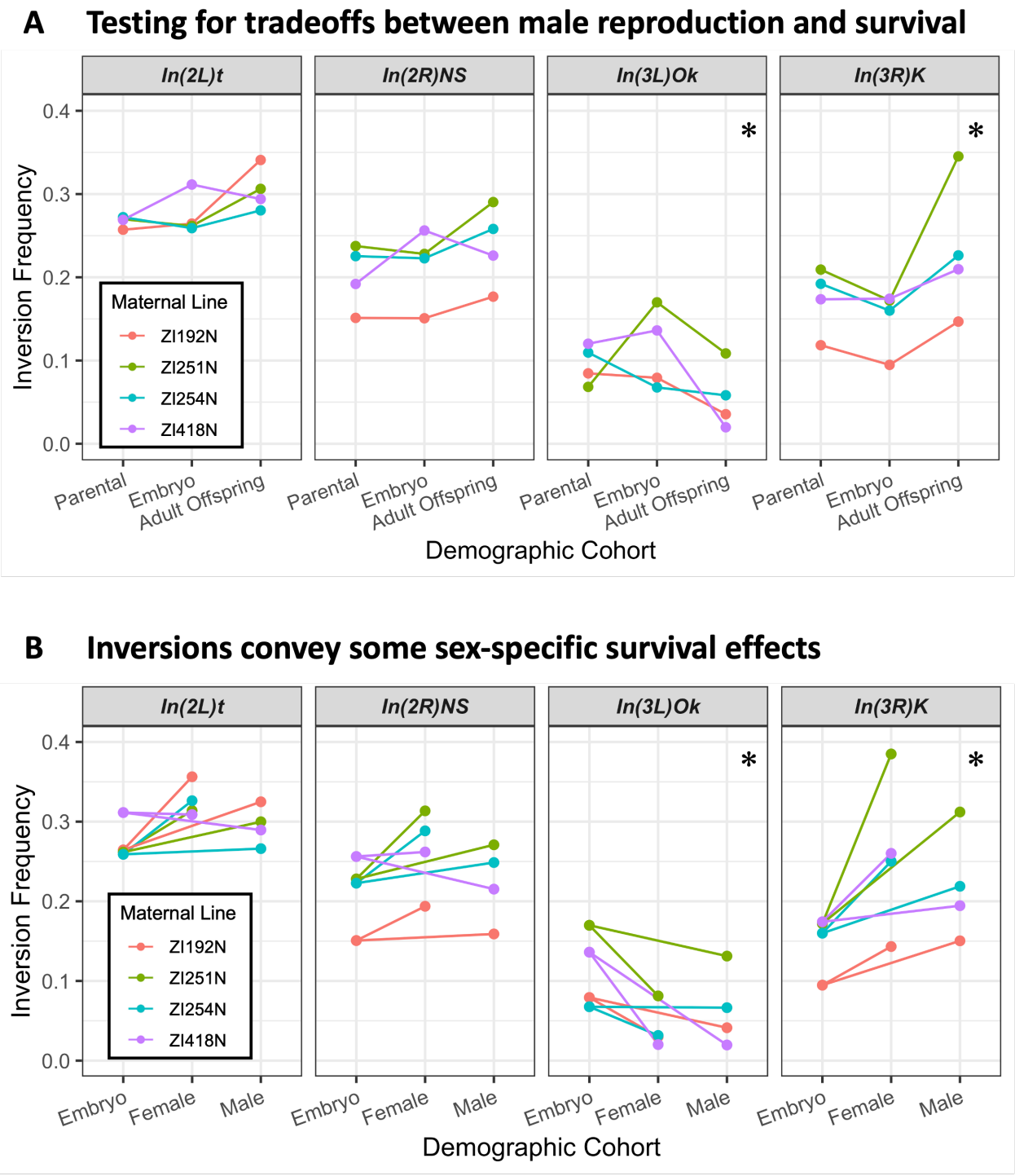
493 and *In(2R)NS* each showed elevated survival with all maternal strains except ZI418N (Figure 9A). In contrast,
494 *In(3L)Ok* was associated with lower survival in each cross (Figure 9A). With regard to male reproduction, the
495 trajectory of inversion frequencies from fathers to embryos to offspring was less consistent between different
496 crosses (Figure 9A), implying that differences in male reproductive success were dependent on the maternal
497 genotype, which is consistent with past findings in *D. melanogaster* (Reinhart et al. 2015). Despite such variability,
498 each of the four inversions had multiple frequency changes that departed strongly from null expectations (Figure
499 9A; Table S3), as might be expected if selection was present, but varied in strength and direction between crosses
500 and life stages.

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

512 We also tested for frequency differences between the different sexes of adult offspring and between the
513 early-eclosing and late-eclosing cohorts of adult offspring. All four inversions displayed fairly consistent differences
514 in the production of female versus male offspring (Figure 9B), suggesting that each of them have differential
515 effects on sex-specific survival. Strikingly, *In(3L)Ok* and *In(3R)K* each displayed sex survival differences in alignment
516 with their antagonistic effects implied above: *In(3L)Ok* favored male survival more-so than female survival ($p =$
517 0.0450), whereas *In(3R)K* favored the survival of females over males ($p = 4.61 \times 10^{-3}$; Table S4; see Materials and

518 Methods). The other two inversions, *In(2L)t* and *In(2R)NS*, each consistently favored female over male survival in
519 all four maternal crosses, although their initial unidirectional cross-combined *p* values (0.0260 and 0.0126) were
520 not significant after multiple test correction (p = 0.282 and 0.137). In contrast, we found no consistent effects of
521 inversions on eclosion time (Figure S3). Overall, we conclude that *D. melanogaster* inversions have distinct and
522 often context-dependent influences on male reproduction and on survival, with some evidence for antagonistic
523 tradeoffs of the sort proposed in our model.

524



[**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

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

537

538 3 Discussion

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

546 3.1 Inversions in *D. melanogaster*

547 We have presented results from a single generation laboratory evolution experiment, in which we crossed outbred
548 male *D. melanogaster* carrying inversions *In(2L)t*, *In(2R)NS*, *In(3L)Ok*, and *In(3R)K* to four separate inbred female
549 lines and tracked the frequency of the inversions in the initial males, embryonic offspring, and aged adults. For
550 most inversions, the strength and direction of the tradeoff varied notably between crosses to different inbred

551 maternal lines, mostly due to maternally-dependent variation in male reproductive-success. However, greater
552 consistency was observed for *In(3R)K*, which was associated with lesser male reproductive success but greater
553 survival. A significant tradeoff was also inferred for *In(3L)Ok*, which favored male reproduction (albeit variably)
554 over survival. These results provide significant but measured evidence for the presence of sexually antagonistic
555 autosomal inversions in this model species, while emphasizing the importance of genetic background.

556 As indicated in the Introduction, models like ours in which inversion frequencies are influenced by balancing
557 selection on pleiotropic tradeoffs may help explain multiple aspects of observed *D. melanogaster* inversion
558 frequencies. This class of models inherently accounts for the intermediate maximal frequency of most *D.*
559 *melanogaster* inversions in natural populations. To the extent that inversions favor male reproduction, this could
560 also help account for the predominantly autosomal locations of common inversions in this species, given that
561 male-specific fitness effects have a weaker influence on the frequencies of X-linked variants since the X
562 chromosome spends only one third of its time in males.

563 This model could also help explain the observed geographic variability of inversion frequencies, as a change in
564 environment may change the fitness value of either pleiotropic effect of the alleles. If a population experiences a
565 novel harsh environment with low survivability, then the population density may decrease, and the number of
566 encountered competing males may decrease as well, reducing the relative fitness advantage of enhanced male
567 display. If there is a genetic tradeoff between mating display and viability, the equilibrium frequency of an
568 antagonistic inversion that favors the genetic extreme of display should be lower. Additionally, in hostile
569 environments, the survival cost of display-favoring variants could be increased, which could further shift their
570 equilibrium frequencies downward or render them not worth carrying at any frequency. These predictions could
571 contribute to the observed negative relationships between some inversion frequencies and latitude or altitude
572 (Kapun et al. 2016; Pool et al. 2017). As a specific example, it is notable that in a high altitude Ethiopian
573 environment with year-round cool weather, which hosts some of this species' most pronounced phenotypic
574 evolution (Bastide et al. 2014; Lack et al. 2016b), inversions are virtually absent - occurring at the lowest
575 frequencies of any analyzed population (Sprengelmeyer et al. 2020). Consistent with such reasoning, a

576 disadvantage of some inversions in challenging thermal environments (both cold and hot) was supported by a
577 previous experimental evolution analysis (Kapun et al. 2014). However, our experimental results include evidence
578 that some inversions, such as *In(3R)K*, may have the opposite tradeoff - that the inverted arrangement is
579 associated with survival success instead of reproductive success. Notably, we have only tested survival and
580 reproduction in a benign laboratory environment, and not under challenging environments such as cold
581 temperatures. Other studies in *D. melanogaster*, including those on focused seasonal evolution, have found
582 evidence for important tradeoffs between reproduction and robustness to environmental challenges (*e.g.*
583 Behrman et al. 2015). Hence, it would clearly be desirable to study the effects of inversions on a wider range of
584 potential tradeoffs. For example, Said et al. (2018) suggested that inversions *In(2L)t* and *In(3R)K* might be
585 maintained by immunity-related tradeoffs, in light of an over-representation of immune-related genes among
586 differentially expressed transcripts between inverted and standard karyotypes.

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

596 As indicated in the Introduction, balancing selection on inversions could also be achieved by 'associative
597 overdominance' (Zhao and Charlesworth 2016; Gilbert et al. 2020), where two haplotypes are balanced by
598 recessive cost to both homozygotes. Inversions may be particularly susceptible to carrying recessive deleterious
599 alleles if population size and/or inversion frequency are low, since under these circumstances inversion
600 homozygotes are rarely generated and hence the crossover rate among inverted chromosomes is low, plus there

601 are fewer inverted chromosomes to facilitate rare double recombination events with standard chromosomes.
602 However, this dynamic is expected to decay with increased age and population size of the inversion, as mutation,
603 gene conversion, or double crossover resolve the negative linkage between recessive deleterious alleles. Most
604 common inversions in *D. melanogaster* are both relatively old and have large effective population sizes. Inversions
605 including *In(2L)t*, *In(2R)NS*, and *In(3R)K* have breakpoints estimated to be several times older than the expansion of
606 the species out of southern-central Africa (Corbett-Detig and Hartl 2012; Sprengelmeyer et al. 2020). Those three
607 inversions, plus *In(3L)Ok*, for which inversion age has not been estimated, all maintain frequencies greater than
608 0.15 within very large populations (Sprengelmeyer et al. 2020), and so it is not clear that the persistence of
609 recessive deleterious or lethal alleles in the inverted regions is expected to be so much higher than for standard
610 arrangements that inversions should be maintained by associative overdominance. In agreement with these
611 predictions, many wild-derived inbred strains of *D. melanogaster* are homozygous for inverted haplotypes (Lack et
612 al. 2016a). Furthermore, as noted in the Results, even an observed lack of inversion homozygotes could reflect
613 strong sexual antagonism as opposed to recessive deleterious variation. Nevertheless, further simulation and
614 experimental testing regarding the potential influence of associative overdominance and other processes on
615 inversion frequencies in *D. melanogaster* and other species would be desirable.

616 3.2 Generality of Balanced Pleiotropic Inversions

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

623 For the specific model explored here, of antagonistic pleiotropy between survival and display in a mate choice
624 system, the prevalence of the dynamic across species and populations is unclear. It seems plausible that a similar

model of antagonistic pleiotropy and balancing selection under mate choice underlies the balanced inversion system in the European ruff *Philomachus pugnax*, where three inverted haplotypes on an autosome determine aggressive, cooperative, and female-mimic male lekking morphs, albeit with a dominance relationship and homozygous inviability for an inverted haplotype (Küpper et al. 2016). More generally, a number of factors contribute to the potential prevalence of this dynamic. Mate competition and choice during reproduction is rather common among animals, and this may mean sexually antagonistic inversions are plausibly also common. The model depends on the distribution of fitness effects of new mutations and the availability of sexually antagonistic variation, but such variation does seem prevalent (Zajitschek and Connallon 2018; Ruzicka et al. 2019). More likely to limit the prevalence of such inversions are the meiotic costs experienced by inversion heterozygotes in some systems (White 1973), whether sexually antagonistic variants are clustered in a region that would allow inversions to increase linkage, and whether other mechanisms like sex-biased expression (Ingleby et al. 2015), gene duplication (Connallon and Clark 2011), and linkage to the sex chromosome or sex-determining locus would resolve the antagonism more frequently (though see following subsection on sex chromosome evolution).

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

Essentially, any population with antagonistic pleiotropy involving some form of negative frequency-dependent balancing selection on one of the phenotypes might exhibit the same dynamic. The population would experience similar divergent selection towards alternate phenotypic extremes, defining an emergent level of epistasis between alleles that trade off between the pleiotropically determined traits. This idea integrates with existing

650 thought on the evolution of recombination and divergent selection on balanced systems. For example, Kopp and
651 Hermisson (2006) have proposed a specific form of frequency dependent divergent selection (FDDS), involving
652 competitive exclusion between similar genotypes for a trait otherwise under stabilizing selection, and this model
653 was expected to result in single or few loci of large effect. Their model did not consider pleiotropy or linkage
654 modifiers, but still had comparable dynamics in that selection in the FDDS regime favored small numbers of large
655 effect loci, while the inversion haplotypes in the model presented here simplify the recombination architecture to
656 approximate a single locus. This simplification of the genetic architecture can be further seen as related to the
657 reduction principle in recombination modifier theory (Feldman et al. 1980; Altenberg et al. 2017). When a
658 population experiences selection towards two (or more) divergent optima, under a selection regime that provides
659 a balancing mechanism to maintain both, indirect selection will favor a reduction of the genetic architectures to
660 prevent the generation of intermediate phenotypes. The likelihood of either model likely rests on the inherent
661 polygenicity of the trait under selection, with simpler traits being more likely to resolve by single locus FDDS
662 dynamics than by inversion association (*e.g.* Yassin et al. 2016).

663 3.3 Autosomal inversions and sex chromosome evolution

664 If balanced, sexually antagonistic, autosomal inversions occur frequently enough, they may also contribute to the
665 evolution of sex chromosomes. Given an established inversion with significant sexual antagonism, the inverted
666 haplotype that favors one sex still experiences an antagonistic cost of transmission to offspring of the opposite sex.
667 Mutations that link the haplotype with the sex it benefits are therefore favored under selection, and this can occur
668 either by fusing to an existing sex chromosome (Charlesworth and Charlesworth 1980), though this only fully
669 resolves the antagonism for linkage to Y and W chromosomes, or by linkage to a new sex determining mutation
670 (van Doorn and Kirkpatrick 2007). In particular, we discovered that strong inversion-linked antagonism could yield
671 a scenario where essentially all highly-successful males were homozygous for a display-favoring karyotype,
672 whereas reproducing females were overwhelmingly heterozygotes (Figure 5). This situation is similar to the ZW sex
673 chromosome system, suggesting that this particular pleiotropic trade-off may favor resolution by linkage to a
674 female sex determination factor. Further, simulations in which only females experienced antagonistic survival cost

usually generated populations with a chromosome essentially lethal in females, thus exclusively transmitted in males similar to a Y chromosome (Figure 7). These simulations sometimes had three states that appeared stable at 20*N* generations, including Y-like, W-like, and shared intermediate arrangements. This three-arrangement stable polymorphism seems unusual for a natural population, but might be interesting to explore in populations with polymorphic sex determination.

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

A basic prediction of the above hypothesis is that at least one inversion difference between sex chromosomes should already exist at the time of new sex chromosome formation, and genomic differentiation therefore begins between inversion karyotypes on an autosome even before sex-specific transmission of the region. This contrasts with current models of sex chromosome evolution in which inversion accumulation and genomic differentiation accrue only after sex chromosome formation, due to advantages in suppressing recombination involving sexual antagonistic variants already associated with a sex-determining locus (Wright et al. 2016). Additional genomic studies of new sex chromosomes are needed to assess the potential role of sexually antagonistic inversions in sex chromosome evolution.

Summary and Future Prospects

We have advanced a model in which inversion polymorphism is maintained by the frequency-dependent fitness effects of a pleiotropic tradeoff associated with inversion-linked genetic variation. We have used sexual antagonism associated with increased male reproduction versus survival as a specific motivating case, and we explore the predictions of this model using a novel simulation algorithm. We confirmed model predictions that

some antagonistic variants can be maintained at an equilibrium frequency, and that inversions can allow multiple antagonistic variants to persist and form more strongly antagonistic haplotypes. We complemented our conceptual and computational modeling with laboratory experiments designed to detect fitness effects of four inversions common in an ancestral range population of *D. melanogaster*. We indeed found some evidence of tradeoffs between male reproduction and survival for these inversions, with two out of four inversions showing significant evidence for such a tradeoff, although we also observed a strong dependence of male reproductive success on female genetic background.

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

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

4 Materials and Methods

4.1 Simulations

We developed a new individual-level forward simulator ('SAIsim' for Sexually Antagonistic Inversion simulator) written in python 3.7 for the simulation of inversions and sexually antagonistic alleles at infinite loci, with uniform rates of mutation, crossing over, and gene conversion (per bp, per generation) over a number of independent chromosome arms. The dynamic we were interested in modeling requires direct accounting of pleiotropy between different components of fitness, crossing over and gene conversion within inversions (as well as the possibility of inversion fixation), and directly modeling male display and female choice, and no existing simulator provided this functionality. These simulations were used to establish the potential for the model to occur in a population generally, and not to compare against specific empirical observations. We have used literature estimates of mutation rate and conversion rate of Zambian *D. melanogaster* for these simulations (Comeron et al. 2012; 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.

In SAIsim, a population is instantiated as a python object, and populated with individuals (also python objects) based on provided (or initially empty) genomes. Each generation, as diagrammed in Figure 5, the population calculates a survival probability for each individual as the product of the survival values of each variant they carry, which range [0,1]. Male and female survival costs can each be included in a sex-specific manner, but only globally in the simulation, not per mutation. For each offspring in the next generation, a mother is selected randomly, weighted to survival values, with replacement. Then a set number of potential fathers (*i.e.* the uniformly specified encounter number) is randomly sampled for each offspring from the full population, weighted by their survival values. The males each have a display quality given as the sum of the reproductive values of their alleles, plus a normally distributed additional noise with standard deviation equal to one display quality unit. SAIsim allows the specification of mutation rates and the distribution of effect sizes of new mutations; our simulations use mutations

with uniformly distributed quality values in the range [0,1]. A noise parameter, normally distributed with a standard deviation of 1, is also applied to each male display value, with the noise effect generated separately for each interaction. Once the male display qualities are calculated based on genotype and noise, the mother of each offspring chooses a father deterministically by the highest quality. As offspring are generated, crossover locations are sampled, and resampled in cases where the resultant gamete would be aneuploid. Gene conversion is modeled as a flat probability per locus, with each event independently affecting just one locus, rather than modeling a conversion tract length. Inversions and new mutations are added to the resultant gamete, with inversions resampled when overlapping an inversion present on the chromosome. The simulation is intended to represent *D. melanogaster* life history and allows removal of recombination in males, though this also removes gene conversion by default.

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

4.2 Experimental Populations and Sequencing

To act as a high-inversion source population for the potential fathers in our experiment, we combined 15 males and 15 females from each of 10 inbred strains derived from wild flies collected in Zambia. These 10 strains were chosen to collectively carry the following common inversions: *In(2L)t*, *In(2R)NS*, *In(3R)K*, *In(3R)P*. 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

770 be inherited through the mothers, inversion frequencies of successfully reproducing F2 males could be inferred
771 from their pooled offspring. Unless otherwise noted, all flies were kept in a lab space of 23°C with around a degree
772 of temperature fluctuation and without a controlled day/night cycle. All fly food used was prepared in batches of
773 4.5 L water, 500 mL cornmeal, 500 mL molasses, 200 mL yeast, 54 g agar, 20 mL propionic acid, and 45 mL tegosept
774 10% (in 95% ethanol). The four parental populations were each allowed to mate freely in clear plexiglass cages of
775 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
776 shut with a binder clip. The parents were given either fly food bottles or grape agar in 7.9 cm petri dishes as food
777 or laying substrate on alternating days over two weeks, after which all surviving male parents were counted and
778 frozen. The grape agar was prepared in batches of 500 mL from FlyStuff.com grape agar powder, and wet Red Star
779 yeast was brushed over the agar plates before use to encourage laying. Embryos were collected from the agar,
780 counted, and frozen for sequencing. Embryos laid in the bottles were allowed to hatch and develop into adults.
781 These adult offspring were separated by sex and by the time after laying at which they eclosed. In our husbandry
782 conditions, the earliest eclosion was 11 days after laying, but this was rare and a first eclosion at 12 days was more
783 common among the bottles. Most flies took several days longer to develop, and we considered flies collected in
784 the first 6 days of a 10 day collection window to be the early-eclosing pool, in order to split the pool sample sizes
785 closer to evenly. The collected adult flies were then aged by sequentially transferring them between bottles
786 weekly for approximately 2.5 months, and then the adult offspring were sexed, counted, and collected. A 10-week
787 aging window was chosen based on the time taken for a control set of four bottles of flies from the paternal
788 outbred population to reach 80% mortality. All adult and embryo samples were frozen at -80°C before DNA
789 extraction. This sampling scheme was chosen to identify inversion frequencies in the set of potential parents, in
790 embryonic offspring, and in older adults, and so to assess potentially selected changes in inversion frequency
791 across male reproduction or in egg to adult viability plus longevity.

792 DNA for each replicate age group was extracted by a phenol-chloroform extraction protocol used previously in
793 preparing DNA for the Drosophila Genome Nexus and *Drosophila* Population Genomics Project. The protocol is
794 similar to protocol 'BPC' used in DPGP phase 1 (Langley et al. 2012), except with the addition of proteinase-K
795 during homogenization of adult flies, and generally based on the common quick preparation protocol presented in

796 *Drosophila Protocols* (Huang et al. 2009). Inversion genotyping was performed by amplicon sequencing, where
797 amplicons associated with each inversion were designed to target at least one single nucleotide polymorphism in
798 the inversion interior that is fixed for alternate alleles between the inversion karyotypes, while still sharing
799 conserved primer regions across karyotypes to prevent primer bias. Genetic variation in these inversions was
800 assessed using previously sequenced and inversion-called haploid genomes from the Zambian *D. melanogaster*
801 population, available from the *Drosophila* Genome Nexus (Lack et al. 2016a). The amplicon region and primer
802 selection was performed with a custom script. The script calls Primer3 version 2.4.0 for primer candidate
803 generation (Koressaar and Remm 2007; Untergasser et al. 2012; Koressaar et al. 2018). Primers thus selected were
804 ordered from IDT as oligos ligated to Illumina stubby adaptors [Table S1]. Illumina sequencing libraries were then
805 prepared for each sample following the Illumina metagenomic amplicon sequencing protocol (Illumina 2013), with
806 50 ng/uL diluted DNA samples. and sequenced 2x150 on an Illumina NovaSeq to at least a depth of 20000 reads.
807 Depth was selected to ensure high likelihood of detecting a 0.025 change in inversion frequency from an initial
808 frequency of 0.10, based on simulated resampling of the collected flies, DNA extraction, and sequencing.

809 4.3 Sequence Analysis and Statistics

810 Sequence preparation and analysis were performed using custom pipeline scripts. Read trimming, alignment, and
811 QC were performed using bwa version 0.7.17-r1188 (Li 2013), GATK version 3.4-46 (Van der Auwera and O'Connor
812 2020), and Picard version 1.79 (Broad Institute 2019). Reads were trimmed at the ends to remove segments with
813 base quality less than 20, and read pairs were filtered to retain only those pairs where both reads aligned with
814 mapping quality >20. SNP identity at the sites of interest was called from those remaining reads with base
815 quality >20 at the SNP of interest. Due to the presence of chimeric reads in the amplicon library sequencing, we
816 chose to use a single discriminant SNP to determine each inversion's frequency in the read pool, instead of
817 incorporating haplotype or additional SNP information.

818 To assess the significance of the difference in inversion frequency between two sampled life stages in the same
819 experiment, we first calculated a chi-square test statistic from the two-by-two table of inverted and non-inverted

allele counts at two life stages. Allele counts were taken by multiplying the inversion-specific SNP frequency in the sequencing reads by the allele sample size. We then generated an approximate distribution of the chi-square statistic when selection is absent (our null hypothesis), using a resampling approach (Figure 6). Comparing the observed statistic to our modeled neutral distribution of test-statistic values allowed us to identify significance thresholds.

To generate the neutral distribution of chi-square test statistics, we took the observed paternal inversion frequencies and then sampled new paternal and offspring read sets to represent a single neutral sample. Ideally, we could sequence from the pool of all fathers in the experiment, but up to 15% of the paternal population died before collection. So, in generating a new paternal read set, we first modeled down-sampling of the paternal allele sample from the initial census size to the collected allele size as a hypergeometric sampling. Then we used a binomial sampling to account for the variation introduced by the extraction and library preparation. The extraction and library preparation appeared to introduce variation independent of allele sample size, possibly due to variation introduced during PCR or sample homogenization. This inference was based on analysis of data from control library preparations (Table S5) and sequencing done in parallel with our experiment from sample fly pools with known inversion frequencies. These control samples included an average of 6 technical replicates (separate library prep and sequencing replicates from the same flies) from four different expected frequencies for *In(2L)t,* as well as one extraction each for *In(2R)NS* and *In(3R)K*. Due to genotyping and husbandry difficulties we did not add a control set for *In(3L)Ok*. Based on these technical replicates, we estimated that taking a binomial sampling of size 316 from the sample allele frequency reflected a maximum likelihood model for the sample-size-independent variation introduced by the library prep and sequencing process. From the preliminary resampled paternal inversion frequency that now accounted for the library variation, we then took a binomial sample of size equal to the read count to account for sequencing variation, thus yielding resampled paternal read counts that account for individual sample size, library preparation, and depth of coverage. For the new offspring read sets, we instead used a binomial sampling of half the offspring alleles from the fully resampled paternal allele frequency for each separate offspring cohort, since the homozygous mothers all contributed non-inverted alleles. Finally, each offspring allele sample was resampled using the same size 316 binomial sampling to account for the variation

introduced during this library's preparation, and binomial sampled again with the read count size to account for sampling reads from the library, thus yielding resampled offspring read counts that accounted for the same three factors as for paternal counts.

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

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

To assess the significance of inversion frequency changes considered across all four maternal line crosses at once, we used Fisher's combined probability test. We applied this to both particular three-cohort frequency change

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

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

Data Availability

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

Acknowledgements

Particular thanks to Shimeng Gao, who assisted with *Drosophila* husbandry and collection of adult flies during the within-generation experiment. We also thank members of the Pool lab 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.

Works Cited

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

Audet T, Krol J, Pelletier K, Stewart AD, Dworkin I. 2023. Sexually discordant selection is associated with trait specific morphological changes and a complex genomic response. :2023.08.31.555745. Available from: https://www.biorxiv.org/content/10.1101/2023.08.31.555745v1

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

963 Hoquet T, Bridges WC, Gowaty PA. 2020. Bateman's Data: Inconsistent with "Bateman's Principles." *Ecol. Evol.*
964 10:10325–10342.

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

967 Huang W, Lyman RF, Lyman RA, Carbone MA, Harbison ST, Magwire MM, Mackay TF. 2016. Spontaneous
968 mutations and the origin and maintenance of quantitative genetic variation.Przeworski M, editor. *eLife*
969 5:e14625.

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

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

975 Innocenti P, Morrow EH. 2010. The Sexually Antagonistic Genes of Drosophila melanogaster. *PLOS Biol.*
976 8:e1000335.

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

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

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

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

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

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

988 Kapun M, van Schalkwyk H, McAllister B, Flatt T, Schlötterer C. 2014. Inference of chromosomal inversion dynamics
989 from Pool-Seq data in natural and laboratory populations of Drosophila melanogaster. *Mol. Ecol.*
990 23:1813–1827.

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

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

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

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

998 Krimbas CB, Powell JR. 1992. Drosophila Inversion Polymorphism. CRC Press

999 Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR. 2014. doublesex
1000 is a mimicry supergene. *Nature* 507:229–232.

1001 Küpper C, Stocks M, Risse JE, Remedios N dos, Farrell LL, McRae SB, Morgan TC, Karlionova N, Pinchuk P, Verkuil YI,
1002 et al. 2016. A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.*
1003 48:79–83.

1004 Lack JB, Lange JD, Tang AD, Corbett-Detig RB, Pool JE. 2016. A Thousand Fly Genomes: An Expanded Drosophila
1005 Genome Nexus. *Mol. Biol. Evol.* 33:3308–3313.

1006 Lack JB, Yassin A, Sprengelmeyer QD, Johanning EJ, David JR, Pool JE. 2016. Life history evolution and cellular
1007 mechanisms associated with increased size in high-altitude Drosophila. *Ecol. Evol.* 6:5893–5906.

1008 Lange JD, Bastide H, Lack JB, Pool JE. 2022. A Population Genomic Assessment of Three Decades of Evolution in a
1009 Natural Drosophila Population. *Mol. Biol. Evol.* 39:msab368.

1010 Langley CH, Stevens K, Cardeno C, Lee YCG, Schrider DR, Pool JE, Langley SA, Suarez C, Corbett-Detig RB,
1011 Kolaczkowski B, et al. 2012. Genomic Variation in Natural Populations of Drosophila melanogaster.
1012 *Genetics* 192:533–598.

1013 Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Available from:
1014 http://arxiv.org/abs/1303.3997

1015 Libbrecht R, Kronauer DJC. 2014. Convergent Evolution: The Genetics of Queen Number in Ants. *Curr. Biol.*
1016 24:R1083–R1085.

1017 Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, Horn D, Kayserili H, Opitz JM, Laxova R, et al.
1018 2015. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer
1019 Interactions. *Cell* 161:1012–1025.

1020 McBroome J, Liang D, Corbett-Detig RB. 2020. Fine-Scale Position Effects Shape the Distribution of Inversion
1021 Breakpoints in Drosophila melanogaster. *Genome Biol. Evol.* 12:1378–1391.

1022 Mérot C, Llaurens V, Normandeau E, Bernatchez L, Wellenreuther M. 2020. Balancing selection via life-history
1023 trade-offs maintains an inversion polymorphism in a seaweed fly. *Nat. Commun.* 11:670.

1024 Mettler E, Voelker RA, Mukai T. 1977. Inversion Clines in Populations of Drosophila melanogaster. *Genetics* 87:8.

1025 Morgan TH. 1910. Sex Limited Inheritance in Drosophila. *Science* 32:120–122.

1026 Mroczek RJ, Melo JR, Luce AC, Hiatt EN, Dawe RK. 2006. The Maize Ab10 Meiotic Drive System Maps to
1027 Supernumerary Sequences in a Large Complex Haplotype. *Genetics* 174:145–154.

1028 O' Donald P, Majerus MEN, Clarke BC, Partridge L, Robertson A, Clarke BC, Partridge L. 1997. Frequency-dependent
1029 sexual selection. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 319:571–586.

1030 Ohta T, Kimura M. 1970. Development of associative overdominance through linkage disequilibrium in finite
1031 populations. *Genet. Res.* 16:165–177.

1032 Otto SP, Lenormand T. 2002. Resolving the paradox of sex and recombination: Evolution of sex. *Nat. Rev. Genet.*
1033 3:252–261.

1034 Otwinowski J, McCandlish DM, Plotkin JB. 2018. Inferring the shape of global epistasis. *Proc. Natl. Acad. Sci.*
1035 115:E7550–E7558.

1036 Pennell MW, Kirkpatrick M, Otto SP, Vamosi JC, Peichel CL, Valenzuela N, Kitano J. 2015. Y Fuse? Sex Chromosome
1037 Fusions in Fishes and Reptiles. *PLOS Genet.* 11:e1005237.

1038 Pool JE, Braun DT, Lack JB. 2017. Parallel Evolution of Cold Tolerance within Drosophila melanogaster. *Mol. Biol.*
1039 *Evol.* 34:349–360.

1040 Poul YL, Whibley A, Chouteau M, Prunier F, Llaurens V, Joron M. 2014. Evolution of dominance mechanisms at a
1041 butterfly mimicry supergene. *Nat. Commun.* 5:5644.

1042 Reinhardt JA, Kolaczkowski B, Jones CD, Begun DJ, Kern AD. 2014. Parallel Geographic Variation in Drosophila
1043 melanogaster. *Genetics* 197:361–373.

1044 Reinhart M, Carney T, Clark AG, Fiumera AC. 2015. Characterizing Male–Female Interactions Using Natural Genetic
1045 Variation in Drosophila melanogaster. *J. Hered.* 106:67–79.

1046 Rest JS, Morales CM, Waldron JB, Opulente DA, Fisher J, Moon S, Bullaughey K, Carey LB, Dedousis D. 2013.
1047 Nonlinear Fitness Consequences of Variation in Expression Level of a Eukaryotic Gene. *Mol. Biol. Evol.*
1048 30:448–456.

1049 Rice WR. 1992. Sexually antagonistic genes: experimental evidence. *Science* 256:1436–1439.

1050 Rice WR. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nat.*
1051 *Lond.* 381:232–234.

1052 Rowe L, Chenoweth SF, Agrawal AF. 2018. The Genomics of Sexual Conflict. *Am. Nat.* 192:274–286.

1053 Ruzicka F, Hill MS, Pennell TM, Flis I, Ingleby FC, Mott R, Fowler K, Morrow EH, Reuter M. 2019. Genome-wide
1054 sexually antagonistic variants reveal long-standing constraints on sexual dimorphism in fruit flies. *PLOS*
1055 *Biol.* 17:e3000244.

1056 Said I, Byrne A, Serrano V, Cardeno C, Vollmers C, Corbett-Detig RB. 2018. Linked genetic variation and not genome
1057 structure causes widespread differential expression associated with chromosomal inversions. *Proc. Natl.*
1058 *Acad. Sci.* 115:5492–5497.

1059 Saunders PA, Neuenschwander S, Perrin N. 2019. Impact of deleterious mutations, sexually antagonistic selection,
1060 and mode of recombination suppression on transitions between male and female heterogamety. *Heredity*
1061 123:419–428.

1062 Schaeffer SW, Goetting-Minesky MP, Kovacevic M, Peoples JR, Graybill JL, Miller JM, Kim K, Nelson JG, Anderson
1063 WW. 2003. Evolutionary genomics of inversions in Drosophila pseudoobscura: Evidence for epistasis. *Proc.*
1064 *Natl. Acad. Sci.* 100:8319–8324.

1065 Sprengelmeyer QD, Mansourian S, Lange JD, Matute DR, Cooper BS, Jirle EV, Stensmyr MC, Pool JE. 2020.
1066 Recurrent Collection of Drosophila melanogaster from Wild African Environments and Genomic Insights
1067 into Species History. *Mol. Biol. Evol.* 37:627–638.

1068 Stearns FW. 2010. One Hundred Years of Pleiotropy: A Retrospective. *Genetics* 186:767–773.

1069 Sturtevant AH. 1921. A Case of Rearrangement of Genes in Drosophila. *Proc. Natl. Acad. Sci.* 7:235–237.

1070 Sturtevant AH, Beadle GW. 1936. The Relations of Inversions in the X Chromosome of Drosophila melanogaster to
1071 Crossing Over and Disjunction. *Genetics* 21:51.

1072 Surmacki A, Ożarowska-Nowicka A, Rosin ZM. 2013. Color polymorphism in a land snail Cepaea nemoralis
1073 (Pulmonata: Helicidae) as viewed by potential avian predators. *Naturwissenschaften* 100:533–540.

1074 Sved JA. 1968. The Stability of Linked Systems of Loci with a Small Population Size. *Genetics* 59:543–563.

1075 Tatarenkov A, Healey CIM, Grether GF, Avise JC. 2008. Pronounced reproductive skew in a natural population of
1076 green swordtails, Xiphophorus helleri. *Mol. Ecol.* 17:4522–4534.

1077 Thain D, Tannenbaum T, Livny M. 2005. Distributed computing in practice: the Condor experience. *Concurr.*
1078 *Comput. Pract. Exp.* 17:323–356.

1079 Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities
1080 and interfaces. *Nucleic Acids Res.* 40:e115.

1081 Van der Auwera GA, O'Connor BD. 2020. Genomics in the Cloud: Using Docker, GATK, and WDL in Terra. O'Reilly
1082 Media, Incorporated

1083 Wang J, Wurm Y, Nipitwattanaphon M, Riba-Grognuz O, Huang Y-C, Shoemaker D, Keller L. 2013. A Y-like social
1084 chromosome causes alternative colony organization in fire ants. *Nature* 493:664–668.

1085 Wellenreuther M, Bernatchez L. 2018. Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends Ecol. Evol.*
1086 33:427–440.

1087 White MJD. 1973. Animal Cytology and Evolution. 3rd ed. Cambridge University Press Available from:
1088 https://wellcomecollection.org/works/meqhhay5

1089 Wright AE, Dean R, Zimmer F, Mank JE. 2016. How to make a sex chromosome. *Nat. Commun.* 7:1–8.

1090 Yassin A, Bastide H, Chung H, Veuille M, David JR, Pool JE. 2016. Ancient balancing selection at tan underlies female
1091 colour dimorphism in Drosophila erecta. *Nat. Commun.* 7:10400.

1092 Zajitschek F, Connallon T. 2018. Antagonistic pleiotropy in species with separate sexes, and the maintenance of
1093 genetic variation in life-history traits and fitness. *Evolution* 72:1306–1316.

1094 Zhao L, Charlesworth B. 2016. Resolving the Conflict Between Associative Overdominance and Background
1095 Selection. *Genetics* 203:1315–1334.

1096